Transcriptome-wide Development and Utilization of Novel Intron-length Polymorphic (ILP) Markers in Common Vetch (Vicia Sativa Subsp. Sativa)

Xingyi Wei (weixy16@lzu.edu.cn)
Lanzhou University https://orcid.org/0000-0002-1334-7359
Qiuxia Wang
Lanzhou University
Yitong Ma
Lanzhou University
Wenxian Liu
Lanzhou University

Research Article

Keywords: Intron-length polymorphism, Vicia sativa subsp, sativa, Molecular markers, Transferability

Posted Date: February 8th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-158950/v1

Abstract

Common vetch (*Vicia sativa* subsp*. sativa*) is one of the most economically important forage legumes with rich nutritional value and multiple uses. Though the large-scale development and application of microsatellite markers have been conducted in common vetch germplasm evaluation, the investigation and exploitation of intron-length polymorphic (ILP) markers have not been systematically analyzed. In this study, the intron positions of common vetch genome were located by aligning the RNA-Seq sequences of common vetch with *Medicago truncatula*, soybean (*Glycine max*), and *Arabidopsis thaliana* genomic sequences, and used for VsILP marker development. A total of 10,400 markers were generated from 44,582 common vetch unigenes. Out of 300 randomly selected VsILP markers, 283 resulted in successful amplification in common vetch. Among these markers, 40 produced length variation in 30 common vetch accessions, collectively yielding 166 alleles with an average of 4.0 alleles per locus. The polymorphic information content (PIC) values extended from 0.06 to 0.81 with a mean of 0.49. Of the 283 VsILP markers, 84.8% exhibited transferability to leguminous and non-leguminous species. We presented here the first large-scale development of ILP markers in common vetch and their utility in germplasm evaluation and transferability, which will be valuable for further comparative genomic studies, genetic relationship assessments, and marker-assisted breeding of leguminous and non-leguminous species.

1. Introduction

Common vetch (*Vicia sativa* subsp. *sativa*) is a self-pollinated, annual and diploid leguminous forage that can adapt to different soil and climate and fix atmospheric nitrogen through its symbiotic relationship with rhizobia to improve soil structure (Chooi 1971; Chung et al. 2013b). As an important forage legume crop, common vetch contains high level of protein, starch, and oil and has been commonly used as green manure, pasture, silage, and hay (Chung et al. 2013a; Mirzapour et al. 2013). In addition to these qualities, the common vetch also has been widely used as good quality animal feedstock and health-promoting foods for human consumption (Uzun et al. 2011). In Turkey, Australia, New Zealand, China and other regions of the world, the common vetch is widely planted and used for agricultural production (Camas and Esendal 2006; Liu et al. 2014).

The population structure and genetic diversity of germplasm are considered critical factors for the discovery of new germplasm characteristics to develop and utilize germplasm resources for plant improvement (Istvanek et al. 2017). At the same time, the studies of genetic diversity can mine new gene resources, improve existing breeding materials and reflect the breeding level of cultivated species (Gowda et al. 2013). DNA-based molecular markers, which are multi-allelic and locus-specific, have applications in marker-assisted selection breeding and genetic diversity studies. In the last decade, the development and use of DNA-based molecular markers have increased remarkably. The variety of molecular marker techniques can be divided into two types: one type is non-polymerase chain reaction (PCR)-based markers (RFLP, restriction fragment length polymorphism) (Williams et al. 1990) and the other is PCR-based markers includes simple sequence repeat polymorphism (SSR) (Tautz and Renz 1984), random amplified polymorphic DNA (RAPD) (Welsh and Mcclelland 1990), amplified fragment length

polymorphism (AFLP) (Vos et al. 1995) and single nucleotide polymorphisms (SNP). These markers have been widely applied in cultivar identification, evolution, linkage mapping, QTL mapping, and comparative genomics for various crop plants (Yang et al. 2015).

The majority of eukaryotic genes have been found to possess abundant, variable and widespread introns (Deutsch and Long 1999). As non-coding regions of a gene, introns are transcribed into mRNA but not translated since they are spliced out during pre-mRNA processing. Introns are less conserved than exon regions and accumulated a larger number of mutations, which can be exploited as genetic markers (Presgraves 2006) such as length and SNPs (Wang et al. 2014). Among the polymorphisms, intron length polymorphism (ILP) is an easily recognizable type due to its easy detection by the PCR method, namely, exon-primed intron-crossing PCR (EPIC-PCR) where primers are designed in exonic regions flanking the introns. Besides,

Compared with other previously reported DNA markers, ILP makers have a lot of advantages such as neutral, co-dominant,convenient, hypervariable, reliable and exhibiting high transferability rates between plant species (Yang et al. 2007). Due to the exon sequences are relatively more conserved, markers developed by this approach may be useful for more extensive applications than those designed in non-coding sequences. Until now, ILP molecular markers have been developed only in a few plants with genome sequences released. (Braglia et al. 2010; Chen et al. 2011; Galasso et al. 2011; Poczai et al. 2010; Shang et al. 2009; Tamura et al. 2012; Wei et al. 2015; Xia et al. 2012), such as rice (Wang et al. 2005), soybean (Shu et al. 2010), *Dasypyrum villosum* (Zhang et al. 2017a), and *Daucus carota* (Stelmach et al. 2017). For most species lacking genome-wide data, the characteristics of the exon-intron structure in homologous genes of different species can be used to infer the intron position of cDNA/EST based on homologous genes from related model organisms (Yang et al. 2007). Recently, with the rapid development of high-throughput transcriptome sequencing, collecting large numbers of nucleotide sequencing reads at the transcription level and having thus made the development of ILP markers more cost-effective and easier (Dong et al. 2017; Liu et al. 2014). For example, 502 ILP markers were successfully developed in *Medicago sativa* based on the alfalfa unigene sequences (Zhang et al. 2017b).

Due to the lack of genome-wide data, the current research on the relationship between species and genetic diversity of common vetch is concentrated on the morphological level (van de Wouw et al. 2003), and the related research at the molecular level remains limited. With the development of high-throughput transcriptome sequencing, using transcriptome data to develop the molecular markers has been made possible in common vetch. In the work described here, we took a large-scale search and developed a set of ILP markers based on the available transcriptome data of common vetch, and the genetic diversity assessments and potential for cross-species transferability of these markers in common vetch accessions were further analyzed.

2. Materials And Methods

2.1 Plant material and DNA extraction

In this study, The leguminous species common vetch (LANJIAN NO.3), barrel medic (*M. truncatula* A17), alfalfa (*M. sativa* cultivar ARC), soybean (*Glycine max* cultivar Dongdou 641), crowtoe (*Lotus corniculatus* cv. Mirabal), *Sophora alopecuroides* (wild material) and yellow sweet clover (wild material), and the non-leguminous species rice (*Oryza sativa* cv. Kitaake), *Arabidopsis* 'Columbia' and tobacco (*Nicotiana tabacum* cv. Samsun NN) were used to examine the transferability of the ILP markers developed. Genomic DNA was extracted from the leaf materials of field plants as described by Yan et al (Yan et al. 2017). A total of 30 common vetch accessions (Table S1) were collected from the National Animal Husbandry Station of the Ministry of Agriculture, the United States Department of Agriculture National Plant Germplasm System (NPGS) and Rural China and the National Grass Germplasm Resource Bank, respectively. The young leaves of 10 individual plants from each accession were used for genomic DNA isolation and polymorphism investigation. The quantity of DNA was determined at 260 nm using a NanoDrop ND1000 (Thermo Fisher Scientific, Wilmington, DE), and the quality was checked on 2% agarose gel. The DNA was normalized to 50 ng/μl for further use.

2.2 Sequence retrieval and primer design

The transcripts and coding DNA sequences (CDS) data of common vetch were downloaded from NCBI Gene Expression Omnibus with accession No. GSE35437 (Liu et al. 2014) and used for the development of the specific intron-based markers. These sequences were aligned with *Arabidopsis thaliana*, *M. truncatula* and soybean (*Glycine max*) genomic sequences using the method described by Yang (Yang et al. 2007) with Perl scripts to predict the intron positions, and the primers flanking both sides of the intron regions were then designed using Array Designer 4.2 software (http://www.softpedia.com/get/Science-CAD/Array-Designer.shtml).

2.3 ILP marker analysis

PCR reactions were performed in a 10-μl volume containing 5 μl 2×Power *Pfu* PCR Master Mix (Bioteke Corporation, Beijing, China), 50 ng template DNA, and 10 ng each of the forward and reverse primers. The touchdown PCR procedure was performed as follows: five cycles of denaturation 95°C for 30 s; the annealing temperature was then decreased by 1°C for each cycle from 60 − 56°C to 55 − 50°C; 72°C for 30 s; 30 cycles of 95°C for 30 s, 56 − 50°C for 30 s, and 72°C for 30 s; and a final extension at 72°C for 7 min. The PCR products were separated by 6% denaturing PAGE (220 V, 90 min) and visualized by silver staining.

The electrophoretic bands produced by the marker VsILP233 in seven leguminous and three non-leguminous species were isolated and sequenced (by the Shanghai Sangon Biotech Company) to examine whether the PCR products amplified the homologous target gene. The ClustalW2 program (http://www.ebi.ac.uk/Tools/msa/clustalw2/) was employed to align the multiple sequences.

2.4 Statistical analysis

Scoring of bands was done as presence (1) or absence (0) for each locus. The allelic diversity of the common vetch ILP markers was evaluated by polymorphism information content (PIC) value, defined as:

$$PIC_i = 1 - \sum_j P_{ij}^2$$

where P_{ij} is the frequency of the j^{th} pattern for the th marker (Anderson et al. 1993). Genetic cluster analysis was performed with NTSYS-pc version 2.10s(Jensen 1989). The data was imported from a binary matrix in Microsoft Excel files to calculate the similarity matrix with the SM coefficient. The similarity matrix was then subjected to the UPGMA method to cluster the relationship of the genotypes and generate the dendrogram. Bayesian model-based software STRUCTURE 2.3.2 was employed to infer the genetic structure without information on the accession origin (Pritchard et al. 2000). The length of the burn-in period and the number of Markov Chain Monte Carlo (MCMC) replications were assigned at 10^5 for each number of clusters (K) set from 1 to 20 for further subclustering after the burn-in.

3. Results

3.1 Intron identification and analysis

Due to the successful implementation of common vetch transcriptome sequencing, a total of 44,582 high-quality unigenes sequences which represented 17% of the predicted 2,025 Mb common vetch genome were available and used in this study. The average sequence size of these common vetch unigenes was 775 bp, and the largest spanned 7,907 bp (Liu et al. 2014). Among them, lengths greater than 200 bp were selected and aligned with the genomic sequences of soybean, *M. truncatula*, and *Arabidopsis*. When at least 200 bp overlapped with 80% similarity, the common vetch unigene sequences were deemed to be homologous to the CDS of *Arabidopsis*, soybean and *M. truncatula*. With the result of a homology search between common vetch unigenes and *Arabidopsis* genome, 76 unigene sequences existed at/on one or more intron insertion sites. Furthermore, a total of 116 introns presented in *Arabidopsis*, with an average of 1.52 introns per unigene sequence. In the case of soybean, 3,446 unigenes with 6,345 introns were identified with an average of 1.84 introns per unigene sequence. In contrast, with the *M. truncatula* genome, a total number of 4,721 unigenes with 8,490 introns were identified with an average of 1.80 introns per unigene sequence. The minimum number of introns presented in single unigene was 13 in the case of *Arabidopsis*, 20 in soybean and 18 in *M. truncatula*. For the introns size comparison, all the unigenes which had at least one intron present were selected. In *Arabidopsis*, 82.8 % of introns were found between 50 bp to 200 bp (Fig. 1). The ratio of soybean and *M. truncatula* were 60.19 % and 59.48 %, respectively. In addition, except for one intron in *M. truncatula*, all the intron sizes were larger than 50 bp.

3.2 Functional classification of VsILPs

To infer the putative gene function of the VsILP-targeted unigenes, we utilized the gene ontology (GO) assignments from *M. truncatula* gene models to deduce the putative functions for the 5,534 unigenes

which were considered potential introns and functionally classified using WEGO 2.0 software (Ye et al. 2018). The categorization of these unigenes was completely summarized in Fig. 2 according to biological process, cellular component, and molecular function. A large number of the unigenes were assigned to GO categories and many of them were found in several categories. As results showed in the biological process category, cellular process (3,531 genes, 63.8%) was the most abundant, followed by single-organism process (3,151 genes, 56.9%) and biological regulation (2,969 genes, 48.7%). In the cellular component category, antioxidant activity (4,019 genes, 72.6%) and cell part (4,013 genes, 72.5%) were the two abundantly-assigned GO term, followed by membrane (3,328 genes,) and organelle (3,310 genes, 59.8%). The two most over-represented GO terms were binding (3,177 genes, 57.4%) and catalytic activity (3,052 genes, 55.7%), followed by transporter activity (464 genes, 8.4%).

3.3 Genetic diversity of ILP markers

A total of 10,400 markers were generated from 44,582 unigene sequences. For more brevity to assess the potential ILP markers, 300 markers were chosen randomly and used for PCR primers development. Of the 300 primers, 283 (94.3 %) amplified marker fragment successfully in four common vetch accessions which came from different countries or regions. All markers of this test were named VsILP and the details were provided in Supplementary material (Table S1).

To further evaluate the polymorphism and molecular diversity potential of these markers, 78 VsILP primer pairs were randomly selected from the 283 markers with successful amplicons in the common vetch as described above and used to examine in 30 common vetch accessions including twenty foreign species and ten domestic species (Table S2). Of these markers, 41 markers showed polymorphisms in all common vetch species (Fig. 3), with a total of 166 alleles collectively yielded (with an average of 4.0 alleles per locus, varying from two to seven), presenting a polymorphic rate of 52.56% (Table S3). In order to examine the extent of the information on diversity, the polymorphic information content (PIC) values were calculated and found that extended from 0.06 to 0.81 with a mean of 0.49. Besides, the expected heterozygosity (He) ranged from 0.06 in primer pairs VsILP176 to 0.83 in primer pairs VsILP160, with an average of 0.54 (Table S3). Based on the data of these 41 VsILP markers, the UPGMA dendrogram of the 30 common vetch accessions was constructed by the software NTSYS-pc. The clustering data clearly showed that these 30 common vetch accessions could be clustered into five major groups with the Jaccard’s similarity coefficient ranging from 0.75 to 0.94 (Fig. 4). The first cluster contained five accessions from the former Soviet Union and accession“CAROLE” from France. Cluster Ⅱ contained accessions “BLANCHEGRAIN” and “CHARKOVSZKAJA N.134”. Nine accessions of Cluster III were originated from Europe, Northern America, and East Asia, respectively. Cluster Ⅳ contained four accessions from France and Northern America, respectively. Cluster V were all originated from China with nine accessions. The accessions “LANJIAN NO.3" and “STRZELECHE RDZOWA” displayed less similarity with other clusters and were thus considered ungrouped. In contrast, the accessions “SUJIAN NO.3(HUIYIN)” and “SUJIAN NO.3(NING)” had the least distance among all the accessions (Fig. 4).

In order to further investigate the evolutionary relationship among the test varieties, STRUCTURE analysis was performed to evaluate the genetic structure of the 30 accessions (Table S1). The value of ΔK statistics was the highest when two clusters were assumed [$\Delta K_{(2)} = 61.64$]. The ΔK value gradually decreases with the increase in the number of assumed clusters [$\Delta K_{(>2)} = 0.11–10.32$]. Twenty cultivated accessions were assigned to cluster 1 with membership coefficients (Q) ranging between 0.533 and 0.994, whereas cluster comprised exclusively wild accessions with the Q value of 0.602–0.998 (Fig. 5).

3.4 Cross-species transferability of VsILP markers

To assess the cross-species transferability potential of the VsILP markers, all the 283 VsILP makers which successfully amplified one or more marker fragments in the two common vetch accessions as described above were tested in seven leguminous (common vetch, alfalfa, barrel medic, soybean, yellow sweet clover, crowtoe and *Sophora alopecuroides*) and three non-leguminous species (rice, *Arabidopsis* and tobacco) (Table 1). The lowest amplification percentage (20.5%) was observed in rice, while barrel medic had the highest amplification percentage (67.1%) and the average was 48.3%. Of these markers, 240 (84.8 %) were found to be transferable to at least one of other nine species except common vetch (the amplification percentage was 100%), 66 (23.3 %) produced amplifications in all of the legume species, and 21 (7.4%) makers produced amplifications in all ten species. Also, the means of transferability of these VsILP markers in leguminous species was appreciably higher than that in non-leguminous, (59.6%) and (21.9%) respectively (Table 1). All the 283 markers showing transferability collectively yielded 786 alleles which clearly separated the leguminous and non-leguminous species into two distinct groups (Fig. 6, Table S4).

Table 1
Percent transferability of the 283 VsILP markers to the different leguminous and non-leguminous species.

NO.	Investigated crop	Transferability
1	*Vicia sativa*	100.0%
2	*Medicago truncatula*	67.1%
3	*Medicago sativa*	65.0%
4	*Melilotus officinalis*	51.6%
5	*Lotus corniculatus*	48.8%
6	*Glycine max*	47.3%
7	*Sophora alopecuroides*	37.4%
8	*Nicotiana tabacum*	23.3%
9	*Arabidopsis*	21.9%
10	*Oryza sativa*	20.5%
Average		48.3%

To detect the sequence variation of VsILP-tagged cross-species polymorphism amplification products, the sequences of the cloned product amplified by the marker VsILP233 in the leguminous and non-leguminous species were isolated and sequenced (Fig. 7). The multiple sequence alignment showed that the intron region in the middle contained large differences, including length variations and point mutations, but high conservation among all the species was observed in the exon regions at both ends of the position (Fig. 8).

4. Discussion

Molecular marker development in forage species is crucial in the facilitation of genomics-based crop improvement. Currently, the intronic regions of the genes are targeted for the development of PCR-based markers since exons are more prone to accumulate more mutations than introns during gene evolution (Andrade-Navarro 2009). Accurate prediction of introns from genomic DNA is critical for the development of intron length polymorphism molecular markers in a species without genomic DNA sequence information. Owing to the limit of genomic data, the development of molecular markers in common vetch is extremely scarce. In the present study, a set of ILP markers have been developed based on the transcriptome data of common vetch, and the polymorphism and transferability of these makers were further evaluated.

The homologous exon among species is conserved, whereas the intron length and number varies dramatically among plants(Guan et al. 2016). In this study, large differences were exhibited in the number

of introns which were identified with reference to *M. truncatula* (8,490), soybean (4,694) and *Arabidopsis* (81) genome, respectively. It also showed that common vetch were closer phylogenetic relatives to *M. truncatula* than to the other two species. In regards to the functional clustering, large numbers of common vetch unigenes were functionally assigned by performing GO analyses, and the remaining unigenes failed to be assigned specific functional annotation maybe they were too short to be matched to known genes during the identification of significant sequence similarity based on query sequence length (Lu et al. 2011).

A total number of 10,400 ILP markers from orthologs were developed. The scale of working markers developed through PCR was about 94.3% (283/300), which helped in detecting 14.5% (41/283) polymorphic loci among 30 common vetch genotypes. The amplification rate of success was higher than those in other plants, such as alfalfa (82%) (Zhang et al. 2017b), cotton (72.71% and 36.45%) (Cai et al. 2017) and sorghum (86%) (Jaikishan et al. 2016), but lower than that in *Dasypyrum villosum* (97.84%) (Zhang et al. 2017a) and carrot (97.1%) (Stelmach et al. 2017), probably due to the length of introns in common vetch was too long to amplify successfully. Large product size is considered the main cause of PCR failure and generally, the successful amplification rate decreases with greater length of the intron (Wang et al. 2010). The PIC values of these VsILP markers ranged from 0.06 to 0.81 with a mean of 0.49, which was comparatively higher than that reported in cowpea (0.34) (Gupta et al. 2012) and foxtail millet (0.20) (Muthamilarasan et al. 2014), but lower than alfalfa (0.60) (Zhang et al. 2017b). These results suggest that the highly polymorphic VsILP markers can be used for molecular-assisted breeding and genetic map construction. Besides, thirty common vetch accessions were grouped into five main clusters based on UPGMA and Bayesian clustering analysis with Jaccard's similarity coefficient ranging from 0.75 to 0.94. The results indicated that most accessions from China (which belongs to the cluster V) may share a more similar genetic background (Fig. 4). Moreover, the clustering patterns of the other geographically closer accessions were unclear. Similar results were concordant with that using cDNA-SSR markers (Chung et al. 2013a). Therefore, more accessions from each geographical location were necessary in order to conclusively confirm the appropriate clustering pattern.

To further check the transferability of these newly developed VsILP markers across other species, two hundred and eighty-three ILP markers were selected and tested in ten species (seven leguminous and three non-leguminous). A total of 84.8% of VsILP markers were found to produce amplified stable products in other species. The results showed that the exons in common vetch species were highly conserved which was consistent with the earlier ILP markers in alfalfa and rice (Wang et al. 2005; Zhang et al. 2017b). This was further confirmed by the sequencing results of the cloned PCR product (Fig. 8). Moreover, due to the high cross-species transferability, these VsILP markers might be very helpful in characterizing species relationships and comparative genomics studies of multiple species.

Declarations

Acknowledgments

This research was supported by the National Natural Science Foundation of China (31502000), and the Special Funding for Open and Shared Large-Scale Instruments and Equipments of Lanzhou University.

Conflict of interest

The authors declare that they have no conflict of interest.

Funding:

This research was supported by the National Natural Science Foundation of China (31502000).

Conflicts of interest/Competing interests:

The authors declare that they have no conflict of interest.

Availability of data and material:

The transcripts and coding DNA sequences (CDS) data of common vetch were downloaded from NCBI Gene Expression Omnibus with accession No. GSE35437.

Code availability

All the software in this article are available on the Internet, and the perl script used cannot be published temporarily.

Authors' contributions:

W.-X.L. designed the study. X.-Y.W., Q.-X.W. and Y.-T.M. performed experiments. X.-Y.W., Q.-X.W. and Y.-T.M. analyzed the data. X.-Y.W. and Q.-X.W. wrote the manuscript. W.-X.L. provided fund for this study. All authors read and approved the final version.

Ethics approval Declarations.

Consent to participate Declarations.

Consent for publication Declarations.

References

1. 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
2. Andrade-Navarro MA (2009) The Dictionary of Genomics, Transcriptomics, and Proteomics BioEssays 31:1367-1369 doi:10.1002/bies.200900121
3. Braglia L, Manca A, Mastromauro F, Breviario D (2010) cTBP: A Successful Intron Length Polymorphism (ILP)-Based Genotyping Method Targeted to Well Defined Experimental Needs Diversity 2:572-585 doi:10.3390/d2040572
4. Cai C, Wu S, Niu E, Cheng C, Guo W (2017) Identification of genes related to salt stress tolerance using intron-length polymorphic markers, association mapping and virus-induced gene silencing in cotton Sci Rep 7:528 doi:10.1038/s41598-017-00617-7
5. Camas N, Esendal E (2006) Estimates of broad-sense heritability for seed yield and yield components of safflower (Carthamus tinctorius L.) Hereditas 143:55-57 doi:10.1111/j.2006.0018-0661.01914.x
6. Chen X, Zhang G, Wu WR (2011) Investigation and utilization of intron length polymorphisms in conifers New Forest 41:379-388 doi:10.1007/s11056-010-9229-5
7. Chooi WY (1971) Variation in nuclear DNA content in the genus vicia Genetics 68:195-211
8. Chung JW, Kim TS, Suresh S, Lee SY, Cho GT (2013a) Development of 65 Novel Polymorphic cDNA-SSR Markers in Common Vetch (Vicia sativa subsp sativa) Using Next Generation Sequencing Molecules 18:8376-8392 doi:10.3390/molecules18078376
9. Chung JW, Kim TS, Suresh S, Lee SY, Cho GT (2013b) Development of 65 novel polymorphic cDNA-SSR markers in common vetch (Vicia sativa subsp. sativa) using next generation sequencing Molecules 18:8376-8392 doi:10.3390/molecules18078376
10. Deutsch M, Long M (1999) Intron-exon structures of eukaryotic model organisms Nucleic Acids Res 27:3219-3228 doi:DOI 10.1093/nar/27.15.3219
11. Dong R et al. (2017) Transcriptome Analyses Reveal Candidate Pod Shattering-Associated Genes Involved in the Pod Ventral Sutures of Common Vetch (Vicia sativa L.) Front Plant Sci 8:649 doi:10.3389/fpls.2017.00649
12. Galasso I, Manca A, Braglia L, Martinelli T, Morello L, Breviario D (2011) h-TBP: an approach based on intron-length polymorphism for the rapid isolation and characterization of the multiple members of the beta-tubulin gene family in Camelina sativa (L.) Crantz Molecular Breeding 28:635-645 doi:10.1007/s11032-010-9515-0
13. Gowda CLL, Upadhyaya HD, Sharma S, Varshney RK, Dwivedi SL (2013) Exploiting Genomic Resources for Efficient Conservation and Use of Chickpea, Groundnut, and Pigeonpea Collections for Crop Improvement Plant Genome-Us 6 doi:10.3835/plantgenome2013.05.0016
14. Guan R et al. (2016) Draft genome of the living fossil Ginkgo biloba Gigascience 5:49 doi:10.1186/s13742-016-0154-1
15. Gupta SK, Bansal R, Gopalakrishna T (2012) Development of intron length polymorphism markers in cowpea [Vigna unguiculata (L.) Walp.] and their transferability to other Vigna species Molecular

Breeding 30:1363-1370 doi:10.1007/s11032-012-9722-y

16. Istvanek J, Dluhosova J, Dluhos P, Patkova L, Nedelnik J, Repkova J (2017) Gene Classification and Mining of Molecular Markers Useful in Red Clover (Trifolium pratense) Breeding Front Plant Sci 8:367 doi:10.3389/fpls.2017.00367
17. Jaikishan I, Rajendrakumar P, Madhusudhana R, Elangovan M, Patil JV (2016) Development and utility of PCR-based intron polymorphism markers in sorghum [Sorghum bicolor (L.) Moench] Journal of Crop Science and Biotechnology 18:309-318 doi:10.1007/s12892-015-0015-y
18. Jensen RJ (1989) Ntsys-Pc - Numerical Taxonomy and Multivariate-Analysis System - Version 1.40 Q Rev Biol 64:250-252 doi:Doi 10.1086/416356
19. Liu Z, Liu P, Luo D, Liu W, Wang Y (2014) Exploiting Illumina sequencing for the development of 95 novel polymorphic EST-SSR markers in common vetch (Vicia sativa subsp. sativa) Molecules 19:5777-5789 doi:10.3390/molecules19055777
20. Lu FH et al. (2011) Transcriptome analysis and SNP/SSR marker information of red pepper variety YCM334 and Taean Sci Hortic-Amsterdam 129:38-45 doi:10.1016/j.scienta.2011.03.003
21. Mirzapour M, Nofouzi F, Mokhtarzadeh S, Kendir H, Khawar KM (2013) Effects of BAP-NAA on plant regeneration from half cotyledon explants of Turkish Common Vetch (Vicia sativa L.) cultivar Kubilay Current Opinion in Biotechnology 24:S121-S121 doi:10.1016/j.copbio.2013.05.381
22. Muthamilarasan M, Venkata Suresh B, Pandey G, Kumari K, Parida SK, Prasad M (2014) Development of 5123 intron-length polymorphic markers for large-scale genotyping applications in foxtail millet DNA Res 21:41-52 doi:10.1093/dnares/dst039
23. Poczai P, Cernak I, Gorji AM, Nagy S, Taller J, Polgar Z (2010) Development of intron targeting (IT) markers for potato and cross-species amplification in Solanum nigrum (Solanaceae) Am J Bot 97:e142-145 doi:10.3732/ajb.1000360
24. Presgraves DC (2006) Intron length evolution in Drosophila Mol Biol Evol 23:2203-2213 doi:10.1093/molbev/msl094
25. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data Genetics 155:945-959
26. Shang W, Zhou R, Jia J, Gao L (2009) RGA-ILP, a new type of functional molecular markers in bread wheat Euphytica 172:263-273 doi:10.1007/s10681-009-0063-9
27. Shu YJ et al. (2010) Genome-wide identification of intron fragment insertion mutations and their potential use as SCAR molecular markers in the soybean Theoretical and Applied Genetics 121:1-8 doi:10.1007/s00122-010-1285-x
28. Stelmach K, Macko-Podgorni A, Machaj G, Grzebelus D (2017) Miniature Inverted Repeat Transposable Element Insertions Provide a Source of Intron Length Polymorphism Markers in the Carrot (Daucus carota L.) Front Plant Sci 8:725 doi:10.3389/fpls.2017.00725
29. Tamura K, Kiyoshi T, Yonemaru J (2012) The development of highly transferable intron-spanning markers for temperate forage grasses Molecular Breeding 30:1-8 doi:10.1007/s11032-011-9593-7

30. Tautz D, Renz M (1984) Simple Sequences Are Ubiquitous Repetitive Components of Eukaryotic Genomes Nucleic Acids Res 12:4127-4138 doi:DOI 10.1093/nar/12.10.4127
31. Uzun A, Gucer S, Acikgoz E (2011) Common vetch (Vicia sativa L.) germplasm: correlations of crude protein and mineral content to seed traits Plant Foods Hum Nutr 66:254-260 doi:10.1007/s11130-011-0239-z
32. van de Wouw M, Maxted N, Ford-Lloyd BV (2003) Agro-morphological characterisation of common vetch and its close relatives Euphytica 130:281-292 doi:Doi 10.1023/A:1022899410696
33. Vos P et al. (1995) Aflp - a New Technique for DNA-Fingerprinting Nucleic Acids Res 23:4407-4414 doi:DOI 10.1093/nar/23.21.4407
34. Wang H, Devos KM, Bennetzen JL (2014) Recurrent loss of specific introns during angiosperm evolution PLoS Genet 10:e1004843 doi:10.1371/journal.pgen.1004843
35. 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
36. Wang YY, Chen J, Francis DM, Shen HL, Wu TT, Yang WC (2010) Discovery of intron polymorphisms in cultivated tomato using both tomato and Arabidopsis genomic information Theoretical and Applied Genetics 121:1199-1207 doi:10.1007/s00122-010-1381-y
37. Wei S, Peng Z, Yang Z (2015) A simple approach based on intron length polymorphism (ILP) for chromosomal localization of GeneWAG-2 Indian Journal of Genetics and Plant Breeding (The) 75:314 doi:10.5958/0975-6906.2015.00049.8
38. Welsh J, Mcclelland M (1990) Fingerprinting Genomes Using Pcr with Arbitrary Primers Nucleic Acids Res 18:7213-7218 doi:DOI 10.1093/nar/18.24.7213
39. 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:DOI 10.1093/nar/18.22.6531
40. Xia S et al. (2012) Mapping of BnMs4 and BnRf to a common microsyntenic region of Arabidopsis thaliana chromosome 3 using intron polymorphism markers Theor Appl Genet 124:1193-1200 doi:10.1007/s00122-011-1779-1
41. Yan Z et al. (2017) Cross-species transferability of EST-SSR markers developed from the transcriptome of Melilotus and their application to population genetics research Sci Rep 7:17959 doi:10.1038/s41598-017-18049-8
42. Yang H, Li C, Lam HM, Clements J, Yan G, Zhao S (2015) Sequencing consolidates molecular markers with plant breeding practice Theor Appl Genet 128:779-795 doi:10.1007/s00122-015-2499-8
43. 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
44. Ye J et al. (2018) WEGO 2.0: a web tool for analyzing and plotting GO annotations, 2018 update Nucleic Acids Res 46:W71-W75 doi:10.1093/nar/gky400

45. Zhang X et al. (2017a) Whole genome development of intron targeting (IT) markers specific for Dasypyrum villosum chromosomes based on next-generation sequencing technology Molecular Breeding 37 doi:10.1007/s11032-017-0710-0
46. Zhang Z, Min X, Wang Z, Wang Y, Liu Z, Liu W (2017b) Genome-wide development and utilization of novel intron-length polymorphic (ILP) markers in Medicago sativa Molecular Breeding 37 doi:10.1007/s11032-017-0659-z

Figures

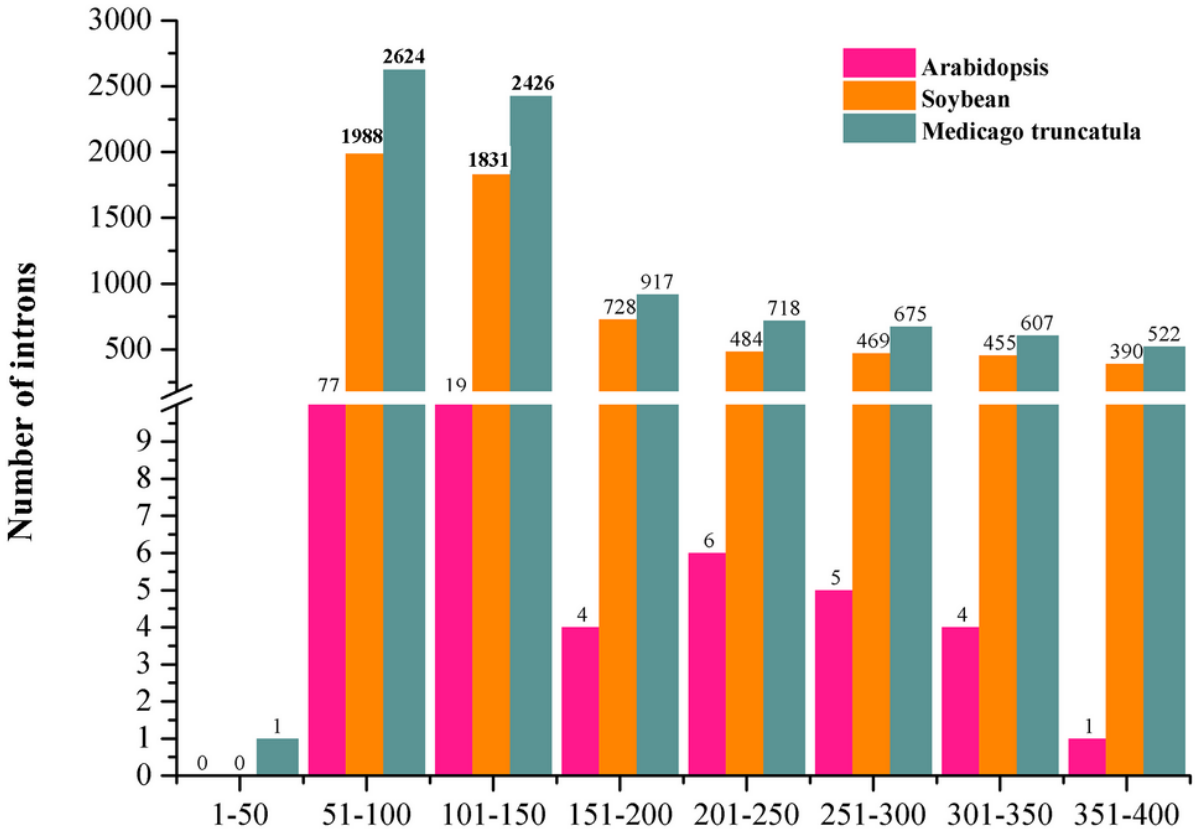

Figure 1

Distribution of intron sizes in Arabidopsis and soybean based on common vetch unigenes.

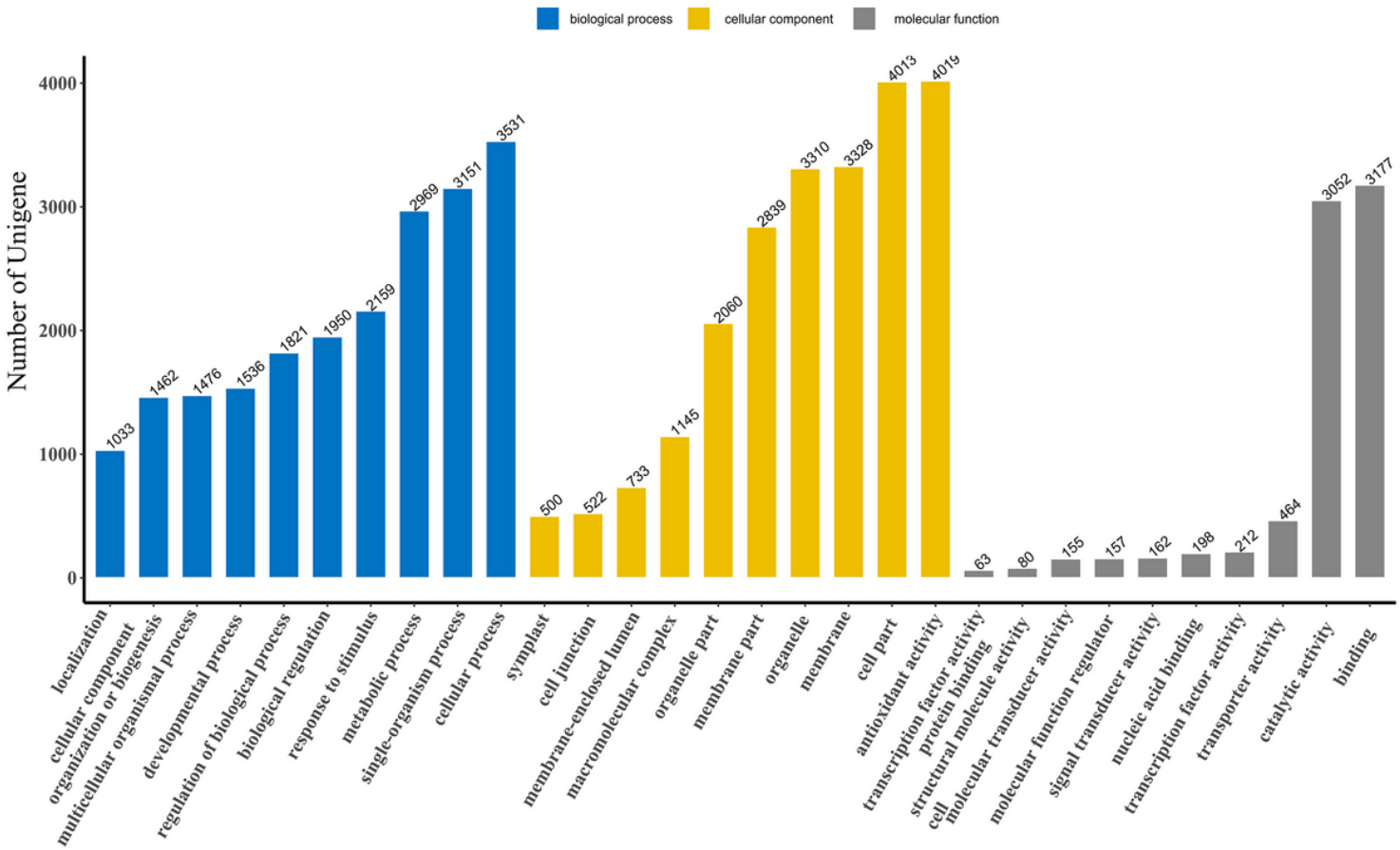

Figure 2

Functional classification of the unigenes.

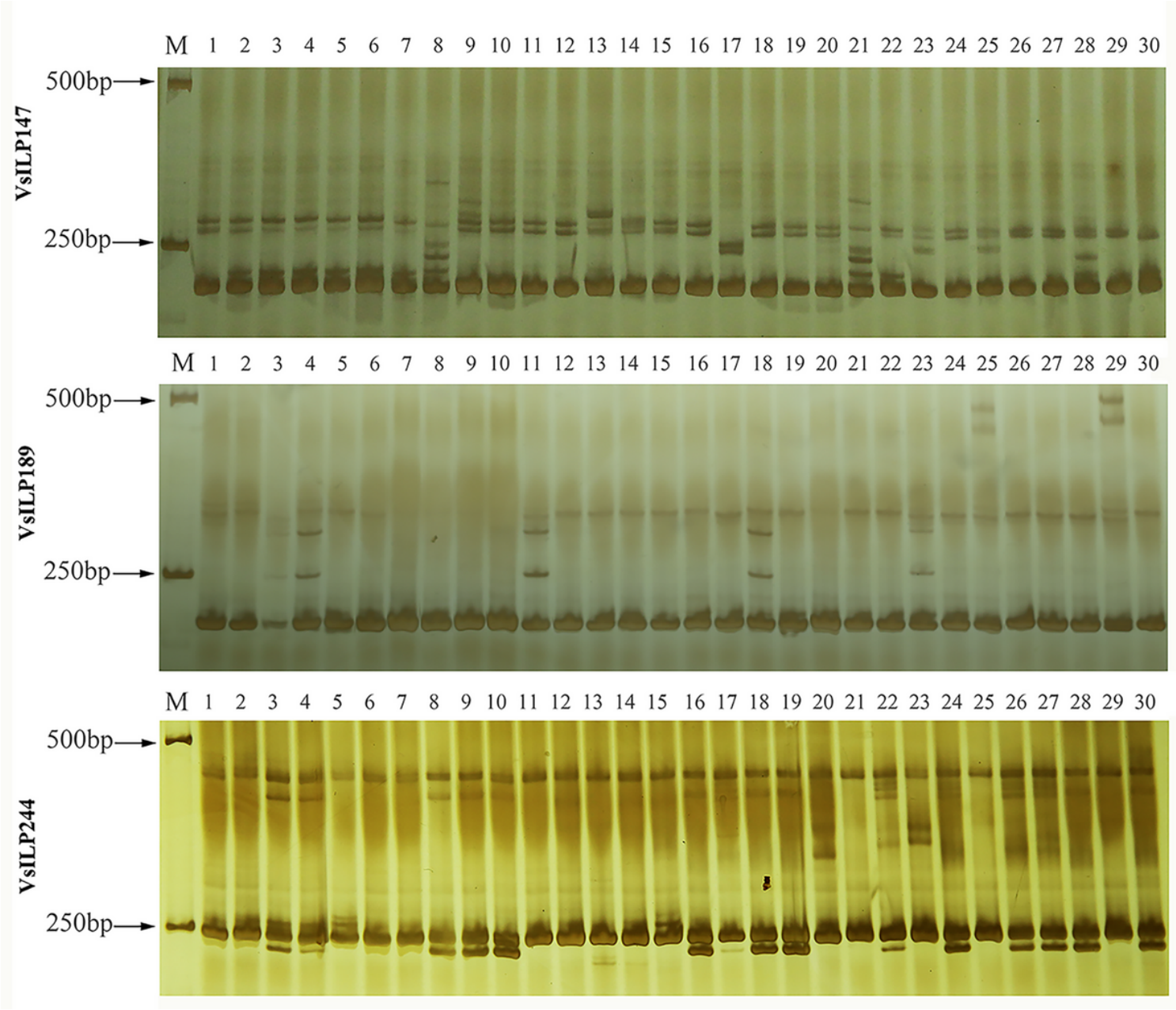

Figure 3

Polyacrylamide gel electrophoresis of the intron length polymorphic (ILP) markers. A, VsILP147; B, VsILP189; C, VsILP244. Lane order is (left to right) DNA marker DL2000, Vicia sativa; Medicago truncatula; Medicago sativa; Glycine max; Lotus corniculatus; Melilotus officinalis; Sophora alopecuroides; Arabidopsis; Oryza sativa; and Nicotiana tabacum.

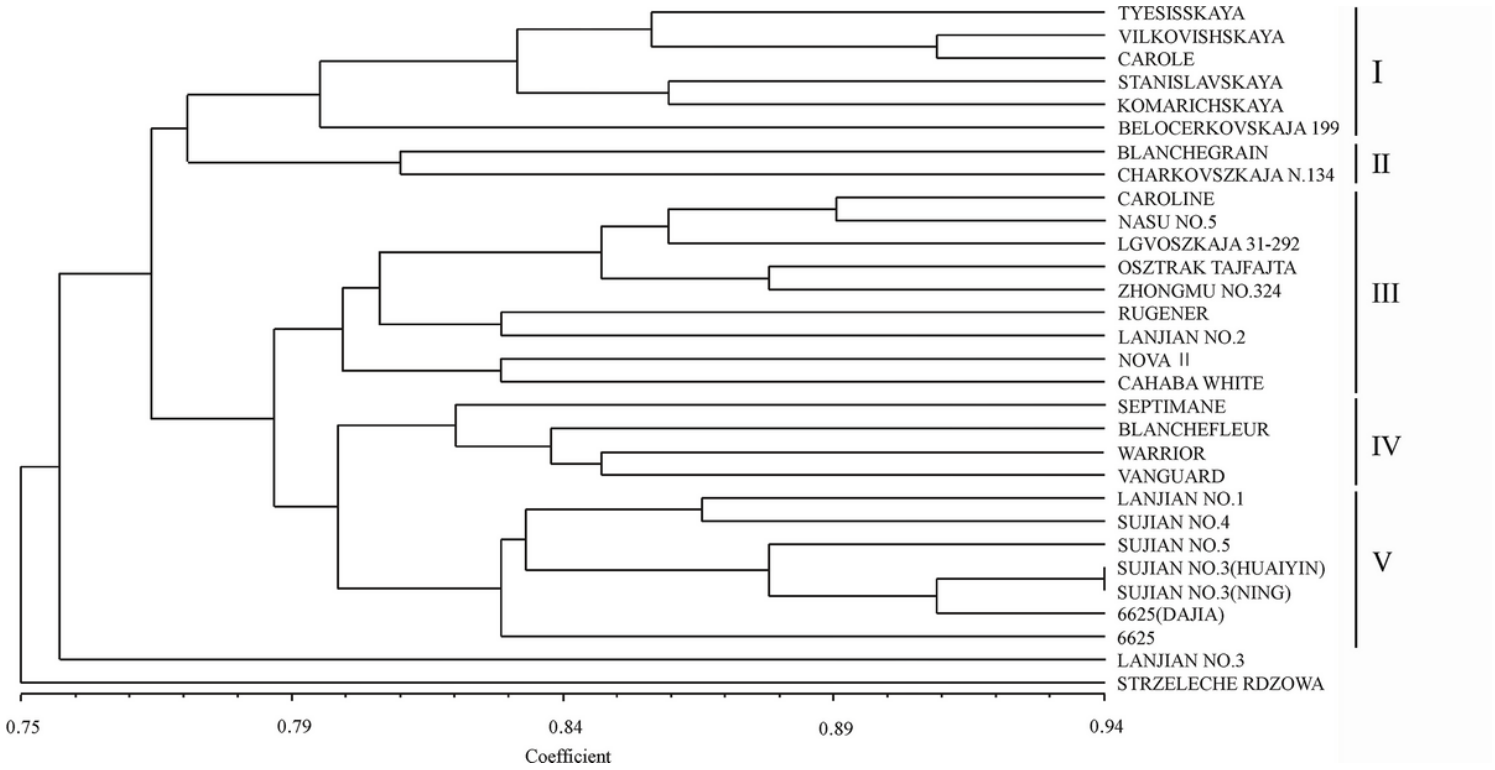

Figure 4

The dendrogram of 21 common vetch accessions based on UPGMA analysis using 41 polymorphic VsILP markers.

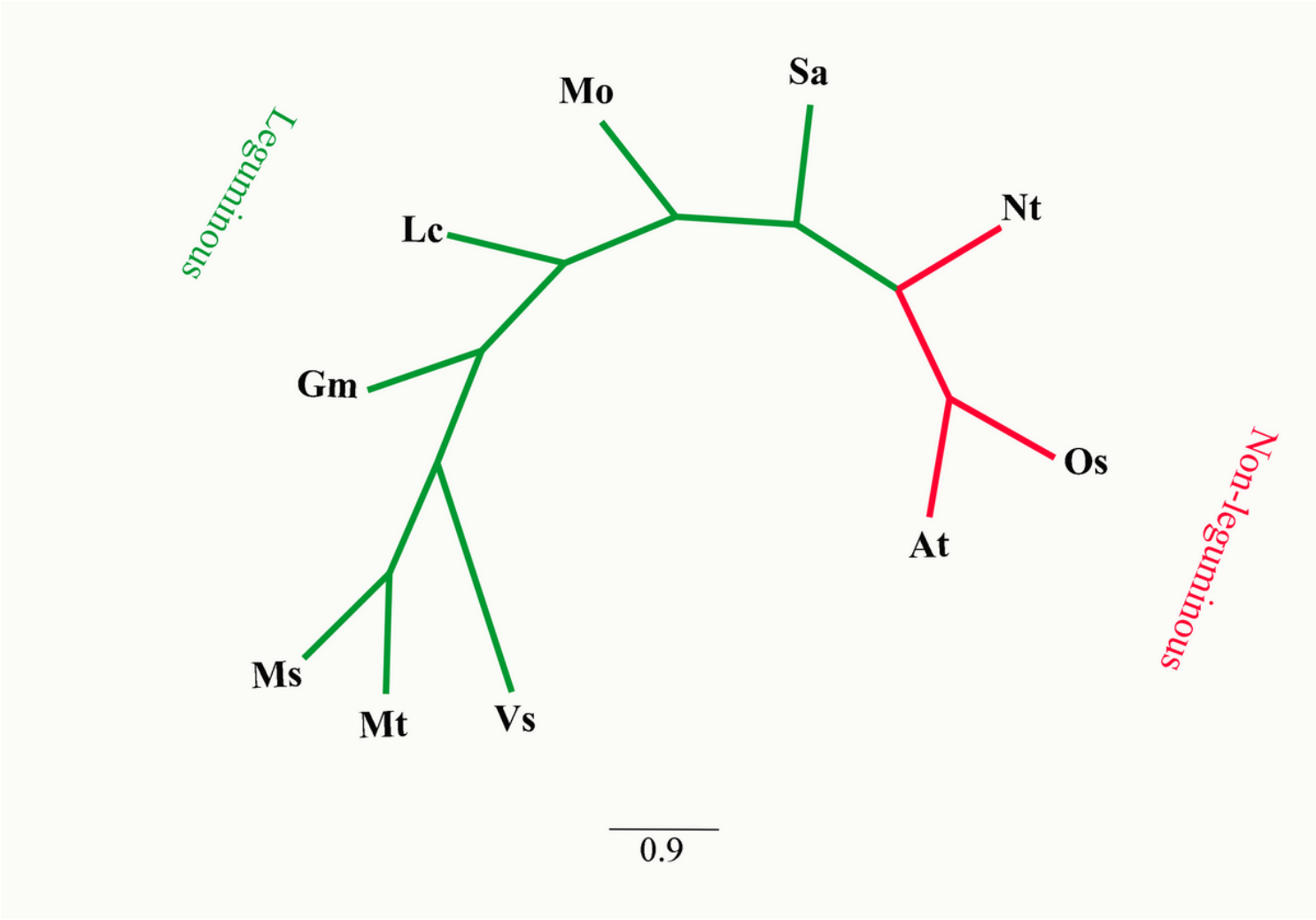

Figure 6

Genetic relationships between the leguminous and non-leguminous species based on 283 VsILP markers using NJoin clustering.

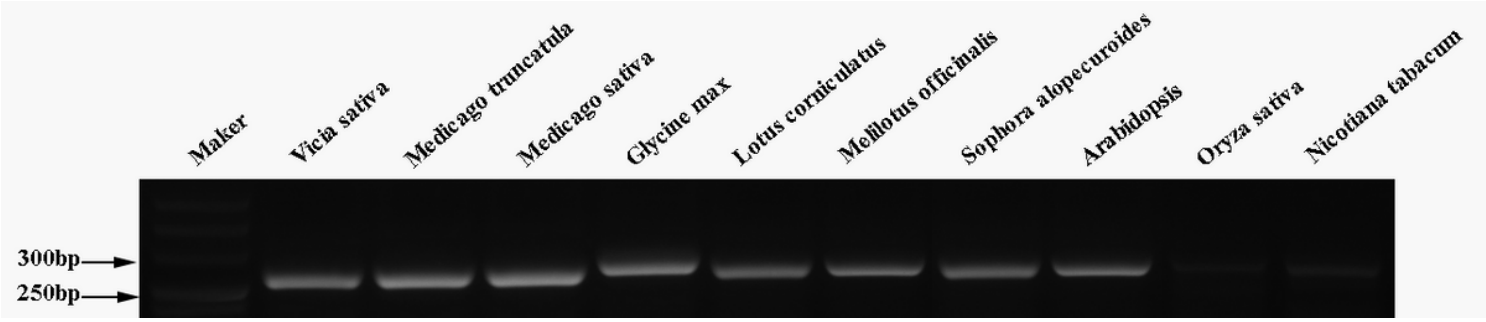

Figure 7

Amplification of VsILP233 in 10 plant species.

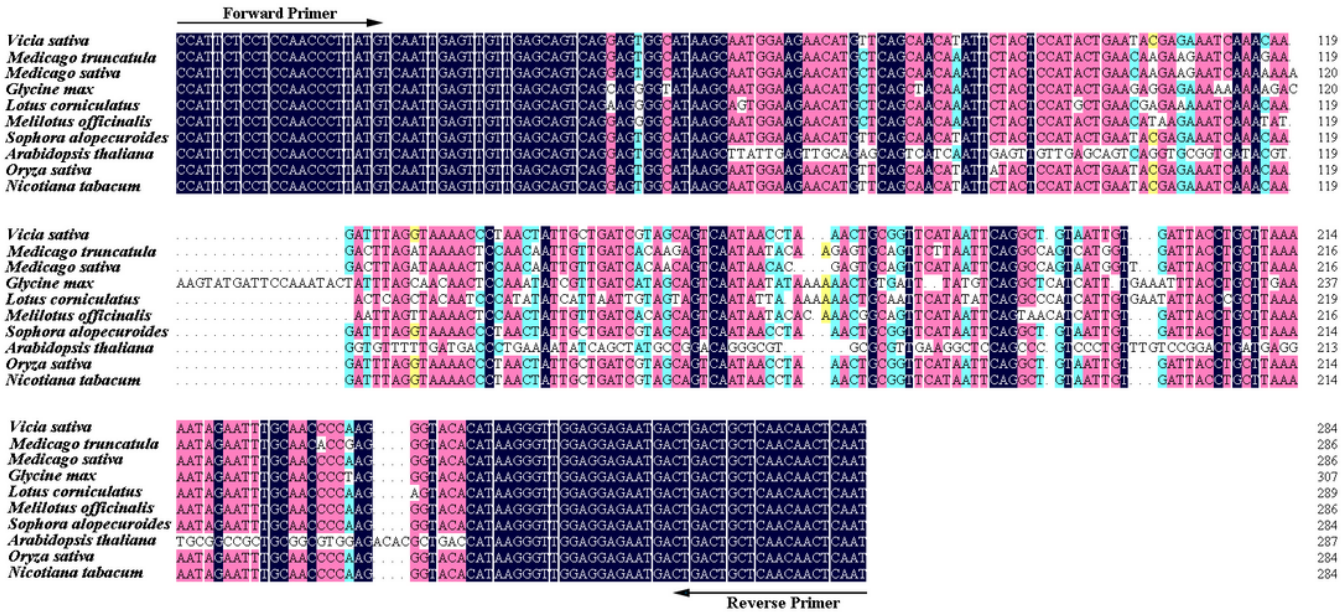

Figure 8

Multiple alignment of sequences amplified from 10 plant species by primer pair VsILP233. The asterisks denote similar sequences, and the points represent deletions.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.docx
- TableS2.docx
- TableS3.docx
- TableS4.doc