GENETIC DIVERSITY AMONG ASPARAGUS SPECIES AND CULTIVARS USING SSR MARKERS

Idrees, M., M. Irshad, M. L. Pathak[1], A. Tariq and R. Naeem

Department of Biotechnology and Genetic Engineering, Kohat University of Science and Technology (KUST). Kohat, Pakistan; CAS Key Laboratory of Mountain Ecological Restoration and Bioresource Utilization and Ecological Restoration and Biodiversity Conservation Key Laboratory of Sichuan Province, Chengdu Institute of Biology, Chinese Academy of Science, P.O Box 416, Chengdu 61004, China; [1]The University of Chinese Academy of Science, Beijing, China

Abstract

The aim of the present study was used to develop SSR markers and to determine genetic relationships among *Asparagus* species and its cultivars. We mined 72953 *Asparagus* nucleotide sequences from NCBI and were analyzed. In total, 143 SSRs from 1943 SSRs containing sequences were identified. Of these, 13.6 % were dinucleotide repeats and 2.3 % were trinucleotide repeats. The most frequent dinucleotide repeat motif was AA/TT (73.9 %). The percentage of tri-nucleotide motifs was highest which coded for stop codon (36 %), whereas the Glycine was least present (4.5%). Among the total of 14 SSR primers used, 10 markers yielded 144 (92.3 %) polymorphic bands with an average of 14.4 alleles per primer. Cluster analysis based on UPGMA grouped the *Asparagus* species and its cultivars into two main clusters. Cluster A contained only A. grecilius which is more diverse than others, while cluster B was further clustered into two sub clusters. Cluster I was comprised of *Asparagus officinalis* cultivars and cluster II was comprised of *Asparagus* wild species. Furthermore, these primers (42.8%) were found to be transferable in other medicinal plants (*Curcuma longa* L. and *Saussurea costus*). The results suggest that SSR markers are sufficiently useful and powerful to assess genetic relationships and diversity analysis in *Asparagus* and *A. officinalis* cultivars. Furthermore, these markers will be particularly useful for evolutionary and genetic mapping studies in *Asparagus*.

Keywords: *Asparagus*, Simple Sequence Repeats, genetic diversity, *Saussurea costus*, *Curcuma longa*.

INTRODUCTION

Asparagus Lindley is a genus of dioecious plants with approximately 150 species throughout the world. The genus has been moved from the family Liliaceace to a newly created family Asparagaceae (Kubitzki and Rudall 1998). It is grown in a wide range of environments including temperate regions, humid tropics and arid tropics up to an altitude of 1200 m (Dutta 2007). *Asparagus* remains valuable as plant species from eras having both therapeutic and nutraceutical importance as well as for food consumption (Shasnay *et al*. 2003). *Asparagus* contains fructans and saponin that have antitumor activity and reduce the risk of disorder such as diarrhea, constipation and helps in diseases like osteoporosis, obesity, cardiovascular disease, rheumatism and diabetes. The root of *Asparagus* possesses antispasmodic, aphrodisiac, diuretic, demulcent and galactagogue properties and has been commonly used for different medicinal purposes. Its roots are also useful in diseases of liver and kidney, gonorrhea, epilepsy, jaundice and disease of blood and eyes. Internally it is used in the treatment of infertility, threatened miscarriage, menopausal problems, stomach ulcer, hyperacidity and loss of libido, while externally it is used to treat stiffness in joints. It also stimulates insulin action, secretion and inhibits starch digestion (Mandal *et al*. 2000).

Simple sequence repeats (SSRs) are stretches of genomic DNA that consist of short tandem repeated motifs of 1-6 nucleotides, and can be found in high frequency of most taxa nuclear genomes (Beckmann and Weber 1992). SSRs are more desirable because of their reproducibility, codominant inheritance, multiallelic nature, high abundant and well genetic coverage, which makes the method highly polymorphic and specific (Bornet and Branchard 2001). These are a PCR-based technique amenable to

automation and require short period of time for their analysis. SSRs are widely used in various aspect of molecular genetics including fingerprinting, marker assisted selection, genetic map development, gene flow characterization, linkage analysis and genetic diversity (Koppolu *et al.* 2010). The conserved flanking sequence of specific SSR loci in genome are shown to be conserved within taxa, genus and even among closely related genera (Varshney *et al.* 2002). SSR sequences frequently change by proofreading errors and slippage during DNA replication and thus changed the length of the repeats (Eisen 1999). However, genomic SSR markers development is expensive and time consuming, particularly if developed from genomic libraries. These SSRs markers previously used to analyze plant genetic diversity (Hafezi-Shahroodian *et al.* 2011). During the past few years, *In-silico* SSR-mining has gone through a rapid evolution. With much advancement in bioinformatics and genomics, various *in-silico* approaches are being used increasingly. These tools were used to design for the development of SSR markers in a very short span of time. These bioinformatics approaches can eliminate the need for screening and costly library construction. Mining SSR from NCBI databases has been streamlined with technological advance to make the process more efficient, cheaper and successful and proved to be an effective approach to develop microsatellites markers for genetic mapping and genetic diversity (Portis *et al.* 2007). In summary, varietal characterization and taxonomy are difficult in *Asparagus*, so molecular markers have proven useful. To the best of our knowledge, there is no study using SSRs markers to evaluat genetic diversity among *Asparagus* species and its cultivars in Pakistan. The aim of the current study was to examine (1) frequency and distribution of SSRs (2) Nature and distribution of encoded amino acid (3) BLAST analysis of amplified markers and 4) to develop SSRs primer and validate their polymorphism/ transferability in other medicinal plants.

MATERIAL AND METHODS

All nucleotide sequences of *Asparagus* were retrieved in Fasta format from National Center Biotechnology Information (NCBI) database. The nucleotide sequences of *Asparagus* were aligned using Clustal X and screened for the presence of Simple Sequence Repeats using microsatellite search module "MISA" which include di, tri and tetra- nucleotide repeats in the Search. These were assembled using CAP3 programme (Huang and Madan 1999).

Table 1. **Plant collections and specimen of *Asparagus* species.**

Species and Cultivars	Collection sites	Altitude (m)	Longitude (E)	Latitude (N)
A. racemosus wild	Charbhage, Swat	909	72° 21'E	34° 46'N
A. capitatus wild	Ghalegay, Swat	950	72° 12' and 72° 32'E	35° 20' to 35° 48'N
A. gracelus wild	Shamozu, Swat	465	72° 12'E	34° 68'N
A. adscenden wild	Jerma, Kohat	508	71° 29'E	30° 40'N
A. setaceus	Bhage Jinnah, Lahore	209	74° 33'E	31° 54'N
A. densiflorus	More green, Lahore	217	74° 34'E	31° 54'N
A. plumosus	Mingora, Swat	950	72° 26'E	34° 69'N
A. officinalus	NARC, Islambad	508	73º 10'E	33° 42' N
A. officinalis Cv Abril	ARI, Mingora	984	72° 26' and 72° 36'E	34° 68'N
A. officinalis Cv Apollo	ARI, Mingora	984	72° 26' and 72° 36'E	34° 68'N
A. officinalis Cv Gersengum	ARI, Mingora	984	72° 26' and 72° 36'E	34° 68'N
A. officinalis Cv Huchel	ARI, Mingora	984	72° 26' and 72° 36'E	34° 68'N
A. officinalis Cv Taranga	ARI, Mingora	984	72° 26' and 72° 36'E	34° 68'N
A. officinalis Cv Para selection	ARI, Mingora	984	72° 26' and 72° 36'E	34° 68'N
Saussurea costus	Shangla	1904	72° 35'E	34° 53'N
Curcuma longa L.	Bannu	371	70° 18'E	33°00'N

The SSRs primers were designed to the nucleotide sequences of *Asparagus* having SSRs motifs using primer3 software, Primer design criteria included, minimum of five or six di, tri, and tetra-

nucleotide repeats. Simple Sequence Repeats were masked as target region including amplicon size ranges from 200-300 bp, 55- 60ºC of optimal annealing temperature, average GC content ranges from 50 to 60%, and length of primer ranges from 18 to 24 bp. The output was displayed as sequence of forward and reverse primers. The Primers were synthesized by Invitrogen Corp. CA, USA. SSR loci putative functions were allowed by similarities with the non-redundant sequence database using BLASTX software (Altschul *et al.* 1997).

A total of 14 different *Asparagus* species and *A. officinalis* cultivars was collected from different regions of Pakistan including Islamabad, Lahore, Kohat and Swat (Table 1 and Fig. 1A). Leaves were collected and immediately stored in silica bags.

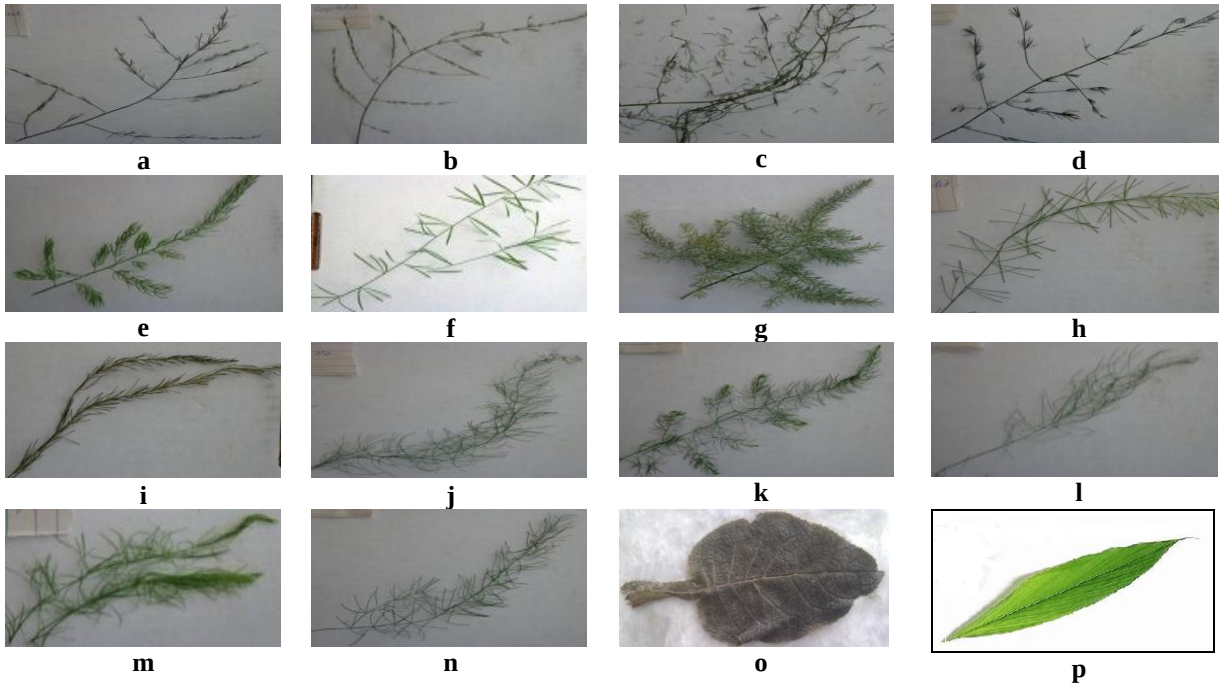


Fig. 1A. Specimen of *Asparagus* species: a. *A. racemosus;* **b**. *A .capitatus*; **c**. *A. gracelus*; **d**. *A. adscenden*; **e**. *A.setaceus*; **f**. *A.densiflorus*; **g**. *A.plumosus*; **h**. *A. officinalus*; **i**. *A. officinalis*; **j**. *A. officinalis*; **k**. *A. officinalis;* **l.** *A. officinalis*; **m**. *A. officinalis*; **n.** *A. officinalis*; **o**. *Saussurea costus*; and **p**. *Curcuma longa*.

Genomic DNA was extracted from the fresh and young leaves of *Asparagus* using standard CTAB method (Cetyl trimethyl ammonium bromide) as described by Doyle and Doyle (1987) with slight modifications, i.e the addition of high concentration of PVP (polyvinyl-pyrrolidine). Fresh leaves of *Asparagus* (100 mg) were ground well using pre-warmed solution containing 800 µl extraction buffers. After grinding, 20 µl of β -mercaptoethanol was added and shaken vigorously by inversion to produce slurry. The tube was then incubated at 65ºC for 30 minutes with intermittent shaking and vortexing every 5 minutes. The tube was kept at room temperature (25ºC) for 7 minutes and then 800 µl of chloroform: isoamylalchol (24:1) was added.

The solution was mixed by inversion for 10 - 15 minutes and kept at room temperature for 7 minutes. The emulsion was centrifuged at 12,000 g at 24ºC for 10 minutes and the upper layer was transferred to another sterile tube. The chloroform: isoamylalchol step was repeated again and the supernatant was centrifuged at 12,000 g for 10 minutes. The upper aqueous phase was placed in a sterile tube. The DNA was precipitated by adding equal volume (0.6V) of chilled isopropanol and 30 µl of 5 M NaCl and then was kept at -20ºC for one hour. The sample was centrifuged at 12,000 g for 10 minutes at 4ºC. The supernatant was discarded and the DNA pellet was first washed with 800 µl of wash buffer

(Tris HCl 5mM, Nacl 25mM and Etnaol 25%) and then with 300 µl of 70% ethanol. The pellet was air dried at room temperature for 15 minutes and finally the purified DNA pellet was allowed to dissolved in 30 µl of TE buffer and stored at -20ºC.

For polymorphism and transferability study, PCR amplifications were carried out in Veriti 96 well thermal cycler (Applied Biosystem Inc, USA) using SSRs markers. Amplifications were carried out in 25 µl master mix containing 2.5 µl of 1X PCR buffer; 0.2 mM dNTPs mix; 2.5 mM $MgCl_2$; 1 Unit of Taq DNA polymerase (Invitrogen); 0.2 µM of each forward and reverse primers and 2 µl of 50 ng of template DNA. Conditions for amplification of the region consists of initial denaturation at 95ºC for 5 minutes and then 35 cycles of denaturation at 94ºC for 1 minute (denaturation), annealing of primer at 55ºC for 1.5 minutes, 72ºC for 1.5 minutes (extension), with the final step of extension at 72ºC for 7 minutes followed by hold temperature at 4ºC. The PCR products were separated on 2% Agarose gel in Tris-Borate EDTA buffer and stained with 0.5 mg/ml ethidium bromide. These were visualized and photographed under a UV transilluminator (WiseDoc). The sizes of the DNA products were calculated by comparison with 1 kb DNA ladder.

Analysis of the electrophoresis patterns of each SSR primers were scored manually as presence or absence of bands. The results were analyzed on the principle that a band is considered to be 'polymorphic' if it is present in some individuals and absent in others, and 'monomorphic' if present in all individuals. Numerical Taxonomy System (NT-SYS), version 2.11 from applied biostatis Inc. (2002), was used to analyze the results obtained from scoring. For measuring genetic similarity and distance among *Asparagus* species and cultivars, Nei and Li genetic similarity coefficient was used (Nei and Li 1979). Cluster analysis was performed by UPGMA (Unweighted Pair wise Methods with Arithmetic averages) to generate a dendogram.

RESULTS AND DISCUSSION

Microsatellite markers were considered as the best genetic markers for the analysis of genetic variation and genetic improvement of plants in comparison to other available markers. Conventional methods for the development of SSR markers were costly and time consuming. However the aim of the present study was to develop cost efficient method by retrieving nucleotide sequences from NCBI database that containing microsatellite to produce SSR markers in *Asparagus* and to evaluate the genetic relationship among *Asparagus* species and its cultivars in Pakistan using SSR markers. Several studies have used traditional approaches for microsatillete markers evolution in various plants species (Thiel *et al.* 2003). Previously, EST-SSR markers have been developed in *Asparagus* (Caruso *et al.* 2008).

Table 2. Summary of SSRs in *Asparagus.*

Screening of Parameter	Generated data by SSR Finder
Total examined Sequences	72953
Nucleotide sequence size	204609
SSR identified	1915
Sequence containing SSR	143
SSR containing more than 1 sequence	96
Dinucleotide	261
Trinucleotide	44
Tetranucleotide	16
Pentanucleotide	5

SSR markers were developed from 72953 nucleotide sequences on NCBI by *In-silico* methods. The frequency of SSRs (91.4%) recognized in our study is much higher as compared to previous studies in *Catharanthus roseus* (10.2%). The difference in the frequencies of SSRs detected depends on the search

module used for SSRs type motif, sequence size analysis and tool for mining database (Portis *et al.* 2007, Gupta and Prasad 2009).

From 261(13.6%) SSR containing sequences, di-nucleotide was the most abundant which account for 13.6 percent, followed by tri-nucleotide, tetra-nucleotide and penta-nucleotide which were 2.3, 0.84 and 0.27 %, respectively (Table 2), similar observations were reported in the plant of *Nelumbo nucifera* (Pan *et al*. 2010). In few studies of animals, dinucleotide repeat was the dominant type reported. The most frequent dinucleotide repeat motifs were AA/TT (73.9%), followed by AT/AT (17.6%), and AG/CT (2.68%). Similar results have been reported in other taxa of plant genomes (Kumpatla and Mukhopadhyay 2005). The dinucleotide motif TA/TA (1.18%) was found as least prevalent while the AT motif was found to be more frequent in plant genome, similarly AG/CT motif was represented as codons GAG, UCU, and CUC in mRNA populations which were further translated into the amino acids Argnine, Alanine, and Leucine, respectively. Alanine and Leucine occurred at high frequency of 8 to10 % in all proteins (Wen 2010). Among the tri-nucleotide repeats, motif AAA/TTT (50%) was most common, followed by ATT/GCG (36.3%) and CCT/TCC (9%). The most common tetra-nucleotide motif was AAAA/TTTT (62.5%) followed by GAGA/CTCT (37.5%) and ATAT/ATAT (6.25%) (Fig. 1a, b, c, d).

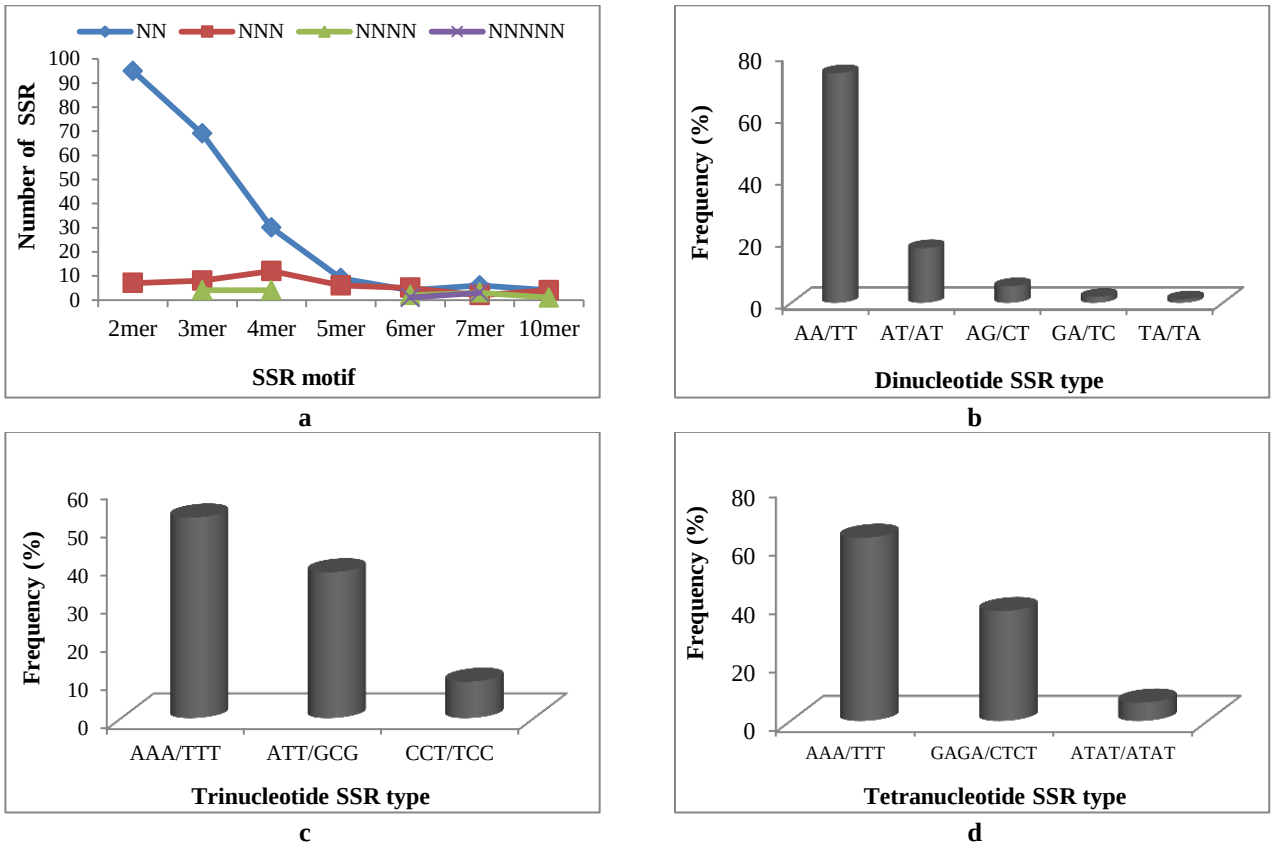


Fig. 1. Frequency distribution of identified SSR sequence: **a**. all different nucleotide repeats in *Asparagus*; **b**. dinucleotide; **c**. Trinucleotide; and **d**. Tetranucleotide.

Each of the tri-nucleotide motifs in SSR loci is coded for a specific amino acid which possesses several biological functions. In the present study, 44 tri-nucleotide repeats motif was found, which code for 4 amino acids, viz. lysine, phenylalanine, arginine and glysine, and one stop codon. The percentage of tri-nucleotide motifs was highest which coded for stop codon (36 %), followed by Lysine (31 %), Phenylalanine (21.4 %), and Arginine (7.1 %), while the Glycine was least present (4.5%) (Fig. 2).

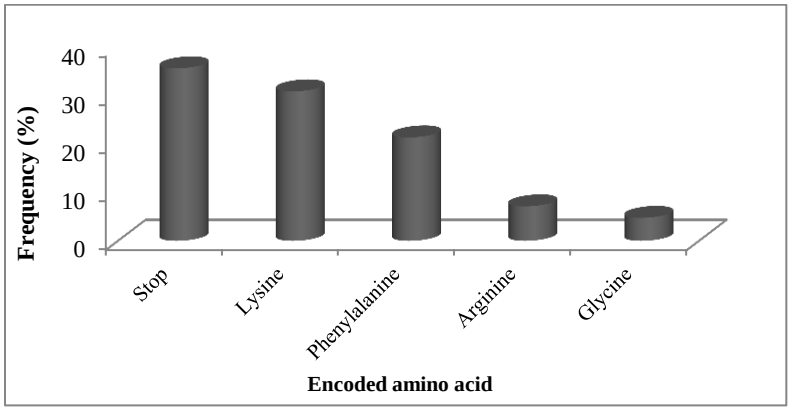


Fig. 2. Distribution of SSR encoded amino acid in *Asparagus*

Arginine in the present result was 21.4% that similar to the previously studied in *Arabidopsis*, *Gnetum*, *Oryza* and *Pinus*, which showed that the predominant amino acid was Arginine (Victoria *et al.* 2011). Till date, the amino acid distribution in SSR loci in *Asparagus* has not reported yet. These encoded amino acids of SSR were divided further into 4 categories, (a) hydrophobic and hydrophilic amino acid, (b) polar and non polar amino acid, (c) acidic, basic and neutral amino acid, and (d) aliphatic and aromatic amino acid (Fig. 3a, b, c, d).

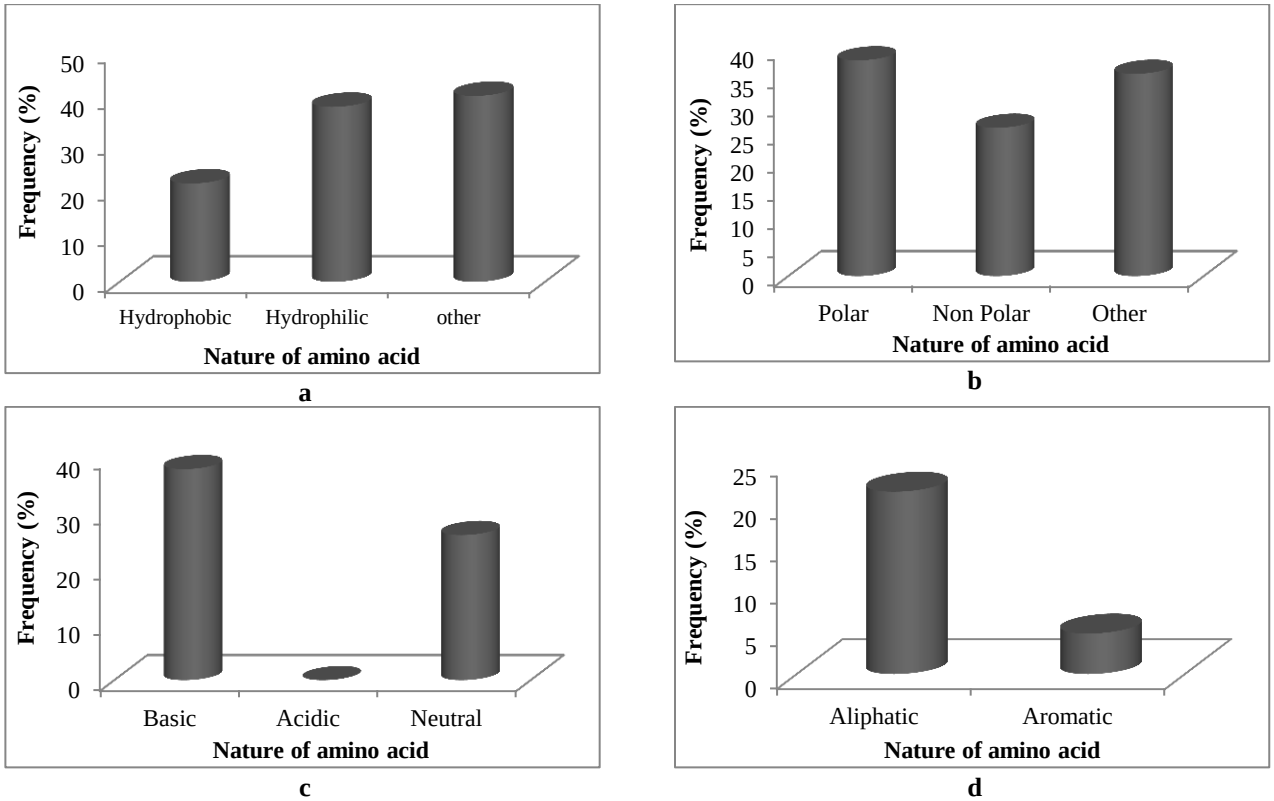


Fig. 3. Distribution of SSR encoded amino acid in *Asparagus:* **a**. hydrophobic and hydrophilic amino acid; **b**. polar and non polar amino acid; **c**. acidic, basic and neutral amino acid; and **d**. aliphatic and aromatic.

From 143 sequences containing SSR, 40 (28.0 %) nucleotide sequences were allowed for primer designing, while remaining 103 (72 %) sequences were unsuitable for primer construction because they have SSR motifs either at the beginning or at the end of nucleotide sequence. Out of 40, 25 primer pairs were developed successfully by utlizing primer 3 softwares (Rozen and Skaletsky 2000). Finally, subsets of 14 primer pairs were selected on the basis of product size for the justification of amplification and evaluation of genetic variations in *Asparagus* species and its cultivars (Table S1).

Table S1. **Characterization of 14 SSRs developed from *Asparagus.***

Primers	ID	SSR motifs	Tmp. (ºC)	GC%	Forward primer	Reverse primer	Product size (bp)
ndhf-rpl32	gi|327442444	$(AT)_4$, $(AA)_8$	58.98	39.05	TTCAAATTTTCCCCTCCTTTC	TTGCTCAAAATGGGTCATTC	235
rps16-trnK	gi|327442522	$(AA)_4$	60.3	44.05	TTCCTTGAAAAAGGTGCTCAA	GGTGGATTCCCACAACAAGA	207
trnQ-rps16	gi|327442389	$(TT)_4$	60.3	38.75	ATGATTCACCATCCCGAAAA	TGAATAGTCATTGGATCAACGGTA	197
psbD-trnT	gi|327442494	$(AA)_5$	59.8	42.5	CGTCCAATGCCCTTTACAAT	AATTTAGGGGCAGGGAAAAA	227
atpB-rbcl	gi|37721940	$(AT)_4$	59.83	43.93	TCTCATTGGCTGCTGTTCTTT	ATATATGGCGCAACCCAATC	245
petB	gi|56603578	$(TT)_4$	60	43.75	AAGAGGCCTGTAACGAGCAA	CACAAATACTGATTTCACCGGATA	183
matK	gi|374975311	$(TT)_4$	60.2	47.5	CCAAAATCGATTCGTTGGAC	CATATGGATGGGATGGGGTA	247
A synth	gi|16075	$(TTGGT)_4$	60	45	GCTGGTCGAACAATTTGGAT	ATACAACGCCCCACACATTT	235
rpl32-trnL	gi|327442338	$(AT)_5$,$(AA)_4$	59.7	35.2	TCTCTTTGACGATTTGTTTTTCA	CCGGAACATTTGGAATTTGA	154
ZHD1	gi|164562228	$(TT)_{10}$	60	47.5	GGAAGAGGGTGCGTGTTTTA	AAACGAACCAAAGTGCCATC	201
ITS1&2	gi|319432428	$(CC)_{9*}$	60	50	CCGTGAACCATCGAGTCTTT	CAGCGTCTTTTGTCTGTCCA	239
AoAS1	gi|383212088	$(TC)_4$	60	52.5	CTCATGCCCACTCCGATATT	TCAGCCTCCACGAACTCTCT	220
AG7	gi|52846592	$(AG)_{10}$	59.6	47.5	TTTTGCTCCGATCATTTTCA	CCTCTTCGTCTTCATCAGCC	194
TC7	gi|52845609	$(TC)_{15}$	59.2	44	CGCCCCGAATCAACTAATAA	TACTGCGGAGGTATGTGGGT	220

Most of the SSR markers developed in *Asparagus* were not allowed to any kind of function. All amplified SSR marker sequences were compared with non-redundant protein data base using BLASTX 2.2.27 program. Of the 14 primers, seven (50%) of the developed SSR markers were found to show significant similarities with known protein sequence, which were in close agreement with previous studies, which showed that 46 % of *L. Vannamei* ESTs sequences corresponded to known proteins (Perez *et al.* 2005). One of the positive hit was similar to ribosomal protein, two of the positive hits similar to transcription factors, one of the positive hit matched to chloroplast gene, one similar to carbon/nitrogen metabolism while two of the positive hits were matched to hypothetical proteins. The results of the present study revealed that such SSR loci showed sequence similarity with other putative functions of protein sequence, which suggested these SSR loci can be used to saturate various plants metabolic pathways (Victoria *et al.* 2011). Putative function of SSR markers are illustrated in Table 3.

Table 3. Putative function of seven amplified SSR markers.

Primer	Gene bank Accession no.	Putative function	Species	E-value	Score
ZHD1	EU200163.1	Transcription factor, putative	Ricinus communis	1e-68	235
A. synthase	X67958.1	glutamine-dependent asparagine synthetase	Triticum aestivum	0	1025
rpl32-trnL	AB613848.1	50S ribosomal protein L32	hoenix dactylifera	2e-08	59.3
matK	HQ180855.1	maturase K (chloroplast)	Hemiphylacus alatostylus	0	959
AoASI	AB673050.1	MYB-related transcription factor PHAN1	Pisum sativum	5e-123	367
rps16-trnK	AB613981.1	hypothetical protein SORBIDRAFT	Sorghum bicolor	0.001	45.4
AG7	CV459406.1	hypothetical protein VITISV_023571	Vitis vinifera	1e-28	116

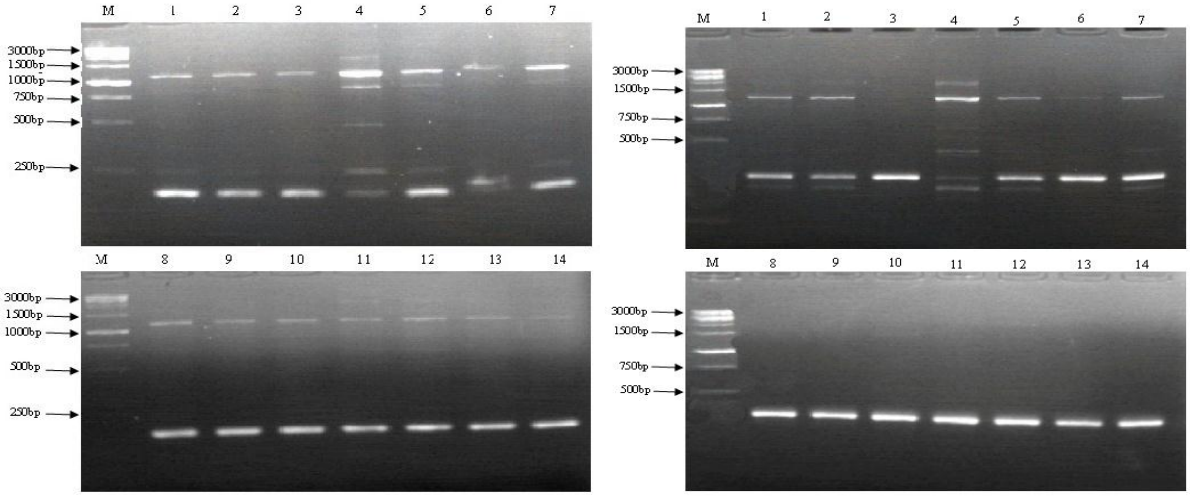
A total of 14 SSR primers was selected for amplification and polymorphism in *Asparagus* species and its cultivars (Table 1 and Fig. 1A). In the present investigations, out of total 14 SSR primers used, 10 SSR primers (AA5, TT10, TC15, TTGGT4, AG10, AA4, AT4, TT4, TC4) showed polymorphism

and produced a total of 156 bands, of which 144 (88.36 %) were polymorphic bands with an average of 14.4 alleles per primer (Fig. S1 and Table 4).

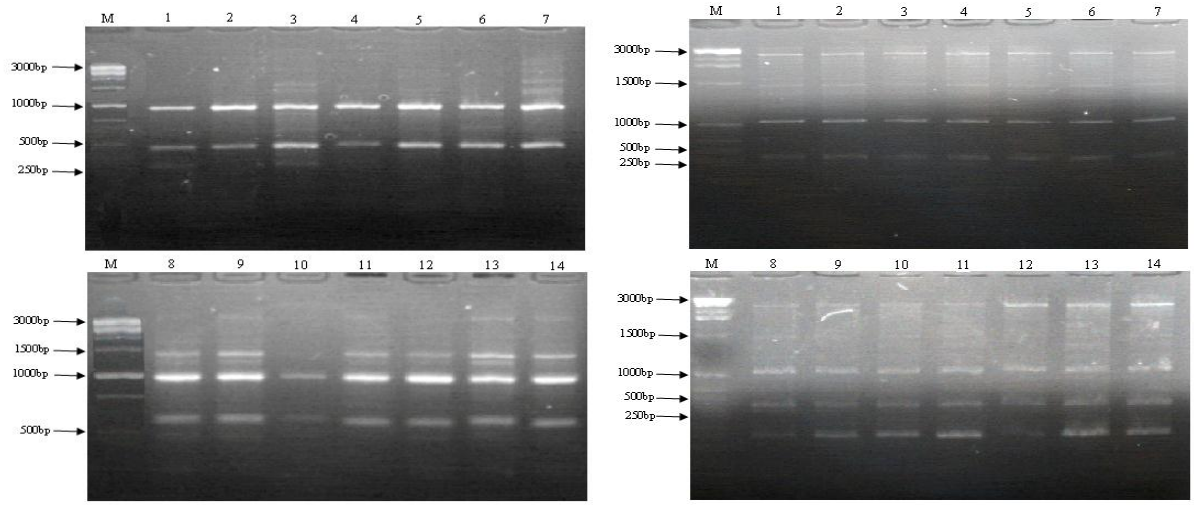
Table 4. SSR motifs and polymorphism index obtained for *Asparagus* species.

Primer	Repeat motifs	No. of bands	Polymorphic bands	Polymorphism Index
psbD-trnT	(AA)5	17	16	94.1
PetB	$(TT)_4$	11	9	81.8
ZHD1	(TT)10	11	7	63.6
TC7	(TC)15	39	39	100
A.synthetase	(TTGGT)4	16	14	87.5
AG7	(AG)10	8	7	87.5
rps16-trnK	(AA)4	4	3	75
atpB-rbcL	(AT)4	17	16	94.1
trnQ-rps16	(TT)4	17	17	100
AoAs1	(TC)4	16	16	100
Total		**156**	**144**	**92.3**
Average		**15.6**	**14.4**	**88.36**

These findings are similar to those of Dow *et al.* (1995) who observed 14.3 alleles per primer in Oak. Genetic diversity among *Asparagus* species was also studied by other investigators, Ray *et al.* (2010) who used 6 ISSR markers for 6 different species of *Asparagus*. They analyzed a total of 110 ISSR fragments with an average of 27.5 fragments per prime. Ginwal *et al.* (2011) applied 18 cpSSR markers in which 5 (27.77%) were identified yielding good amplification in *A. racemosus*. In the present study, the genetic similarity coefficients ranging from 0.52 to 0.94 reveal a high level of genetic variations among *Asparagus* species which was in close agreement with Pan *et al.* (2004) who found genetic similarity matrix using RAPD markers which ranged from 60.5 % to 88.5% for the collection of *S. spontaneum* clones. The polymorphism observed in present study by using SSR markers for *Asparagus* species and its cultivars was effective to determine genetic variation at species level (Caruso *et al.* 2008).



(PetB) (atpB-rbcL)



(A. synthetase) (ZHD1)

Fig. 4. Electrophoretic pattern of SSR product generated with primers (A) A. synthetase, (B) psbD-trnL and (C) petB (D) matK: Lane 1-14 represents 1) *A.adscenden*, 2) *A.racemosus*, 3) *A.capitatus*, 4) *A. gracilis*, 5) *A. densiflorus*, 6) *A. setaceus*, 7) *A. plumosus*, 8) *A.officinalis*, 9) *A.officinalis* cv. Appollo, 10) *A.officinalis* cv. Abril, 11) *A.officinalis* cv. Gersengum, 12) *A.officinalis* cv. Huchels, 13) *A.officinalis* cv. Para selection 14) *A. officinalis cv.* Taranga.

Cluster analysis of SSR data using UPGMA revealed *Asparagus* species and *A. officinalis* cultivars into 2 main clusters, one cluster comprised of only *Asparagus* species and other cluster comprized of *A. officinalis* cultivars (Fig. 4). Clustering, using SSR markers, showed that *A. gracelis* was quite separated from other species genotypes, because of its different genetic background, suggesting that these might be due to mutations, genetic changes and interspecific hybridization. Our result is consistent with Ray *et al.* (2010) that also cluster *Asparagus* species into 2 subclusters.

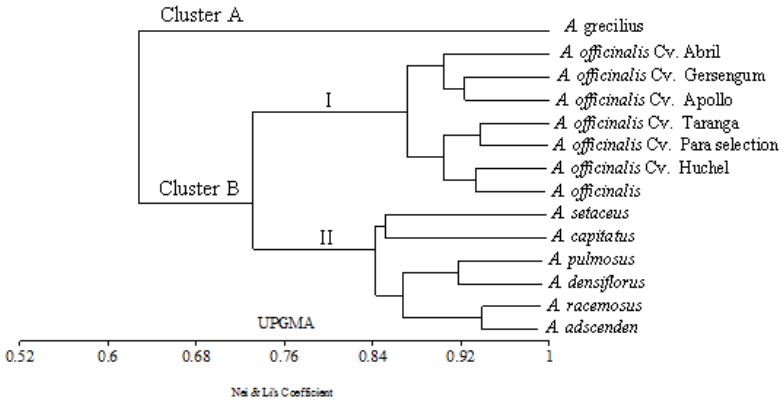


Fig. 5. Cluster analysis of *Asparagus* species using SSR markers.

Simple sequence repeat information for the transferability of one species to another species was used to determine the genome homology of the species and chance to increase success in PCR amplification utilizing heterologous primer pairs paticularly designed for the plant taxa (Pierantoni *et al.* 2004). In the present study, fourteen SSR markers were used to determine amplification and polymorphism in other medicinal plants (viz. *Saussurea costus* and *Curcuma longa* L.). Of these 14 primers, six primer pairs (42.8%) showed successful amplification in *Curcuma longa* L. and *Saussurea costus*. Our result was consistent with previous studies reported in *Jatropha curcas* and *Catheranthus roseus*. In several

studies, transferability level was found to be only 20-30 % (Tahan *et al.* 2009). Thus, SSR markers generated in the present study are satisfactory and provide a new tool for its further use in characterization and genetic variation in the related Asteraceae and Zingiberaceae families (Fig. S2).

Table S2. Genetic similarities index based on Nei and Lie Coefficient using SSR markers for *Asparagus* species and *A. officinalis* cultivars.

	A. adscenden	***A. racemosus***	***A. apitatus***	***A. gracilis***	***A. densiflorus***	***A. setaceus***	***A. plumosus***	***A. officinalis***	***A. officinalis Cv. Apollo***	***A. officinalis Cv. Abril***	***A. officinalis C. Gersengum***	***A. officinalis Cv. Huchel***	***A. officinalis Cv. Para***
A .racemosus	0.94												
A. capitatus	0.86	0.83											
A. gracilis	0.75	0.76	0.62										
A. densiflorus	0.89	0.87	0.78	0.75									
A. setaceus	0.83	0.85	0.85	0.70	0.84								
A. plumosus	0.86	0.84	0.88	0.70	0.92	0.86							
A. officinalis	0.75	0.73	0.84	0.59	0.68	0.78	0.74						
A. officinalis Cv. Apollo	0.73	0.71	0.82	0.56	0.69	0.80	0.76	0.93					
A. officinalis Cv. Abril	0.75	0.76	0.80	0.60	0.75	0.86	0.77	0.88	0.91				
A. officinalis Cv. Gersengum	0.72	0.74	0.82	0.54	0.72	0.79	0.78	0.86	0.92	0.90			
A. officinalis Cv. Huchel	0.70	0.72	0.76	0.57	0.71	0.77	0.73	0.93	0.89	0.87	0.85		
A. officinalis Cv. Para	0.63	0.64	0.71	0.52	0.63	0.73	0.69	0.89	0.88	0.86	0.88	0.92	
A. officinalis Cv. Taranga	0.63	0.64	0.71	0.52	0.63	0.73	0.69	0.89	0.88	0.83	0.84	0.92	0.94

The similarity matrix was computed using SSR markers based on Nei and Lie coefficients used in NTSYS software. In the present study, the genetic similarity coefficients ranged from 0.10 to 0.87 revealing a high level of genetic variations among *Asparagus* species, *Curcuma longa* and *Saussurea costus* which was consistent with previous study using RAPD markers, which found genetic similarity matrix ranged from 60.5 % to 88.5 % for a collection of *S. spontaneum* clones (Pan *et al*. 2004). The genetic similarity highest value was 0.87 between *A. officinalis* and *A. capitatus*, whereas the genetic similarity lowest value was 0.10 between *A. adscenden* and *Curcuma longa*. Dendogram based on UPGMA analysis grouped the *Asparagus* species, *Curcuma longa* and *Saussurea costus* into 2 major clusters. Cluster I consists of *Asparagus* species (*A. adsceden*, *A. capitatus*, *A. officinalis* and *A. racemosus).* Cluster II consists of *Curcuma longa* and *Saussurea costus*. (Table S2 and S3).

Table S3. Genetic similarities index based on Nei and Lie Coefficient using SSR markers.

	A. adscenden	***A. capitatus***	***A. officinalis***	***A. racemosus***	***Saussurea costus***
A. adscenden					
A. capitatus	0.81				
A. officinalis	0.76	0.87			
A. racemosus	0.79	0.69	0.77		
Saussurea costus	0.19	0.22	0.28	0.29	
Curcuma longa	0.10	0.12	0.17	0.17	0.67

To the best of our knowledge, this is the first study in Pakistan to use SSR markers for the determination of genetic diversity among *Asparagus* species and *A. officinalis* cultivars. For preserving this valuable plant, more *Asparagus* samples should be gathered, cultivated and domesticated in collections. Prohen *et al*. (2008) showed that Asia, Africa and Europe are the main centers for the genetic diversity of edible *Asparagus*. The results of current study showed that Pakistan could be considered as another center for the genetic diversity of *Asparagus*.

The present study showed that SSR markers developed in *Asparagus* by *in-silico* methods is more effective and laborious as compared to classical methods. These *in-silico* methods play a major role in contributing to the evolution and progress of SSR development. The present study revealed that SSR markers are sufficiently useful and powerful to assess genetic relationships and diversity analysis in *Asparagus* and *A. officinalis* cultivars. SSR method effectively discriminates *Asparagus* species and *A. officinalis* cultivars. These SSR markers will be particularly useful for evolutionary and genetic mapping studies of Asparagaceae, Asteraceae and Zingiberaceae.

ACKNOWLEDGEMENTS

We are grateful to Dr. Maryam Shinwari (Quaid-E-Azam University) and Mr. Maazullah (Agriculture Research Institute) for the facilities provided during sampling of *Asparagus* species and cultivars. The authors acknowledge the technical support of all technicians at the Department of Biotechnology and Genetic Engineering, KUST, Kohat, Pakistan.

REFERENCES

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. H. Zhang, Z. H. Zhang, W. Miller and D. J. Lipman. 1997. Gapped BLAST and PSIBLAST: a new generation of protein database search programs. *Nucleic Acids Res*. **25**: 3389-3402.

Beckmann, J. S. and J. L. Weber. 1992. Survey of human and rat microsatellites. *Genomics*. **12**: 627-631.

Bornet, B. and M. Branchard. 2001. Nonanchored inter simple sequence repeat (ISSR) markers: Reproducible and specific tools for genome fingerprinting. *Plant Mol. Biol. Rep*. **19(3)**: 209-215.

Caruso, M., T. F. Claire, L. Mikeal and Roose. 2008. EST–SSR markers for *Asparagus* genetic diversity evaluation and cultivar identification. *Mol. Breed*. **21**: 195-204.

Dow, B. D., M. V. Ashley and H. F. Howe. 1995. Characterization of highly variable (GA/CT) n microsatellites in the bur oak (*Quercus macrocarpa*). *Theor. Appl. Genet*. **91**: 137-141.

Doyle, J. J. and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Focus*. **12**: 13-15.

Dutta, I. C. 2007. *Non timber forest products of Nepal* (*identification, classification, ethnic uses and cultivation*). Institute of Forestry, Pokhara, Nepal. **44**: 436-438.

Eisen, J. A. 1999. Mechanistic basis for microsatellite instability. In: D. B. Goldstein and C. Schlotterer (eds). *Microsatellites: Evolution and applications*. Oxford University Press, Oxford, UK., pp. 34-48.

Ginwal, H. S., M. Neha, S. M. Shalini, B. Santan and B. Preeti. 2011. Genomic DNA isolation and identification of Chloroplast microsatellite markers in *Asparagus racemosus* Wild. Through cross amplification. *Ind. J. biotechnol*. **10**: 33-38.

Gupta, S. and M. Prasad. 2009. Development and characterization of genic SSR markers in *Medicago truncatula* and their transferability in leguminous and non-leguminous species. *Genome*. **52**: 761-771.

Hafezi-Shahroodian, S., D. Azadfar, H. Soltanloo and S. Ramezanpour. 2011. Genetic variability in natural Iranian populations of *Cupressus sempervirens* var. horizontalis in Caspian Sea coastward assessed by SSR markers. *Plant Omics. J*. **4**(1): 19-24.

Huang, X. and A. Madan 1999. CAP3: a DNA sequence assembly program. *Genome Res*. **9**: 868-877.

Koppolu, R., D. U. Hari, L. D. Sangam, A. H. David and K. V. Rajeev. 2010. Genetic relationships among seven sections of genus Arachis studied by using SSR markers. *BMC Plant Biol*. **10**: 15.

Kubitzki, K. and P. J. Rudall. 1998. Asparagaceae. In: K. Kubitzki (ed.). *The families and genera of vascular plants*. *Monocotyledons: Lilianae (except Orchidaceae)*. Vol. 3. Springer-Verlag, Berlin, Germany., pp. 125-126.

Kumpatla, S. P. and S. Mukhopadhyay. 2011. Mining and survey of simple sequence repeats in expressed sequence tags of dicotyledonous species. *Genome*. **48**: 985-998.

Mandal, S. C., C. K. A. Kumar, M. S. Lakshmi, S. Sinha, T. Murugesan, B. P. Saha and M. Pal. 2007. Antitussive effect of *Asparagus racemosus* root against sulfur dioxide induced cough in mice. *Fitoterapia*. **71**: 686-689.

Nei, M. and W. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleasis. *Proc. Natl. Acad. Sci*. **76**: 5269-5273.

Pan, L., X. Qiuju, Q. Zhiwu, L. Honggao, K. Weidong and D. Yi. 2010. Development of Novel EST–SSRs from Sacred Lotus (*Nelumbo nucifera Gaertn*) and Their Utilization for the Genetic Diversity Analysis of *N. nucifera*. *J. Hered*. **101**(1): 71-82.

Perez, F., O. Juan, Z. Mariuxi, G. Cesar, C. Jorge and F. A. M. J. Volckaert. 2005. Development of EST-SSR Markers by Data Mining in Three Species of Shrimp: *Litopenaeus vannamei*, *Litopenaeus stylirostris*, and *Trachypenaeus birdy*. *Marin Biotechnol.* **7**: 554-569.

Pierantoni, L., K. H. Cho, I. S. Shin, R. Chiodini, S. Tartarini, L. Dondini, S. J. Kang and S. Sansavini. 2004. Characterization and transferability of apple SSRs to two European pear F1 populations. *Theor. Appl. Genet.* **109**: 1519-1524.

Portis, E., I. Nagy, Z. Sasva, A. Stagelri, L. Barchi and S. Lanteri. 2007. The design of *Capsicum* spp. SSR assays via analysis of *In silico* DNA sequence, and their potential utility for genetic mapping. *Plant Sci*. **172**: 640-648.

Prohen, J., F. Nuez and M. J. Carena. 2008. Handbook of plant breeding. Springer Publlishing. 364p.

Ray, S., J. M. Madhumita and M. Sandip. 2010. Phylogenetic Relationship among six Economically Important Species of *Asparagus* Utilizing RAPD, ISSR and Isozyme Polymorphism. *Biores. Bull*. **3**: 153-160.

Rozen, S. and H. J. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. In: S. Krawetz and S. Misener (eds.). *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ., pp. 365-386.

Shasnay, A. K., M. P. Darokar, D. Sakia, S. Rajkumar, V. Sundaresan and S. P. S. Khanuja. 2003. Genetic diversity and species relationship in *Asparagus* spp. using RAPD analysis. *J. Med. Arom. Plant Sci.* **25**: 698-704.

Tahan, O., Y. P. Geng and L. Y. Zeng. 2009. Assessment of genetic diversity and population structure of Chinese wild almond, *Amygdalus nana*, using EST- and genomic SSRs. *Biochem. Syst. Ecol*. **37**: 146-153.

Thiel, T., W. Michalek, R. K. Varshney and A. Graner 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, R. K., T. Thiel, N. Stein, P. Langridge and A. Graner. 2002. *In silico* analysis on frequency and distribution of microsatellites in ESTs of some cereal species. *Cell Molecul. Biol*. **7**: 537-546.

Victoria, F. C., C. Luciano and C. O. Antonio. 2011. In silico comparative analysis of SSR markers in plants. *BMC Plant Biol*. **11**: 15.

Wen, M., W. Haiyan, X. Zhiqiang, Z. Meiling, L. Cheng and W. Wenquan. 2010. Development of EST-SSR and genomic-SSR markers to assess genetic diversity in *Jatropha Curcas* L. *BMC Res Note*. **3**: 42.