Plant Omics Journal POJ

POJ 6(4):295-301 (2013) ISSN:1836-3644

Development of EST-SSR markers to study genetic diversity in hyacinth bean (*Lablab purpureus* L.)

Guwen Zhang[1+], Shengchun Xu[1+], Weihua Mao[2], Yaming Gong[1*], Qizan Hu[1]

[1]Institute of Vegetables, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, People's Republic of China
[2]Center of Analysis and Measurement, Zhejiang University, Hangzhou 310023, People's Republic of China
[+]These authors are equally contributed to this work

***Corresponding author: gongym07@126.com**

Abstract

The development of expressed sequence tags (ESTs) has provided a useful source for the mining of simple sequence repeat (SSR) markers. The hyacinth bean (*Lablab purpureus* L.) is one of the most ancient and important legume crops in the world, and is widely distributed in tropical and sub-tropical regions, but currently only limited molecular markers are available. The aim of the present study is to develop EST-SSRs for hyacinth beans to investigate the genetic structure and diversity of different populations originating from China and Africa. A total of 459 hyacinth bean ESTs from the National Center for Biotechnology Information (NCBI) database were downloaded and analyzed to search for SSRs with a minimum of 12 repeating nucleotides. Finally, 22 microsatellites were identified in 420 unigenes, indicating that merely 4.79% of the sequences contained SSRs. EST-SSR loci were subsequently screened on 24 hyacinth bean accessions collected from both China and Africa. Among 22 EST-SSRs, 11 loci showed polymorphism and revealed two to four alleles per locus. Tri-nucleotide motifs were the most common type of repeats, accounting for 54.55% of the total, followed by di-nucleotide repeats. The polymorphic information content (PIC) values ranged from 0.0767 to 0.4864, with a mean of 0.286. Furthermore, both principal coordinate analysis (PCA) and phylogenetic tree analysis indicated that all accessions were clustered into two main groups, and that all 19 Chinese accessions were clustered into the single group. These results suggest that there is a narrow genetic basis for Chinese hyacinth bean accessions. In summary, the results of this study suggest that the data mining of EST databases is a feasible method for generating EST-SSR markers for hyacinth bean, and these newly developed EST-SSR markers will be useful for further taxonomy, molecular breeding and comparative mapping studies.

Key words: Expressed sequence tag (EST); genetic diversity; hyacinth bean (*Lablab purpureus* L.); simple sequence repeats (SSR); polymorphism.
Abbreviations: EST-SSRs_expressed sequence tag-derived simple sequence repeats; NCBI_National Center for Biotechnology Information; PIC_polymorphic information content; H_O_observed heterozygosity; H_E_expected heterozygosity; PCA_principal coordinate analysis; RAPD_random amplified polymorphic DNA; AFLP_amplified fragment length polymorphisms; CTAB_hexadecyltrimethylammonium bromide; N_a_number of alleles.

Introduction

The hyacinth bean (*Lablab purpureus* L.), one of the most ancient and important legume crops in the world, is widely distributed throughout tropical and sub-tropical regions (Kimani et al., 2012). It is tolerant to drought and salinity, and thus can be cultivated in a wide range of climate and soil types with many uses such as human food, animal forage, cover crops and intercrops (D'Souza and Devaraj, 2010). As human food, it is consumed in the form of mature seeds, green pods or leaves; as animal fodder, it is used as feed or mixed with other feed as silage; and as cover crops or intercrops, it is used as a weed suppressor and soil erosion retardant (Maass et al., 2010). Despite all these advantages, the potential value of this crop has not been fully utilized, and there has been almost no research regarding this bean performed in China.

In modern plant breeding studies, molecular markers have become important and efficient tools. Molecular markers linked to agronomic traits are capable of increasing the accuracy of selection, and reducing the field workload, therefore the selection of suitable markers has become one of the key factors in the success of molecular breeding programs. At present, molecular markers such as random amplified polymorphic DNA (RAPD) (Kamakshi et al., 2012) and amplified fragment length polymorphisms (AFLP) (Venkatesha et al., 2007; Kimani et al., 2012) have been used to study the genetic diversity and population structure of hyacinth bean. These studies have provided valuable information for researchers and breeders working with hyacinth bean. However, these studies also have many limitations. For example, RAPD markers rely on large, intact DNA template sequences, which may not be available for RAPD development if the DNA is of lower quality or has been degraded. Other disadvantages of RAPD include lower resolving power for complex amplified products, difficulty in reproducing results, and the subjective nature of determining the precise characteristics of bands on a gel (Lee et al., 2002). Similarly, AFLP markers also have several disadvantages, such as complexity, low time efficiency, high cost, scoring of the

Table 1. Unigene sequences and SSR frequencies of hyacinth bean ESTs.

Parameter	No.	Repeat type	No of SSRs	Total length (bp)	Average length (bp)
Total ESTs	459	Di	7 (31.82%)[a]	90 (32.61%)	12.86
Total EST containing SSRs	22	Tri	12 (54.55%)	144 (52.17%)	12.0
Sequences containing 1 SSRs	22	Tetra	2 (9.09%)	24 (8.70%)	12.0
Sequences containing 2 SSRs	0	Hexa	1 (4.54%)	18 (6.52%)	18.0
Total SSRs identified	22	Total	22 (100%)	276 (100%)	12.55

[a]Numbers in parentheses are percentages

Table 2. Occurrence of different types of EST-SSRs in hyacinth bean.

SSR motif		Number of repeats						
		3	4	5	6	7	8	Total
Di-nucleotides	TC/GA	-	-	1	1	2	-	4
	TA/TA	-	-	1	-	1	1	3
Tri-nucleotides	TAA/TTA	3	-	-	-	-	-	3
	TGG/CCA	2	-	-	-	-	-	2
	GTG/CAC	2	-	-	-	-	-	2
	CTT/AAG	1	-	-	-	-	-	1
	CGA/TCG	1	-	-	-	-	-	1
	CTA/GAG	1	-	-	-	-	-	1
	CTA/TAG	1	-	-	-	-	-	1
	ACC/GGT	1	-	-	-	-	-	1
Tetra-nucleotides	TGGA/TCCA	1	-	-	-	-	-	1
	CTAA/TTAG	1	-	-	-	-	-	1
Hexa-nucleotides	TGATGC/ GCATCA	1	-	-	-	-	-	1
	Total	15	-	2	1	3	1	22

'-' denotes no repeats of the SSR motif

presence or absence of an AFLP band yielding dominant markers, and inaccurate quantitation of band intensities (Tang et al., 2009). Moreover, both RAPD and AFLP are dominant markers which are not able to differentiate homozygous from heterozygous genotypes (Zane et al., 2002). All these disadvantages significantly limit the wide application of these markers.

Compared with AFLP and RAPD, simple sequence repeat (SSR) markers have the advantages of simplicity, effectiveness, abundance, hypervariability, reproducibility, co-dominant inheritance, extensive genomic coverage, and ease of detection by polymerase chain reaction (PCR), and are considered to be the most powerful molecular markers for resolving genetic diversity and phylogenetic relationships in modern plant breeding systems (Powell et al., 1996). Initially, SSRs were mainly developed by one of two methods, i.e. SSR-enriched libraries or genomic sequence databases, therefore the isolation of SSRs and establishment of the specific primers require large amounts of both funding and time (Nakatsuji et al., 2011). Moreover, SSRs have poor transferability, which is generally limited to closely related species (Ma et al., 2009). It was reported that the transfer rate of SSR markers across the legume family was lower than 30%, and despite the large number of SSR markers available in legume crops including chickpea, soybean, pea and faba bean (Lichtenzveig et al., 2005; Nayak et al., 2010; Gong et al., 2010a; Gong et al., 2010b), there have been no reports on the development of SSR markers in hyacinth bean (Kamakshi et al., 2012).

Recently, the rapid progress in expressed sequence tag (EST) studies has generated a new source for the development of SSRs. Compared with conventional genomic SSR markers, SSRs from EST sequences have several intrinsic advantages: (1) they are embedded in functional gene sequences and directly associated with transcribed genes; (2) they have a lower cost of identification; and (3) they have a high transferability to related species (Varshney et al., 2005). In recent years, an increasing number of EST-SSRs have been identified in many crops and used extensively for comparative mapping, DNA fingerprinting, genetic diversity and transferability. The rapid and inexpensive development of SSRs from EST databases has become a feasible way to obtain high-quality nuclear markers (Gutiereea et al., 2005). However, there have yet to be any EST-SSR studies for the hyacinth bean reported, and the lack of suitable markers has become one of the most restricting factors for hyacinth bean breeding studies. The objectives of the present study are to develop EST-SSR markers for the hyacinth bean, examine their polymorphism, and evaluate the genetic diversity of hyacinth bean accessions, the results of which may be useful for marker-assisted selection and hyacinth bean breeding in the future.

Results

Distribution and frequencies of EST-SSRs in hyacinth bean

A total of 459 hyacinth bean ESTs were obtained from dbEST of NCBI. After clustering and assembly, 420 unigenes were used to search for SSRs. SSR screening with the SSRIT program resulted in the identification of 22 unigenes containing 22 microsatellites, suggesting that approximately 4.79% of the sequences contained SSRs. All 22 unigenes contained 1 SSR, the average length of which was 12.55 bp (Table 1). The 22 SSRs consisted of 31.82% di-nucleotide, 54.55% tri-nucleotide, 9.09% tetra-nucleotide and 4.54% hexa-nucleotide. TC/GA was the most common motif among the di-nucleotide repeats, accounting for 57.14%, followed by TA/TA (42.86%). TAA/TTA was the dominant motif among tri-nucleotide repeats, accounting for 25.00%, followed by TGG/CCA and GTG/CAC accounting for 16.67%, and CTT/AAG, CGA/TCG, CTA/GAG, CTA/TAG and ACC/GGT, each with a frequency of 8.33%. Further details of the different repeat motifs are listed in Table 2.

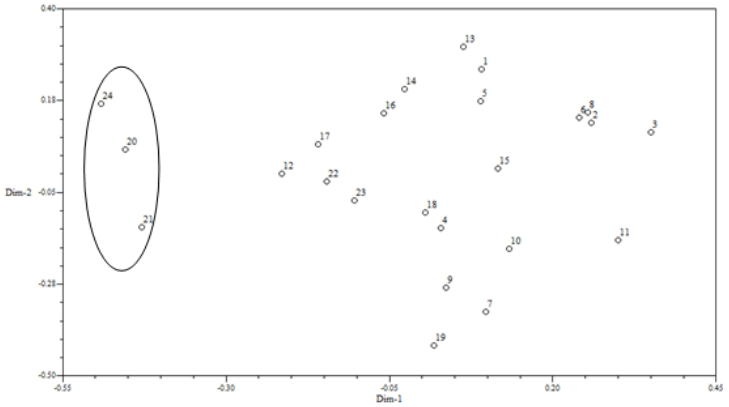

Fig 1. Two-dimension principle coordinate analysis of 24 hyacinth bean accessions. The numbers representing accessions are as listed in Table 5.

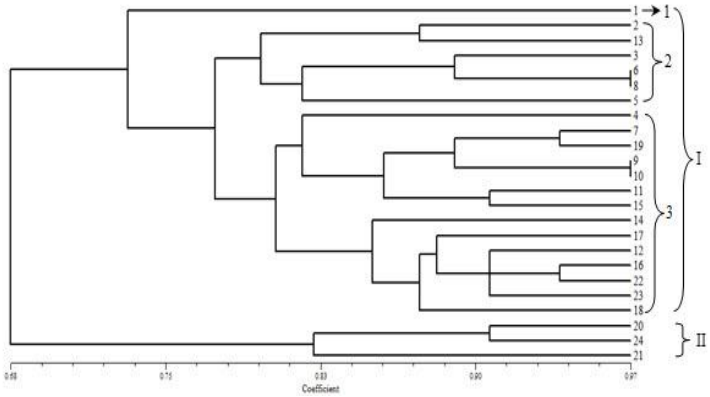

Fig 2. Dendrogram of 24 hyacinth bean accessions based on 11 EST-SSR data using the UPGMA clustering method. The major clusters and sub-clusters identified are indicated on the right. Between the major clusters, group I is comprised of two accessions from Africa and 19 accessions from China, and group II contains three accessions from Africa. The numbers representing accessions are as listed in Table 5.

Polymorphism of EST-SSRs

Of the 22 loci, 16 (72.73%) were found to have products with the expected size, two exhibited unexpected products, and four had no significant products. Among the 16 pairs of primers with products of expected size, 11 (68.75%) loci were polymorphic. A general functional classification of the 11 polymorphic EST-SSRs is shown in Table 3. BLAST analysis of these SSR-ESTs against the NCBI non-redundant sequence database revealed that nine of these ESTs matched known genes involved in ribosomal protein, heat shock protein, calmodulin-like protein, gibberellin receptor protein, zinc-dependent protease, auxin response factor and RNA-binding protein. Two ESTs represented unknown proteins. For each locus, the number of alleles ranged from 2 (Lpxu-002, Lpxu-003, Lpxu-006, Lpxu-012, Lpxu-015 and Lpxu-016) to 5 (Lpxu-011), with an average of 2.73. PIC values varied from a low of 0.0767 (Lpxu-003 and Lpxu-006) to a high of 0.4864 (Lpxu-009), with a mean value of 0.286 (Table 4). As a self-fertilizing crop, the hyacinth bean was expected to have a low heterozygosity. However, the H_O values ranged from 0 to 1.00, with an average of 0.3482, and the H_E values ranged from 0.0799 to 0.5785, with a mean of 0.3426. Eight EST-SSR loci (72.73%) were homozygous in all accessions, and only three loci (Lpxu-010, Lpxu-011 and Lpxu-013) did not significantly depart from the Hardy–Weinberg equilibrium ($P < 0.01$).

Assessment of genetic diversity in 24 hyacinth bean accessions

In order to evaluate the usefulness of these newly developed EST-SSRs, a genetic variation study was conducted in 24 hyacinth bean accessions. A cluster graph produced by two-dimension principle coordinate analysis indicated that the African accessions 20, 21 and 24 were clustered on the left side and were distinguished from the Chinese accessions, but accessions 22 and 23 were not distinguished from the Chinese accessions (Fig. 1). Consistent with this result, phylogenetic tree analysis demonstrated that at the coefficient of 0.70, all accessions were clustered into two main groups, i.e. Chinese type (I) and African type (II) (Fig. 2). However, African accessions 22 and 23 were clustered in the Chinese group (Figs. 1 and 2), indicating that these two cultivars were most likely an intermediate type between the two groups. At the coefficient of 0.78, accessions of the Chinese group were clustered into three sub-groups. Sub-group 1 contained only 1 accession from Xinjiang Uyghur Autonomous Region of China, sub-group 2 comprised 5 cultivars mainly from northern China, and sub-group 3 contained 14 accessions mainly from southern China. Finally, at the coefficient of 0.97, all 24 accessions could be distinguished from each other, except accessions 6, 8 and 9, 10. These results indicate that cultivars sharing the same or close origins tend to cluster together, suggesting that although some accessions are geographically distant from one another, their average genetic similarity is high.

Discussion

In this study, 459 hyacinth bean ESTs obtained from NCBI were used to search for SSRs. The results indicate that a total of 22 EST-SSRs were detected with a frequency of 4.79%, which was higher than previous results for oats (3.2%) (Becher, 2007), flax (3.5%) (Cloutier et al., 2009) and lettuce (4.5%) (Simko, 2009). However, this frequency was lower than those of wheat (7.41%) (Peng and Lapitan, 2005), dwarf lilyturf tuber (16.8%) (Li et al., 2011) and coffee (18.5%) (Aggarwal et al., 2007). These results indicate that the frequency of SSRs in EST sequences may vary among different plant species, and this difference may be resulted from the number and length of ESTs, identification criteria, size of the database, or database mining tools used for SSR research (Yan et al., 2009).

In general, the frequencies of different repeat motifs in different plant species show uneven. In the present study, tri-nucleotide repeat was the most dominant SSR in hyacinth bean. This result is in agreement with previous reports involving other legume crops, such as soybean, wild soybean, pea and faba bean (Kuroda et al., 2009; Kuroda et al., 2009; Gong et al., 2010a; Gong et al., 2010b). In previous studies, tri-nucleotide repeats were shown to be more abundant than other repeats in EST-derived SSRs, differentiating them from genomic SSRs derived from enriched genomic libraries, in which di-nucleotide repeats were predominant (Wang et al., 2010). The abundance of tri-nucleotide SSRs in plant ESTs may be attributed to the absence of frame-shift mutations within transcribed genes, owing to the suppression of non-trinucleotide SSRs in coding regions (Metzgar et al., 2000). Furthermore, the high frequency of tri-nucleotide repeats in coding regions is due to the mutation pressure and putative selection for specific amino acid stretches (Morgante et al., 2002). TAA/TTA was found to be the most abundant tri-nucleotide repeat motif, constituting for

Table 3. Characteristics of the 11 hyacinth bean EST-SSR markers

Locus name	EST No.	Primer sequence(5′→ 3′)	Motifs	TM (°C)	Expected size (bp)	Practical size (bp)	Number of alleles	BLAST
Hbp-001	GR463486.1	Forward: GACAACGACAGGGATAA Reverse: GAGTAAGAAGGGAAGTGA	$(ACC)_5$	58	122	106–143	3	Unknown
Hbp-002	GR463566.1	Forward: TTCCGCAAAGACAAGTT Reverse: CGTCAGCGAGAAGGGTA	$(CTA)_4$	60	140	138–161	2	Ribosomal protein
Hbp-003	GR463608.1	Forward: CCAAATCCGAATCAGCA Reverse: CTAAGCAGGAAAGCAAA	$(TGG)_8$	56	146	140–165	2	RNA-binding protein
Hbp-006	GR463523.1	Forward: GGTTCTCCTGCTGTTTG Reverse: CCATCTTGCTGCCCTCA	$(TGATGC)_3$	60	194	214–226	2	Heat shock protein
Hbp-009	GR463571.1	Forward: GCCCAGCTAAGATTGAG Reverse: GTTCTGATCCTATGACCG	$(TC)_6$	55	237	174–199	3	Calmodulin-like protein
Hbp-010	GR463556.1	Forward: AGCCTGACATTTCACCTG Reverse: TGCCACTTCAATCTCCC	$(GTG)_4$	58	242	232–264	3	Gibberellin receptor protein
Hbp-011	GR463544.1	Forward: GATGACCCTCCTACAG Reverse: GAGTTAACATAAATAAAGCT	$(TTA)_4$	56	246	180–265	5	Zinc-dependent protease
Hbp-012	GR463550.1	Forward: CCAAATCCGAATCAGCAG Reverse: TCCCTACCAAACACCGT	$(TGG)_4$	56	252	263–271	2	RNA-binding protein
Hbp-013	GR463463.1	Forward: CTCTACTATCATCCGTCTC Reverse: TCGGTCCATACTCTTCT	$(TC)_5$	60	284	303–364	4	Unknown
Hbp-015	GR463598.1	Forward: TTCTGGGCATTTCTCGG Reverse: GTAGTGGAAGGGTGGGT	$(TC)_7$	58	323	344–351	2	Predicted protein
Hbp-016	GR463546.1	Forward: ATTCGGGTTCGTGTTGG Reverse: CGATTGCTGTGGTGGTT	$(TC)_7$	58	323	345–351	2	Auxin response factor

Table 4. Characterization of the 11 SSR loci of hyacinth bean, including observed heterozygosity (H_O), expected heterozygosity (H_E), polymorphism information content (PIC) and chi-square test for Hardy–Weinberg equilibrium (P_{HW}).

Locus name	H_O	H_E	PIC	P_{HW}
Lpxu-001	1.0000	0.5000	0.3750	**
Lpxu-002	0.9583	0.4991	0.3746	**
Lpxu-003	0.0000	0.0799	0.0767	**
Lpxu-006	0.0000	0.0799	0.0767	**
Lpxu-009	0.1818	0.5785	0.4864	**
Lpxu-010	0.4583	0.4141	0.3561	
Lpxu-011	0.6667	0.4575	0.3701	
Lpxu-012	0.0000	0.1528	0.1411	**
Lpxu-013	0.5652	0.4924	0.4425	
Lpxu-015	0.0000	0.2268	0.2011	**
Lpxu-016	0.0000	0.2873	0.2460	**

** denotes significant deviation from Hardy–Weinberg equilibrium ($P < 0.01$).

Table 5. Cultivar names and origins of the 24 hyacinth bean accessions used in the study.

No.	Cultivar name	Origin	No.	Cultivar name	Origin
1	Hb1101	Xinjiang, China	14	Hb1114	Jiangsu, China
2	Hb1102	Hebei, China	15	Hb1115	Jiangsu, China
3	Hb1103	Hebei, China	16	Hb1116	Zhejiang, China
4	Hb1104	Qinghai, China	17	Hb1117	Zhejiang, China
5	Hb1105	Shanxi, China	18	Hb1118	Shanghai, China
6	Hb1106	Beijing, China	19	Hb1119	Guangzhou, China
7	Hb1107	Beijing, China	20	Hb1120	Nairobi, Kenya
8	Hb1108	Henan, China	21	Hb1121	Nairobi, Kenya
9	Hb1109	Hunan, China	22	Hb1122	Nairobi, Kenya
10	Hb1110	Hunan, China	23	Hb1123	Nairobi, Kenya
11	Hb1111	Hunan, China	24	Hb1124	Nairobi, Kenya
12	Hb1112	Jiangsu, China			
13	Hb1113	Jiangsu, China			

one-fourth of the total tri-nucleotide repeats identified. This result was consistent with previous findings in studies regarding other legume species, such as chickpea, soybean and *Medicago truncatula* (Lichtenzveig et al., 2005; Nayak et al., 2010; Cregan et al., 1994; Mun et al., 2006).

Furthermore, EST-SSRs generally show lower polymorphism compared with genetic SSR markers, due to the fact that they are located within expressed genes (Saha et al., 2006). However, a study by Woodhead et al. (2005) indicated that the differentiation of populations based on EST-SSRs was comparable to that based on genomic SSRs and AFLPs. In addition, Clark et al. (2003) reported that only a small percentage of genes are under positional selection. In the present study, 11 of 22 EST-SSRs were polymorphic, resulting in a polymorphism rate of 50%, which was comparable even higher than that of genomic SSRs (25-50%) (Zhong et al., 2008). A total of 30 alleles were detected in 11 loci, with an average of 2.73 per locus. The most frequent number of alleles per marker was two. The reason for which two alleles per marker was most frequent is they may include STS markers. PIC value is the most important index of molecular marker polymorphism. In the present study, the PIC values among seven di-nucleotide motifs were shown to range from 0.2011 to 0.4864, with an average of 0.344, and their values among 12 tri-nucleotide motifs ranged from 0.0767 to 0.375, with a mean of 0.2823, indicating that the polymorphism frequency was higher for di-nucleotide than for tri-nucleotide motifs, which was consistent with previous observations in soybean (Hisano et al., 2007). As self-fertilizing crops, legume crops are expected to have low heterozygosity. However, in this study, the H_O value among 11 loci ranged from 0 to 1, with a mean of 0.3482, and nine loci (81.82%) exhibited significant deviations from Hardy-Weinberg equilibrium corrected for multiple comparisons ($P < 0.01$), which was significantly higher than the results for soybeans (Li et al., 2008). Moreover, nine loci matched known genes, which were involved in different functional types of proteins and directly assisted in marker-trait associations (Thiel et al., 2003). In addition, SSRs with long motifs are typically ignored during the generation of genomic SSRs, due to their rarity in plant genomes (Wang et al., 2010). However, in this study one EST-SSR with a tetra-nucleotide repeat motif was found to be polymorphic, suggesting that long-motif SSRs are also useful for generating informative markers.

Assessment of genetic diversity within a population is essential for characterizing germplasm and provides insight into the evolutionary aspects, conservation, utilization and establishment of breeding programs for breeders (Li et al., 2011). Geographically widespread species generally tend to retain more genetic diversity than species with small geographical ranges. In previous studies, Venkatesha et al. (2007) and Kamakshi et al. (2012) used AFLP and RAPD markers to study the genetic diversity of Indian hyacinth bean accessions, and their results indicated that there was low genetic diversity among them. Furthermore, Kimani et al. (2012) used AFLP markers to study 50 Kenyan *L. purpureus* accessions and reported similar low diversity among Kenyan accessions. In China, hyacinth bean production regions may be divided into two main regions, i.e. the northern region (north of the Yangtze River) and southern region (south of the Yangtze River). Owing to the differences in climate and soil types, each main production region can be further divided into smaller sub-regions, and the different growing conditions and accessions show high agronomic variation among these sub-regions. However, the results of this study (Figs. 1 and 2) indicate that most accessions from the northern or southern regions are clustered into the same subgroups and that all accessions from China are clustered into a single group. These results suggest that genetic variation among Chinese populations is low too and that there is a narrow genetic basis for hyacinth bean in China, similar to that of Indian and Kenyan accessions. The reason for this phenomenon may be due to the narrow genetic base for breeding stock and extensive exchange of germplasm among smallholder farmers across the country. Therefore, more foreign germplasm must be introduced into Chinese breeding programs to broaden the genetic base of hyacinth bean, which will benefit the future hyacinth bean projects.

Materials and methods

Plant material and DNA extraction

In the present study, a total of 24 hyacinth bean varieties which are widely cultivated in China and Africa were used (Table 5). DNA was extracted from 20-day-old seedlings grown in a glasshouse. A total of 0.2 g of fresh leaves was used for each repeat and DNA was extracted by the CTAB method according to the manufacturer's instructions. The DNA concentration was adjusted to 20 ng/mL and determined using an ND-1000 spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE, USA).

Database search and primer design

L. purpureus ESTs were acquired from the EST database of NCBI (for which the data was available to May 2012). In total, 459 available ESTs were obtained for this study. In a preliminary phase, the ESTs sequences were assembled into unigenes using DNASTAR software, with the parameters for clustering sets at a minimum of 95% identity in 40-bp overlaps. Then SSRIT (http://www.gramene.org/gramene/searches/

ssrtool) was used for identifying SSRs in unigenes with the criteria of a minimum of 5, 4, 3 and 3 repeat units for di-, tri-, tetra- and higher order nucleotides, respectively. The putative functions of the EST-SSR markers were determined using the basic local alignment search tool X (BLASTX) analysis at the NCBI website with an *E*-value threshold of 10^{-10}.

The primers were designed using the Primer Premier 5.0 software. The major parameters for primer pair design were as follows: primer length range from 18 to 24 bases (optimal 20 bases), annealing temperature of 50–60°C (optimal 55°C), G–C content of 40–70% (optimal 50%), and PCR product sizes ranging from 100–400 bp (optimal 200 bp).

SSR amplification and PCR product analysis

The reagents for PCR amplification were purchased from TaKaRa Biotechnology Co., Ltd. (Dalian, China). PCR was performed on a PTC-225 thermal cycler (MJ Research Inc., Waltham, MA, USA). The forward primers of each pair were labelled with 6-FAM fluorescent dye. All PCR amplifications were performed in 20 μL reaction mixtures containing a 20 ng genomic DNA template, 1× PCR buffer (containing Mg^{2+}), 0.2 mM dNTPs, 0.2 mM forward and reverse primers, and 1 U of Taq polymerase (TaKaRa, Dalian, China). The PCR program was as follows: 95°C for 5 min, 36 cycles each at 94°C for 30 s, 56°C for 30 s, 72°C for 60 s, and a final extension at 72°C for 10 min. The PCR products were diluted and detected on a MegaBACE 1000 DNA Analysis System as previously described by Gong et al. (2010a). Analysis of amplified fragment sizes was performed using methods also as described by Gong et al. (2010b).

Data and statistics

The number of alleles (N_a), observed heterozygosity (H_O) and expected heterozygosity (H_E) were calculated using POPGENE software, version 1.3 (Choudhary et al., 2009). The polymorphism information content (PIC) value was calculated using the formula developed by Anderson et al. (1993). A dendrogram was constructed using the unweighted pair group method with arithmetic means (UPMGA).

Conclusion

The 11 EST-SSR markers newly developed in this study showed high polymorphism and were successfully used to investigate the genetic diversity among 24 Chinese and African hyacinth bean accessions. The study results demonstrate that hyacinth bean ESTs are a valuable resource for developing SSR markers. With the rapid development of functional genomics, these new EST-SSRs markers will be useful in further studies for diversity analysis, genetic map construction and gene discovery in hyacinth bean.

Acknowledgments

This work is supported by the Project of National Natural Science Foundation of China (31101538; 31000942; 31000676), and the Grand Science and Technology Special Project of Zhejiang Province (2012C12903-4-1; 2012C12903-6-3).

References

Aggarwal RK, Hendre PS, Varshney RK, Bhat PR. Krishnakumar V, Singh L (2007) Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. Theor Appl Genet. 114: 359-372

Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. Genome. 36: 181-186

Becher R (2007) EST-derived microsatellites as a rich source of molecular markers for oats. Plant Breeding. 126: 274-278

Choudhary S, Sethy NK, Shokeen B, Bhatia S (2009) Development of chickpea EST-SSR markers and analysis of allelic variation across related species. Theor Appl Genet. 118: 591-608

Clark AG, Glanowski S, Nielsen R, Thomas PD. Kejariwal A, Todd MA, Tanenbaum DM, Civello D, Lu F, Murphy B, Ferriera S, Wang G, Zheng XG, White T J, Sninsky JJ, Adams M D, Cargill M (2003) Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. Science. 302: 1960-1963

Cloutier S, Niu Z, Datla S, Duguid S (2009) Development and analysis of EST-SSRs for flax (*Linum usitatissimum* L.). Theor Appl Genet. 119: 53-63

Cregan PB, Bhagwat AA, Akkaya MS, Rongwen J (1994) Microsatellite fingerprinting and mapping of soybean. Meth Mol Cell Biol. 5: 49-61

D'Souza MR, Devaraj VR (2010) Biochemical responses of Hyacinth bean (*Lablab purpureus*) to salinity stress. Acta Physiol Plant. 32: 341-353

Gong YM, Xu SC, Mao WH, Hu QZ, Zhang GW, Ding J, Li YD (2010a) Developing new SSR markers from ESTs of pea (*Pisum sativum* L.). J Zhejiang Univ Sci B. 11: 702-707

Gong YM, Xu SC, Mao WH, Hu QZ, Zhang GW, Ding J, Li ZY (2010b) Generation and characterization of 11 novel est derived microsatellites from *Vicia faba* (Fabaceae). Am J Bot. 97: 69-71

Gutierrez MV, Vaz Patto MC, Huguer T, Cubero JI, Moreno MT, Torres AM (2005) Cross-species amplification of *Medicago truncatula* microsatellites across three major pulse crops. Theor Appl Genet. 110: 1210-1217

Hisano H, Sato S, Isobe S, Sasamoto S, Wada T, Matsuno A, Fujishiro T, Yamada M, Nakayama S, Nakamura Y, Watanabe S, Harada K, Tabata S (2007) Characterization of the soybean genome using EST-derived microsatellite markers. DNA Res. 14: 271-281

Kamakshi N, Sudhakar P, Muralikrishna T (2012) Analysis of genetic diversity of field bean genotypes showing resistance to pod borers and pod bugs using RAPD technique. Curr Biotica. 5: 421-427

Kimani EN, Wachira FN, Kinyua MG (2012) Molecular diversity of kenyan lablab bean (*Lablab purpureus* (L.) Sweet) accessions using amplified fragment length polymorphism markers. Am J Plant Sci. 3: 313-321

Kuroda Y, Tomooka N, Kaga A, Wanigadeva SMSW, Vaughan DA (2009) Genetic diversity of wild soybean (*Glycine soja* Sieb. et Zucc.) and Japanese cultivated soybeans [*G. max* (L.) Merr.] based on microsatellite (SSR) analysis and the selection of a core collection. Genet Resour Crop Evol. 56: 1045-1055

Lee SW, Ledig FT, Johnson DR (2002) Genetic variation at allozyme and RAPD markers in *Pinus longaeva* (Pinaceae) of the White Mountains, California. Am J Bot. 89: 566-577

Li G, Ra WH, Park JW, Kwon SW, Lee JH, Park CB, Park YJ (2011) Developing EST-SSR markers to study molecular diversity in *Liriope* and *Ophiopogon*. Biochem Sys Eco. 39: 241-252

Li YH, Guan RX, Liu ZX, Ma YS, Wang LX, Li LH, Lin FY, Luan WJ, Chen PY, Yan Z, Guan Y, Zhu L, Ning XC, Smulders MJM, Li W, Piao RH, Cui YH, Yu ZM, Guan M, Chang RZ, Hou AN, Shi AN, Zhang B, Zhu SL, Qiu LJ

(2008) Genetic structure and diversity of cultivated soybean (*Glycine max* (L.) Merr.) landraces in China. Theor Appl Genet. 117: 857-871

Lichtenzveig J, Scheuring C, Dodge J, Abbo S, Zhang HB (2005) Construction of BAC and BIBAC libraries and their applications for generation of SSR markers for genome analysis of chickpea, *Cicer arietinum* L. Theor Appl Genet. 110: 492-510

Ma KH, Kim NS, Lee GA, Lee SY, Lee JK, Yi JY, Park YJ, Kim TS, Gwag JG, Kwon SJ (2009) Development of SSR markers for studies of diversity in the genus *Fagopyrum*. Theor Appl Genet. 119: 1247-1254

Maass BL, Knox MR, Venkatesha SC, Angessa TT, Ramme S, Pengelly BC (2010) *Lablab purpureus*—a crop lost for Africa? Trop Plant Biol. 3: 123-135

Metzgar D, Bytof J, Wills C (2000) Selection against frameshift mutations limits microsatellite expansion in coding DNA. Genome Res. 10: 72-80

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nat Genet. 30: 194-200

Mun JH, Kim DJ, Choi HK, Gish J, Debellé F, Mudge J, Denny R, Endré G, Saurat O, Dudez AM, Kiss DB, Roe B, Young ND, Cook DR (2006) Distribution of microsatellites in the genome of *Medicago truncatula*: a resource of genetic markers that integrate genetic and physical maps. Genetics. 172: 2541-2555

Nakatsuji R, Hashida T, Matsumoto N, Tsuro M, Kubo N, Hirai M (2011) Development of genomic and EST-SSR markers in radish (*Raphanus sativus* L.). Breeding Sci. 61: 413-419

Nayak SN, Zhu HY, Varghese N, Datta S, Choi HK, Horres R, Jüngling R, Singh J, Kishor PBK, Sivaramakrishnan S, Hoisington DA, Kahl G, Winter P, Cook DR, Varshney RK (2010) Integration of novel SSR and gene-based SNP marker loci in the chickpea genetic map and establishment of new anchor points with *Medicago truncatula* genome. Theor Appl Genet. 120: 1415-1441

Peng JH, Lapitan NL (2005) Characterization of EST-derived microsatellites in the wheat genome and development of eSSR markers. Funct Integr Genomics. 5: 80-96

Powell W, Machray GC, Porvan J (1996) Polymorphism revealed by simple sequence repeats. Trends Plant Sci. 1: 215-222

Saha MC, Cooper JD, Mian MA, Chekhovskiy K, May GD (2006) Tall fescue genomic SSR marker: development and transferability across multiple grass species. Theor Appl Genet. 113: 1449-1458

Simko I (2009) Development of EST-SSR markers for the study of population structure in lettuce (*Lactuca sativa* L). J Hered. 100: 256-262

Thiel T, Michalek W, Varshney R, Graner A (2003) Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L). Theor Appl Genet. 106: 411-422

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 23: 48-55

Venkatesha SC, Gowda MB, Mahadevu P, Rao AM, Kim DJ, Ellis THN, Knox MR (2007) Genetic diversity within *Lablab purpureus* and the application of gene-specific markers from a range of legume species. Plant Genetic Resources-Charac. 5: 154-171

Wang XB, Mulock B, Guus B, Mccallum B (2010) Development of EST-derived simple sequence repeat markers for wheat leaf rust fungus, *Puccinia triticina* Eriks. Can J Plant Pathol. 32: 98-107

Woodhead M, Russell J, Squirrell J, Hollingsworth PM, Mackenzie K, Gibby M, Powell W (2005) Comparative analysis of population genetic structure in *Athyrium distentifolium* (Pterdophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. Mol Ecol. 14: 1681-1695

Yan QL, Zhang YH, Li HB, Wei CH, Niu LL, Guan S, Li SG, Du LX (2008) Identification of microsatellites in cattle unigenes. J Genet Geomics 35:261–266

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. Mol Ecol. 11: 1-16

Zhong SB, Yang BJ, Alfenas AC (2008) Development of microsatellite markers for the guava rust fungus, *Puccinia psidii*. Mol Ecol Notes. 8: 348-350