Journal of Plant Physiology & Pathology

Research Article

A SCITECHNOL JOURNAL

Development and Validation of EST-SSR Markers in *Gymnema sylvestre* R.Br.

Kuldeepsingh A. Kalariya*, Lipi Poojara, Dipal Minipara, Parmeshwar Lal Saran, Ram Prasnna Meena, Sinora Macwan, Ponnuchamy Manivel

Abstract

Known for its anti-diabetic properties, the *Gymnema sylvestre* belonging to the family *Asclepiadaceae* is native to South-Indian forests. It is also found in tropical Africa and in Australia. All arial parts of this plant contain alkaloids, flavones and saponins, but the leaves are mainly used for its medicinal properties. The diversity based on morphological, anatomical and chemo-profiles, and Randomly Amplified Polymorphic DNA (RAPD) and Inter Simple Sequence Repeats (ISSR) markers is reported however, the genetic diversity at molecular level through the most efficient, the expressed sequence tags-SSR (EST-SSR) markers is lacking in this plant. A huge transcriptome data was generated and 5276 SSRs loci were identified in this study. The frequency of SSRs in *Gymnema sylvestre* was 1/12.16 kB. The AAG/CTT repeats were nearly ten times higher than the CCG/CGG repeats. Total 40 pairs of primers were synthesized, and 27 primers gave polymorphic amplification in 25 genotypes of *G. sylvestre* collected from different parts of India. Genotypes DGS 16 and DGS 34 were the most dissimilar genotypes. This was the first study revealing genetic diversity and high polymorphisms in *G. sylvestre* with the help of EST-SSR markers having a higher transfer rate of 67.5%.

Keywords: EST-SSR, *Gymnema sylvestre*, PAGE, Transcriptome, Genetic Diversity

Introduction

Gymnema sylvestre is a climbing herb belonging to the family *Asclepiadaceae* and Class Dicotyledoneae is a native to South-Indian forests. It is also found in tropical Africa and in Australia [1]. It has many therapeutic applications in Ayurvedic system of medicine but, it is mainly known for having anti-diabetic properties. Popularly it is known as 'gurmar' or 'sugar destroyer'. Other medicinal uses of *G. sylvestre* includes lowering serum cholesterol, triglycerides blood glucose level (hypoglycaemic or antihyperglycemic), weight loss, hypolipidaemic, liver diseases, constipation, stomach ailments and water retention. It is also used for maintaining normal blood pressure, tachycardia or arrhythmias, prevention of dental caries, cataract and as anticancer-cytotoxic agent.

Leaves, flowers and fruits of this plant contain alkaloids, flavones and saponins.The stigmasterol having main principle bioactive compounds viz. gymnemic acids, gymnemasides, gymnemagenin, gurmarin, gymnemosides, gymnemanol, gymnemasins, gypenoside and conduritol which act as therapeutic agent and play vital role in many therapeutic applications. Gymnemic acids from the extract are shown to stimulate insulin release from the pancreas and thereby, are thought to be responsible for its antidiabetic activity. Diversity based on morphological and anatomical traits and chemo-profiling among some natural population of gymnema genotypes is documented. Recently, transcriptomic data based putative pathway leading to biosynthesis of polyoxypregnane [2] gymnemic acid [3] identification of micro RNAs [4] and characterization of SQS gene [5] in *G sylvestre* is reported. Variation based on photosynthetic efficiency in genotypes of *G. sylvestre* is also documented [6]. Thus, gymnema genotypes are studied for morphological, anatomical, chemo and transcriptome profiling. However, diversity at molecular level is explained through RAPD and ISSR markers only in gymnema genotypes [7-10] with only few locations with small sample size.

Development of molecular markers has an important role in crop improvement programme. Exploitation of diversity at molecular level through more efficient molecular markers in such a high value medicinal plant is need of the hour.

PCR based molecular markers were used for characterizing and evaluating genetic diversity. Recently, SSR markers have been widely used in genetic diversity analysis, fingerprint construction, and molecular marker-assisted breeding because of their repeatability, high polymorphism, and codominant inheritance. EST-SSRs are not available in gymnema genotypes and hence, the main objective of this work was to develop EST-SSR markers to study the genetic diversity among twenty-five genotypes of *G. sylvestre*.

Materials and Methods

Plant Materials used for RNA isolation, cDNA library preparation and Quantity and Quality check (QC)

Leaf sample ofDGS 22 genotype of *G. sylvestre* were collected during the month of November 2017. From this sample, total plant RNA was isolated through Norgen total RNA isolation kit following manufacturer's instructions. To calculate RNA Integrity Number (RIN) of the isolated total RNA, Agilent RNA 6000 Nano chip was used. Following standard protocol of Truseq stranded total RNA ribo zero library preparation kit (Illumina), adapter ligated libraries were prepared which was sequenced using Illumina Hiseq 2500 platform (Illumina Inc., San Diego, CA, USA) (2 × 150 PE) using standard protocol.

Transcriptome sequencing

Based on the Qubit concentration and mean peak size, the library was loaded into Illumina HiSeq 2500 platform for cluster generation. Sequencing through 2x150 Paired-End sequencing facilitated the template fragments to be sequenced in both the forward and reverse directions. The library molecules bind to complementary adapter oligos on paired-end flow cell. The adapters were designed to facilitate selective cleavage of the forward strands after re-synthesis of the reverse strand during sequencing process and the copied reverse strand was

***Corresponding author:** Kuldeepsingh A Kalariya, ICAR-Directorate of Medicinal and Aromatic Plants Research, Boriavi, Anand-387310, Tel: +91 9427710751; E-mail: Kuldeep.Kalariya@icar.gov.in, Kuldeep_ka@yahoo.co.in

Received: May 17, 2019 **Accepted:** June 14, 2021 **Published:** June 21, 2021

then used to sequence from the opposite end of the fragment.

De novo assembly and unigene prediction from transcripts

Trimmomatic-0.36 [11] was used to filter raw data and process at Min Phred Score (QV) 20. Trinity software [12] at default parameter. For the preparation of *de novo* assembly containing high quality reads, Trinity software [12] at default parameters was used. Contigs were prepared by assembling reads and minimally overlapping contigs were clustered into connected components. Unigene prediction from transcripts was done through CD-HIT package using CD-HIT-EST executable to remove the shorter redundant transcripts.

CDS prediction, SSR search and primer designing

Trans decoder [13] at default parameters was used to predict CDSs from unigenes. Minimum of 100 amino acids for the encoded protein length was set along with homology search from swiss-prot and pfam databases. SSR were identified from CDS sequences with the MISA perl script [14] Criteria for SSR repeat length for di-nucleotides was kept at least six while for other tri, tetra, penta and hexa nucleotides repeat criteria was five and more than it. Flanking regions of SSRs, 75 base pair upstream as well as downstream were used for primer designing. Primer pairs were designed using Primer 3 [15]. The major parameters for primer pair design were primer length of 18–22 bases (optimal 20 bases), PCR product size of 150-200 bp, GC content of 40–80% (optimal 50%), and annealing temperatures of 55–60°C (optimal 55°C). Based on these parameters, 40 pairs of primers were designed and synthesized for identification of polymorphism between different *G. sylvestre* genotypes.

Marker validation in various genotypes of *Gymnema sylvestre*

Total 25 genotypes of *G. Sylvestre* originally collected from different parts of India in year 2009 used for genomic diversity analysis in this study are listed in Table 1. These plants were maintained with standard package of practices of cultivation at the experimental field of ICAR-Directorate of medicinal plant and Aromatic plant research, Boriavi, Anand, Gujarat, India. A total of 40 EST-SSR markers were validated using Genomic DNAs of these 25*G. sylvestre* genotypes. Cetyl trimethyl ammonium bromide (CTAB) method was used to extract DNA from leaves of these genotypes. For PCR amplification, 10-µL reaction mixtures containing 20 ng of template DNA, 2× Dream Taq Green master mix (Thermo scientific) and 0.25 µM of each primer was used. Thermal cycling was carried out on Bio-Rad make Thermal Cycler. The PCR steps used were a pre-denaturing (95°C for 5 min) followed by denaturing (95°C for 30s), annealing (55-60°C for 45s), extension (72°C for 45s) for 35 cycles, and a final extension at 72°C for 10 min. Amplified PCR products were initially visualized to check respective product size on 1.2% agarose gel. Only agarose confirmed products was separated on 8% denaturing polyacrylamide gel using a vertical electrophoresis device.Identification of EST-SSR bands was performed using the silver staining method. Different primer pairs resulted in different number of bands and hence, only the first two bands close to the product size were considered as co-dominant markers. Based on the locus observed on the gels, well-defined bands were scored as 1 (present) or 0 (absent) in a binary matrix prepared on excel (Microsoft office 2007) work sheet. These data set was used to prepare dendrogram following Wards dissimilarity method using default parameters in Past (4.0). This method utilizes the distance between central points in each cluster and results in nicely balanced clusters and hence, preferred for cluster analysis in this study.

Table 1: Summary of EST-SSRs present in the *Gymnemasylvestre* transcriptome.

Summary	Number
Total number of sequences examined	71676
Total size of examined sequences (bp)	64186092
Total number of identified SSRs	5276
Number of SSR containing sequences	4448
Number of sequences containing more than 1 SSR	692
Number of SSRs present in compound formation	417

Results and Discussions

Development process of *Gymnema* EST-SSR marker

Total 9.14 Gb raw data leading to 9,143,594,400 raw reads are deposited at NCBI under Project SUB2977090 as SAMN07528738. After filtration of data at QV 20, 8.48 Gb high quality data accounting 8,484,325,024 quality readswere generated which were further used to generate unigenes. A total of 112,583 unigenes were obtained with an N50 length of 2532 bp. Prediction of CDSs using Trans decoder at default parameters with the encoded protein length set to a minimum of 100 amino acids and homology search with swiss-prot and pfam databases resulted in 71,676 CDSs with an N50 length of 896 bp. From these CDSs, the tandem repeats of nucleotide motifs of the sizes 2-6 bp were identified using the MISA perl script which resulted in identification of a total of 5276 SSRs loci.

The occurrence rate and distribution of Gymnema SSR loci

There were 4448 CDSs containing 1 SSR locus and692 CDSs had more than 1 SSR locus. Total number of SSRs with 75bp flanking were 2685. The summary of EST-SSRs along with 75 bp up and down stream sequences are presented in Table 2. Using these information, total 40 pairs of primers were designed with the help of PRIMER 3 software.

The proportion and types of repeat motifs

The statistical analysis on repeat motifs of all SSR loci showed that, in the *G. sylvestre* EST-SSRs, the type of repeat nucleotide forming SSR was 2 to 6, but the occurrence rate of different SSR types was different. Total 41 different types of repeat sequence motifs were detected in SSRs loci, including three types of Di-nucleotide repeat motifs (total 1534), ten types of Tri-nucleotide repeat motifs (total 3651), ten types of Tetra-nucleotide repeat motifs (total 52), five types of Penta-nucleotide repeat motifs (total 11) and thirteen types of Hexa-nucleotide repeat motifs (total 28). As depicted (Figures 1-3) the most frequent repeat motifs were Tri-nucleotiderepeats (69.20%) followed by di-nucleotide (29.07%), tetra-nucleotide (0.99%), hexa-nucleotide repeats (0.53%) and penta-nucleotide repeats (0.21%). The results above indicated that the major repeat sequence type of *G. Sylvestre* SSR loci was tri-nucleotide repeats. The tri-nucleotide repeat motif AAG/CTT was the most abundant (26.79%) among the tri-nucleotide repeats which was followed by ATC/ATG (17.91%) and ACC/GTT (11.50%).

The Times of SSR repeat motif repetition

The number of times of SSR motifs repetition of *G. sylvestre* splicing sequences ranged from from 5 to 49 (Table 3, those >15 times are excluded). Total 5230 SSRs accounting more than 99% of the repetitions were ranged between 5 to 15 repetitions and only and 46

Table 2: Characteristics of di, tri, tetra, penta and hexa nucleotides repeat motifs in the *Gymnemasylvestre* transcriptome.

Repeat type	No. of EST-SSRs	Proportion in all SSRs (%)	Repeat motif	Total number	Proportion (%)
Di-nucleotide	1534	29.07	AC/GT	178	11.60
			AG/CT	696	45.37
			AT/AT	660	43.02
Tri-nucleotide	3651	69.20	AAC/GTT	188	5.15
			AAG/CTT	978	26.79
			AAT/ATT	225	6.16
			ACC/GGT	420	11.50
			ACG/CGT	92	2.52
			ACT/AGT	121	3.31
			AGC/CTG	487	13.34
			AGG/CCT	387	10.60
			ATC/ATG	654	17.91
			CCG/CGG	99	2.71
Tetra-nucleotide	52	0.99	AAAC/GTTT	3	5.77
			AAAG/CTTT	2	3.85
			AAAT/ATTT	25	48.08
			AAGG/CCTT	2	3.85
			ACAG/CTGT	1	1.92
			ACAT/ATGT	13	25
			ACTG/AGTC	1	1.92
			AGAT/ATCT	1	1.92
			ATCC/ATGG	1	1.92
			ATGC/ATGC	3	5.77
Penta-nucleotide	11	0.21	AAAAC/GTTTT	2	18.18
			AAAAG/CTTTT	1	9.09
			AAATC/ATTTG	3	27.27
			AAGAG/CTCTT	4	36.36
			AGAGG/CCTCT	1	9.09
Hexa-nucleotide	28	0.53	AAAGGG/CCCTTT	1	3.6
			AACGGG/CCCGTT	3	10.7
			AAGGAG/CCTTCT	3	10.7
			AATCCC/ATTGGG	1	3.6
			AATTCC/AATTGG	3	10.7
			ACAGGC/CCTGTG	1	3.6
			ACCGCC/CGGTGG	4	14.3
			ACCTCC/AGGTGG	3	10.7
			ACGGAG/CCGTCT	1	3.6
			AGATGG/ATCTCC	4	14.3
			AGCAGG/CCTGCT	2	7.1
			AGCCTC/AGGCTG	1	3.6
			ATCGCC/ATGGCG	1	3.6

motifs had repetition >15, accounting for 0.88% only. Irrespectively of the type of SSRs, the number of repetitions was maximum (5 times) accounting total 2119.

Processing Results using Primer 3.0 Software.

PRIMER3.0 software was used to design primers for 40 SSRs sequences. As per the data the tri-nucleotide repeats were the maximum in gymnema and hence, out of total 40 primers, we had 30 primers representing tri-repeats SSRs, 4 primers representing Di-repeats SSRs, 3 primers representing hexa-repeats SSRs, 2 primers representing tetra-repeats SSRs and 1 primer representing penta-repeats SSRs (Table 4). These primers were qualified with no possibilities of primer-dimer complex and self-complementary.

Marker validation

From these 40 primer pairs the amplification was obtained in 27 pairs of EST-SSR primers. Based on the banding pattern there were total 25 polymorphic primers. The dendrogram following Wards dissimilarity method was prepared (Figure 2) and there were 2 separate main clusters named A and B. There were 9 genotypes in cluster A, and 16 genotypes were there in cluster B. Cluster A and B were further divided in Clusters Aa, Ab and Ba, Bb. Genotypes collected from Gujarat and Karnataka were there in both the clusters A and B. Cluster Ba and Bb contained most of the genotypes collected from different regions of Andhra Pradesh and Tamil Nadu. According to the dendrogram it can be presumed that the genotype DGS 16 (collected from Veerappaaiyanar, Teni, Tamil Nadu, India) and DGS 34 (collected from Eastern Ghats Forest, Jamshedpur,

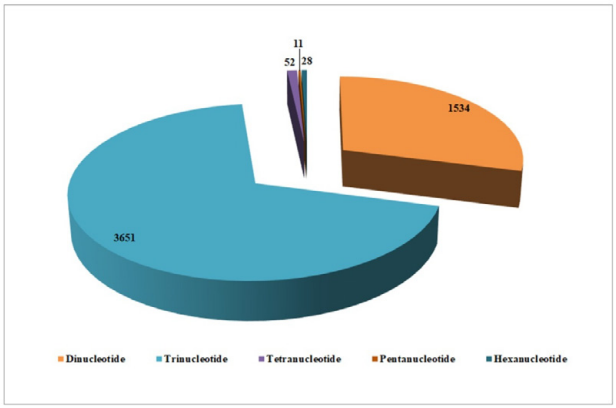

Figure 1: Pie chart showing distribution of SSRs in *Gymnema sylvestre*.

Jharkhand) were the most dissimilar genotypes. Genotypes DGS 28 and DGS 30 both were collected from Madhya Pradesh and both were falling under the same cluster Ab also. However, genotypes DGS 15 (collected from Tumkur Road, Karnataka) and DGS 8 (collected from Dharmasthala Road, Udipi, Karnataka) thoughboth collected from the same state of Karnataka plotted in different clusters but with close similarity between these two clusters. The minimum span tree of the principal component analysis (PCA) with showed that there was no specific group of one particular state (Figure 3). It is interesting to note that the genotype collected for Rajasthan state was at the central point of the span tree.

Diversity can be categorized into genetic diversity, species diversity and ecosystem diversity [16,17]. The conservation of biological diversity should emphasise on preventing the disappearance of genetically distinct populations rather than the sole prevention of the extinction of species. This will also lessen the risk of extinction,

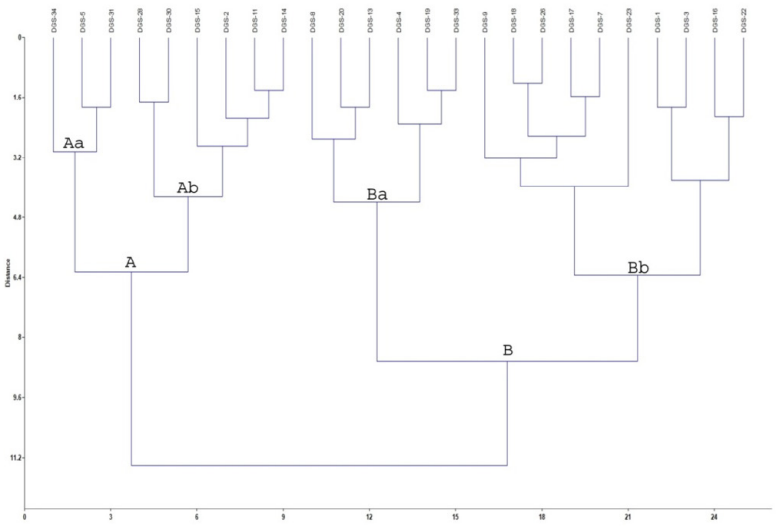

Figure 2: Dendrogram showing Wards dissimilarity among 25 different genotypes *Gymnema sylvestre*.

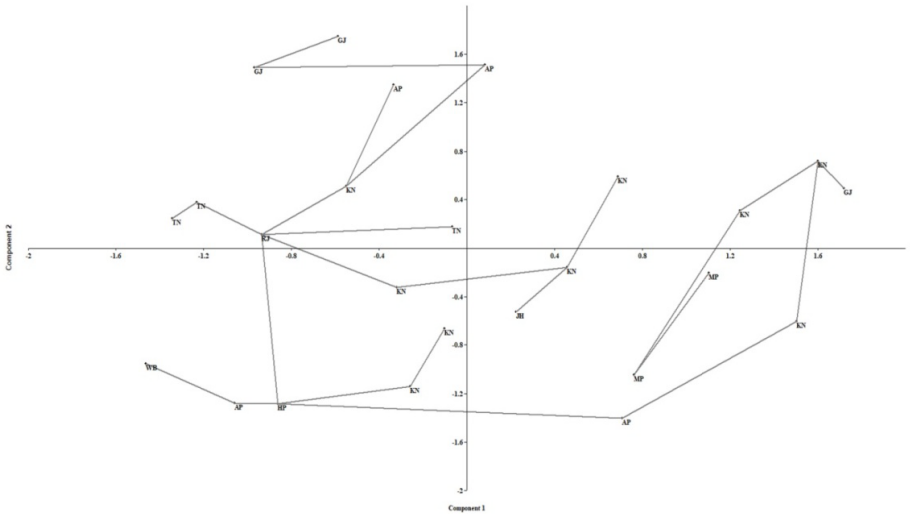

Figure 3: PCA analysis in *Gymnema sylvestre* genotype.

Table 3: Contains details from where 25 different *Gymnema sylvestre* genotypes were collected across India.

Sr.no	Genotype	Geographical location
1	DGS-1	Waghai,Dang, Gujarat
2	DGS-2	Waghai,Dang, Gujarat
3	DGS-3	Waghai, Dang, Gujarat
4	DGS-4	Kalyani, West Bengal
5	DGS-5	Agumbe, Shimoga, Karnataka
6	DGS-7	Manglore Road, Chikmanglore , Karnataka
7	DGS-8	Dharmasthala Road, Udipi, Karnataka
8	DGS-9	Karkal,Udipi, Karnataka
9	DGS-11	Kalsa Road, Shimoga, Karnataka
10	DGS-13	N.R. Pura Road, Shimoga, Karnataka
11	DGS-14	Karnataka
12	DGS-15	Tumkur Road, Karnataka
13	DGS-16	Veerappaaiyanar,Teni, Tamil Nadu
14	DGS-17	Anakaraipatti, Madurai, Tamil Nadu
15	DGS-18	Tanipurai, Virudhnagar, Tamil Nadu
16	DGS-19	Central Region of Eastern Ghats, Visakhapatnam, A.P
17	DGS-20	Central Region of Eastern Ghats, Visakhapatnam, A.P
18	DGS-22	Central Region of Eastern Ghats, Visakhapatnam, A.P
19	DGS-23	Central Region of Eastern Ghats, Visakhapatnam, A.P
20	DGS-26	BankiSisarvula, Udaipur, Rajasthan
21	DGS-28	Khanpura Tola, Raisen, Madhya Pradesh
22	DGS-30	Agera Kheda, Dewas, Madhya Pradesh
23	DGS-31	Badra Dam, Shimoga, Karnataka
24	DGS-33	Behal village, Solan, Himachal Pradesh
25	DGS-34	Eastern Ghat Forest, Jamshedpur, Jharkhand

even in a longer time perspective as the ability of a population to adapt to the environmental changes depends on genetic variability or diversity of the population [18]. *G. sylvestre* has a very high medicinal value which makes it one of the highly marketed plants. It is therefore a very important plant species from the medicinal and economical perspective. The interactions of various processes such as long-term evolutionary history of the species which includes habitat fragmentation, population isolation and shifts in distribution along with gene flow, genetic drift, mutation, and natural selection all together decides the genetic makeup of plant populations [19]. In a genetic diversity study, a meaningful conclusion drawing estimate of heterozygosity is possible provided it has enough loci generated even in small sample size [20]. Because of their repeatability, high polymorphism, and codominant inheritance, the SSR markers have been widely used in genetic diversity analysis, fingerprint construction, and molecular marker-assisted breeding [21-24]. The traditional methods for developing SSR markers are not very efficient [25] and therefore, next-generation sequencing technologies can generate massive data, which are good resources for developing SSR markers in many species including *G. sylvestre*. Nevertheless, SSRs in *G. Sylvestre* have not been reported previously because of lack of transcriptome or genome sequences. There were 4448 CDSs containing 1 SSR locus and 692 CDSs had more than 1 SSR locus. Total number of SSRs with 75bp flanking were 2685. The summary of EST-SSRs along with 75 bp up and down stream sequences are presented in Table 2. Using these information, total 40 pairs of primers were designed with the help of Primer 3 software.

The frequency of SSRs in *G. sylvestre* is 1/12.16 kB, excluding the mononucleotide repeats. It was close to Arabidopsis (1/13.83) [26] but, significantly lower than wheat having 1/5.46 [27] and *P. Violascens* having 1/4.55 kB [28]. The SSR frequency can be affected by various factors, including software, search criteria as well as species properties. Tectonic activities and climate fluctuations during species evolutionary history contribute to the production and accumulation of genetic variations. The high frequency of SSRs detected in *G. sylvestre* may be due in part to the long and complex evolutionary process of *G. sylvestre.*

In this study, we identified 4448 EST-SSR loci, designed 40 primers and report 27primers with polymorphic locis. Thus, we have successfully made diversity analysis of 25 genotypes of *G. Sylvestre* collected from different location in India. Clear bands were generated and higher transfer rate of about 67.5% was detected compared with *Elymus sibiricus* (22.40%) [29] *Chrysanthemum nankingense* (20%), *Juglans mandshurica* (30.8%) [30] and *Neolitsea sericea* (16.3%) [31] and comparatively lower than the transfer rate observed in *Phyllostachys violascens* (100%) [28].

This study also reports different types of EST-SSR repeat motifs in *G. sylvestre* however, with disorderly distribution. Tri repeats were the most predominant and they were accounting for 69.20 % of total SSR repeats. In addition, the proportion of Di, tetra, Penta, Hexa repeat motifs were significantly lower.In the EST-SSR loci, each tri-nucleotide repeat motif codes a specific amino acid, which plays an important role in various cellular, biological, and metabolic processes in plants. The percentage of tri-nucleotide motifs AAG/CTT, which codes for leucine and lysine was the highest (26.79%) followed by isoleucine and methionine coding repeats ATC/ATG (17.91%). Among tri nucleotides AAG/CTT was the most predominant accounting for 26.79% of the total SSR repeat motifs. The value is comparatively higher than the monocot plants like Taro [32], Bamboo [28], Rice and Maize [26] in which the tri repeat SSR motif values were about 5.91%. Previous studies showed that the tri-repeat type of CCG/CGG was a rare motif in dicotyledonous plants but the most abundant repeat type among the tri-repeats in monocots [33,32] In our study in *G. sylvestre* which is a dicot plant, CCG/CGG repeats was only 2.71% whereas AAG/CTT repeats were maximum of all tri SSR repeat motifs. This is in support of the available data base which shows high CCG/CGG repeats in monocot plants and AAG/CTT repeats are common in Dicot plants [34].

To device management strategies for conservation of plant, the complete knowledge of genetic variation within and among populations of plant species is essential. But the genetic diversity of *G. sylvestre* has been relatively unclear. ISSR and RAPD markers were used to study the genetic diversity and population genetic structure in *G. sylvestre*. Several studies have been carried out on genetic diversity in India. The study on 18 samples of *G. Sylvestre* from Kerala using 15 RAPD primers has revealed high polymorphism [7].Similarly, polymorphism on 11 progenies of *G. sylvestre* from Uttar Pradesh using 40 RAPD primers have been reported [9]. In Maharashtra genetic diversity was carried out on 22 accessions of *G. sylvestre* using ISSR and RAPD markers that resulted high level of gene differentiation [8]. Highpolymorphism is reported on 5plants samples from Haryana using ISSR marker [10]. It is interesting to note that all the above authors have obtained high genetic diversity within the populations of *G. Sylvestre*. However, available genetic diversity reports explored few locations, small samples size, and RAPD fingerprinting, no elaborate data on population diversity is yet available. Recently, transcriptomic data based putative pathway leading to biosynthesis of polyoxypregnane is proposed [2]. However, as far as the diversity

Table 4: Details of 40 polymorphic primers synthesized from 5276 identified SSR.

Sr.No	CDS ID	SSR	CDS_SIZE	Start	End	Sequences with flanking region	Forward	Reverse	Product size
1	CDS_12917	(GTCTCC)6	1302	81	116	attgcgaagccaccacatagctcttgctccgtcgccagaccactacaaccat-gattcttggtggtttccttttctGTCTCCGTCTCCGTCTCCGTCTC-CGTCTCCGTCTCCtgtggctgcagccttttacagtcaacaaattaca-cagacccgtctacgtcatctccaccttccaaattcccggt	agccacca-catagctcttgc	tggaaggtggagat-gacgta	169
2	CDS_53451	(GAAAGG)5	357	106	135	gcgaagagaagagcttgttccaagatcaacaactcagcacagaacagag-gcagtggcgaaaggaacaaggatagaGAAAGGGAAAGGGAAA GGGAAAGGGAAAGGcaaagggataggaaggggagagagta-aaggcacgtgaatgtgataggggaagggaatctgacagggaacgagaa	cgaagagaagagctt-gttcca	cccttcccctatca-cattca	159
3	CDS_2152	(TCCACC)6	666	238	273	gtcagttgcatgggtcaggtgaagagaaacagcaaggtcattggattttc-gacgccttacagactcacctcttctTCCACCTCCACCTCCACCTC-CACCTCCACCTCCACCgccaccaacacttcaagaaa-caacccatctcacggtaatctcaaatatgtgaagctcaaacgcttcttttccggc	gcaaggtcattg-gattttcg	ggaaaagaagcgttt-gagctt	152
4	CDS_40417	(AAACA)5	339	109	133	tgtgtggaaacagagcttcccacgcctttatatataaatcaacatcttcct-caagaacagaaccagtgttcaaccAAACAAAACAAAACAAAA-CAAAACAaaaccaaacccagcaataaaaatacaaaaatgccttgcctt-gatatatcaactaatgtaaacttggaagaagttg	agagcttcccac-gcctttat	aacttcttccaagtt-tacattagttga	163
5	CDS_45606	(TACA)5	384	243	262	ctcggagtcagccctggaaacaatctggactgtgccaaccaaaggccattt-gcttagatcactactttacatatgTACATACATACATACATACAg-gtctcctaataatatccctgcaaataaatatatattgctgcggcttgtgttttatg-gcagcaataaacaaagag	agtcagccctggaaa-caatc	tgctgccataaaaca-caagc	153
6	CDS_63713	(ATTT)5	1560	1192	1211	acttttctgggtaagcctggcatcatcatagttgacgctatttttaagtatc-gctttcaggttatgagtcatcaaATTTATTTATTTATTTATTTtctttttg-gtctttctgggaaagaaaccttgaccgctctttcagatgatagtttg-aagctcttcacccgtcttta	ggtaagcctggcat-catcata	cgggt-gaagagcttcaaacta	155
7	CDS_43444	(AGA)11	2046	1832	1864	aattagaagcccgacaacattatgagaaaatggaaaaagaactatctttctct-gaatcggggaacttgaatcaacCAGCAGCAGCAGCAGCAGtag-gaagcaatgaatgcagctcagagtcggtggaggatgagaatgagatg-gaagaggaagccgaatgtagagggc	gcccgacaacattat-gagaaa	attcggcttcctcttc-catc	151
8	CDS_2712	(ATC)10	819	137	166	tcttgttcttgatcatttcccagcaatccatggccattatttcagctcgaagtctcgg-gaatccatccacagtggATCATCATCATCATCATCATCATCAT-CATCgtcatcattatgcccaacttgcattctccttttacatagtagatgtgcttg-cagaggagaagcatcaatatcctt	atcatttcccagcaatc-cat	gatgcttctcctctg-caagc	160
9	CDS_17686	(AAG)10	1716	1039	1068	gagattcctgaacttatccccattgataatggtgagagtggagttgcagtt-gtgacagaagattctcagctgaaaAAGAAGAAGAAGAAGAA GAAGAAGAAGAAGgcaaataaagacagcataaataatgata-agaaggcaactgggaaaagtggggttgctggcaataatctcgaggaa	tcctgaact-tatccccattga	cgagattattgccag-caacc	170
10	CDS_57493	(ATG)9	1923	1181	1207	agcatgaagagatttccaataacagctatgaagaagaatataatggatatg-gctcagaggataatgagggtaggtATGATGATGATGATGAT-GATGATGATGctgctgctgctgatggtgatgcagatgtagaagaaca-gatagaagatcttggggttgtggataatgatgattcaa	gcatgaagagatttc-caataaca	tcattatcca-caaccccaaga	168CDS_45501
11	CDS_45501	(TCC)8	543	144	167	gaaccaaagttaccatcttggtctgaataatagcagaagcagttgttac-caaatttctttggtaccggtggctctTCCTCCTCCTCCTCCTCCTCC TCCtcttcggcaacagccacaacaacaacaggcggcggtgggcggaga-actaattgttcgacccggagagcaatttga	ccaaagttaccatcttg-gtctg	tctccgggtcgaa-caattag	162
12	CDS_49932	(TTC)8	1089	186	209	cactgcctcccaatcttcttcttctcccaaaagactcacatcaatccgtcccact-gttgtttcagcaaaacttaaTTCTTCTTCTTCTTCTTCTTCTTC cctccctagtctcatcagagaccagcctgtttttgctgcccctgctcccgtcat-caccccccattctgagagaaga	cact-gcctcccaatcttctt	ctctcagaatgggggt-gatg	170
13	CDS_67863	(GCG)8	450	191	214	gcggaggagtctctaagagggttaccaccaccgcaaatcgcacca-cagctctggctttggctgtcgtcaccatgaGCGGCGGCGGCGGCG-GCGGCGGCGgcaataacactggaaaatgcagcgccgagtgtttgt-gcccaaatttcacggctacggtttcggcatcttcttctt	ggttaccaccaccg-caaat	aagaagatgcc-gaaaccgta	152
14	CDS_21921	(TCT)8	357	213	236	gtcatcattatcatcaccaccaacaactacgagga-caaaggtaaagatgaaaagcaaccgtcttaaagttgtt-gcTCTTCTTCTTCTTCTTCTTCTTCTggtgatgatgatgta-caagttctgttgcttattgatgaagtccaacaaccaatactctttcctcgctttcatca	accaccaacaactac-gagga	gcgaggaaagag-tattggttg	151
15	CDS_62182	(GAA)8	903	309	332	ctttgggatgaaagggatattaacaagctttgtca-caccagctgcaccttccaaagaaaatcacaattct-gaaggGAAGAAGAAGAAGAAGAAGAAGAAcaggctgtatgt-gtgcaaccacagaacacttttgatccaatttgtatttccgtggcactcag-gaaacctgtaat	tgggat-gaaagggatat-taacaa	tgagtgccacg-gaaatacaa	158
16	CDS_2169	(CGG)8	1008	591	614	ccgtagaaacaagaaaagcaaaagcggtagcagctcgaaatcttcagc-tagctctgatcatcaccggcagattggTAATAATAATAATAActcaac-cagtacggctagtccttccagctgcaccacggatatggtcggtcatcatttccc-acagccgccgtcaac	ccgtagaaa-caagaaaagcaaaa	ggctgtgggaaat-gatgac	156
17	CDS_14006	(AAT)8	1287	733	756	gactgtgttgtcgtcgagagctgctgctatggcggtaatagtcttggaacct-gttcttcagcaacaagctatggaAATAATAATAATAATAATAATAAT-gctttacttatcagtcgctcaatgagctatccctctgagaattttctcagtttctgcta-caaatgcaagaagaat	tgtgttgtcgtcgaga-gctg	tgcatttgtag-cagaaactgaga	163

18	CDS_23082	(TGC)7	1155	117	137	taataataatcacgagatgaagagaatcccttcggagttggcaatggac-gagctgttcaagcacacgagagccgaTGCTGCTGCTGCTGCT-GCTGCtcagatcggcccgaataatgatcagaaagacgcggctaaatc-gagaacaactgatcatcagacctttgggtgcag	aagagaatcccttcg-gagttg	gcacccaaaggtct-gatgat	150
19	CDS_53326	(CTC)7	999	134	154	ccccgattctctctcttcccaccaactcctcttcttcttccatttctctcaagcata-aattctcattctcattcaCTCCTCCTCCTCCTCCTCCT-Catatcttcttctccacttcttcggcgccaccgcgacagctacaacaactac-cgtcacctctgcctatatctgcct	tctctcttcccac-caactcct	aggcagaggtgacg-gtagtt	152
20	CDS_62796	(ACC)7	648	145	165	gagaaagaagaaattaaagggatgatgatgtcatctgaggctacctcat-catcccagggattggtaatcagtagtACCACCACCACCACCAC-CACCactgataaggaagagcatgaaaagggaaagaactatatcggtgttc-gtaagaggccatggggaaaatttgctgct	gaagaaatta-aagggatgatgatg	ccatggcctcttac-gaacac	150
21	CDS_37800	(ATA)7	858	200	220	ccctggctcagcgtttccaggttctggttttcgactggagcttttcgggtgcggta-attaataatgataatgatgATAATAATAATAATAATAATAaggactc-gacgaagaaccagctgttcgatgcagccaaatacacttgctatgatgcttttg-cagatgacttgatag	tccaggttctggttttc-gac	tcaagtcatctg-caaaagca	153
22	CDS_9020	(CCA)7	675	221	241	ccccaactgctctactcaccccaaccaccgccactcccaagccactgactc-caattaagtcgaaactccttcctcCCACCACCACCACCACCAC-CAatgcagccacaggtccttcaccttgtactcgccgccgtgacttgttatctctgg-ctgtcggaatagtcgtcgcac	ctgctctact-caccccaacc	tccgacagccaga-gataaca	152
23	CDS_16387	(CGA)7	600	225	245	tctaaacaccccaaaaggcttcggaacctcaccaccaaagaaatc-gaagaaaccaaaaaagggctacaacaaattCGACGACGACGAC-GACGACGACggcgaagatgaagaagaagaagagcgagaagag-gacgttataccagagatcgtgacgaatagaatgatgaacag	taaa-caccccaaaaggcttc	cgtcacgatctctg-gtataacg	152
24	CDS_61610	(CCT)7	384	226	246	ggggaaaacatgggtagttctggaagctacaaggcggc-cagcttgggggatatttcaggtataaaggagtttgggCCTCCTCCTCCT CCTCCTCCTcttactataatctgcaagcaacagcccccaccactcctctg-tatttcctttcctcctctatcagagtcgctcgag	ggggaaaacatggg-tagttc	tcgagcgactct-gatagagga	170
25	CDS_62309	(CAG)7	2064	247	267	ccactgagtcctcctacgacttcagcttcatcacagactgttgctccc-gtttcttctccgatgtctactcgactaCAGCAGCAGCAGCAGCAG-CAGgattattttacttcggaagaggagtatcaggtgcaattagcccttgcctt-gagtgcgtcggattcatcaggccac	cctcctac-gacttcagcttca	ctgatgaatccgac-gcact	157
26	CDS_24775	(CTG)7	726	266	286	cccccgtactatcaagggatttctttgctgtagatcctccacaatctgag-gcccaaatcaacaacaaccaactgcCTGCTGCTGCTGCTGCT-GCTGcttcagcaatggatgaagaagttgaatcagtcgactctccacactc-catcgaatccgaagtgttgccacccaaca	ccccgtactat-caagggattt	cttcggattcgatg-gagtgt	155
27	CDS_46528	(GCA)7	537	278	298	aacaaactcctctcactcaatgtgagtggtttcgagatttcaatccggt-gttcagtcgtcatcaatcacagagtgGCAGCAGCAGCAGCAG-CAGCAacaacaatggttctccagctcaacgagctttccactacatggctgt-gaagttcaaaaggctgttcagagagctgg	tcaatgtgagtggtttc-gagat	ccagctctctgaa-cagccttt	155
28	CDS_14113	(TAA)7	1239	342	362	tgcgacggtgcctttgactgggaagagaagcctggtaagcccaaaat-gaaatcacccgccgtcataggtggttcTAATAATAATAATAATAATA-Atgaaggagaaggtggaggaggatatgactttgcttttgaagtgagtgag-gatttcgacagagtctctgtttctgc	gtgcctttt-gactgggaaga	ctctgtcgaaatcct-cactca	151
29	CDS_13865	(TTA)7	519	364	384	gatatggaatctaaattatctaacagtaatactgactccaccgagaataatg-gaagagttatttgcaaggtccgtTTATTATTATTATTATTATTAtttata-attttaaaaaatatttatatcatgctttcttataagctttactcttcactttattgcttgtt-tactgtt	tgactccaccgaga-ataatgg	caagcaataaagt-gaagagtaaagc	130
30	CDS_12534	(TGA)7	1968	462	482	tgaggatattgttgacacggagacagaggaatcatgtcacaatactgac-gacgataaatatgaagatagctttatTGATGATGATGATGATGAT-GAactggaagtttttcacattcacctgtttcgagtgatcaaggtaaaacaaa-gatgatagagcagaaaggcaacag	tgttgacacggaga-cagagg	tgttgcctttctgctc-tatca	161
31	CDS_24774	(CTG)7	1068	608	628	cccccgtactatcaagggatttctttgctgtagatcctccacaatctgag-gcccaaatcaacaacaaccaactgcCTGCTGCTGCTGCTGCT-GCTGcttcagcaatggatgaagaagttgaatcagtcgactctccacactc-catcgaatccgaagtgttgccacccaaca	ccccgtactat-caagggattt	cttcggattcgatg-gagtgt	155
32	CDS_35537	(GAC)7	792	658	678	gagaagttgaaaagggaggcaaaggggattccattgaaacagtatgat-gaagaagatgaaagtcaagatgaagatGACGACGACGACGACGAC-GACGACtttgattatagcatcctggctgatccaaatgtgaatgtatatgagc-cagtgcaacctcatgttaatggcactgaa	gttgaaaagggag-gcaaagg	tgaggttgcactg-gctcata	150
33	CDS_8492	(TCA)7	2241	706	726	gagttgactcaattgtcgcacttactcactctccgtctcgagtttaactcgtt-taatggaactctttcctctgtcTCATCATCATCATCATCATCAtc-tactctggcctcctcactttccgacttcaatgtctcagaaaatggcctcgccgg-taggatcccagactggtta	ttgtcgcacttact-cactctcc	ac-cagtctgggatcctac-cg	157
34	CDS_33255	(GAG)7	1317	1115	1135	aagaagacgttctgagacttcagagccagtgcatggtaatgcaaggaca-gattgagaggttaatggagaaaaagcGAGGAGGAGGAGGAG-GAGGAGggcttttcagttggaagaagatcggaatgctgtctcttagagctgc-taataacagtaagtttggggaagttgacg	tcagagccagtgcatg-gtaa	cgt-caacttccccaaactt	152
35	CDS_9535	(GAT)7	1581	1429	1449	ttggtggagagatcatgtacccaattgattaaaatacctggaaactctccaat-gagtataattgctggagaaaacGATGATGATGATGATGATGATa-cagagagtgttaggttggatcacatgtttccaggtcaagagtttaggtcctttctg-gctgtagagagtctgaat	tggtggagagatcat-gtaccc	cagc-cagaaaggaccta-aactc	156

36	CDS_10965	(CAC)7	1971	1510	1530	agaaatccaccacccttgggcggtggtggcggcattggcagccatttatttg-gtcttagcggcggaagcagcagtAACAACAACAACAACAACatgat-gagcttaggtttatctcaattgggttcttctcaaattgaccaaaaccagag-tagtagtactactccagca	agaaatccac-caccccttgg	ctactactctggttttg-gtcaattt	152
37	CDS_15968	(AG)12	303	139	162	ttaggacaggaaacgctggaccctccgagaaaacctcactcaaagtcagtt-gttttcaactcccgaggacagcacAGAGAGAGAGAGAGAGA-GAGAGAGaaggggagtctctttctcagcagtgaatggtggtggtgttg-gtgggcagagaaagctagattacgccagtggaca	taggacaggaaac-gctggac	ctagctttctct-gcccacca	157
38	CDS_24144	(AT)12	1758	245	268	tttcattcaattatctatacgccccttcttattccg-caataacgttacattcaaaaactctgttctacatgta-cATATATATATATATATATATATATTtcctgttcttgtttcagttcatta-cagtgtaaggcttcaagatggttcttactttgaatccagcagcaagaaag	atacgccccttct-tattcc	ttcttgctgctg-gattcaaa	156
39	CDS_52371	(CT)10	546	121	140	atatacttcttcgattgcattcgtctaccattgatctcaaggttttgctttcatttttgta-atcctagagtcaagCTCTCTCTCTCTCTCTCTCTgatcacac-taatgaagaaggtggttgtggaagtaggcgtccgtgatgagaaagataaa-cagaaggcgatgaaagc	tcgattgcattcgtc-tacca	acgcctacttcca-caaccac	124
40	CDS_4021	(AC)9	888	217	234	cttcctaaccacaaagggctgtcacttttaactcaagaaatgacctatatagact-tatttttacccaaaaaaaaaACACACACACACACACACaaattatta-gattgttttatccttaaattggtagaaaattcaccctccaaaaaaggacaaaat-gactaccagaca	aaccacaaagggct-gtcact	ggagggtgaattttc-taccaa	135

Table 5: Frequency of classified repeat types (considering sequence complementary) in the analysed 4448 splicing sequences.

Repeat motif	Number of repeat units												
	5	6	7	8	9	10	11	12	13	14	15	>15	**Total**
AC/GT	-	65	37	25	16	7	18	3	3	3	-	1	178
AG/CT	-	215	125	68	74	58	44	46	16	6	5	39	696
AT/AT	-	247	92	94	76	71	54	11	11	1	3	0	660
AAC/GTT	117	42	9	11	5	3	1	-	-	-	-	0	188
AAG/CTT	470	260	124	89	5	16	6		1	1	-	6	978
AAT/ATT	122	70	21	11	-	-	-	-	-	1	-	0	225
ACC/GGT	238	103	53	19	2	-	5	-	-	-	-	0	420
ACG/CGT	63	14	9	4	-	2	-	-	-	-	-	0	92
ACT/AGT	73	27	19	2	-	-	-	-	-	-	-	0	121
AGC/CTG	352	71	38	25	1	-	-	-	-	-	-	0	487
AGG/CCT	162	91	67	34	14	12	6	1	-	-	-	0	387
ATC/ATG	403	159	59	24	3	3	-	-	-	3	-	0	654
CCG/CGG	59	31	4	5	-	-	-	-	-		-	0	99
AAAC/GTTT	3		-	-	-	-	-	-	-	-	-	0	3
AAAG/CTTT	1	1	-	-	-	-	-	-	-	-	-	0	2
AAAT/ATTT	25		-	-	-	-	-	-	-	-	-	0	25
AAGG/CCTT	-	2	-	-	-	-	-	-	-	-	-	0	2
ACAG/CTGT	1		-	-	-	-	-	-	-	-	-	0	1
ACAT/ATGT	10	3	-	-	-	-	-	-	-	-	-	0	13
ACTG/AGTC	1	-	-	-	-	-	-	-	-	-	-	0	1
AGAT/ATCT	1	-	-	-	-	-	-	-	-	-	-	0	1
ATCC/ATGG	1	-	-	-	-	-	-	-	-	-	-	0	1
ATGC/ATGC	3	-	-	-	-	-	-	-	-	-	-	0	3
AAAAC/GTTTT	1	1	-	-	-	-	-	-	-	-	-	0	2
AAAAG/CTTTT	1	-	-	-	-	-	-	-	-	-	-	0	1
AAATC/ATTTG	3	-	-	-	-	-	-	-	-	-	-	0	3
AAGAG/CTCTT	4	-	-	-	-	-	-	-	-	-	-	0	4
AGAGG/CCTCT	1	-	-	-	-	-	-	-	-	-	-	0	1
AAAGGG/CCCTTT	1	-	-	-	-	-	-	-	-	-	-	0	1
AACGGG/CCCGTT	-	3	-	-	-	-	-	-	-	-	-	0	3
AAGGAG/CCTTCT	3	-	-	-	-	-	-	-	-	-	-	0	3
AATCCC/ATTGGG	-	1	-	-	-	-	-	-	-	-	-	0	1
AATTCC/AATTGG	3	-	-	-	-	-	-	-	-	-	-	0	3
ACAGGC/CCTGTG	-	1	-	-	-	-	-	-	-	-	-	0	1
ACCGCC/CGGTGG	1	3	-	-	-	-	-	-	-	-	-	0	4
ACCTCC/AGGTGG	1	2	-	-	-	-	-	-	-	-	-	0	3
ACGGAG/CCGTCT	-	1	-	-	-	-	-	-	-	-	-	0	1
AGATGG/ATCTCC	-	4	-	-	-	-	-	-	-	-	-	0	4
AGCAGG/CCTGCT	-	2	-	-	-	-	-	-	-	-	-	0	2
AGCCTC/AGGCTG	-	1	-	-	-	-	-	-	-	-	-	0	1
ATCGCC/ATGGCG	1	-	-	-	-	-	-	-	-	-	-	0	1

at genetic level is concerned, no data was available revealing polymorphism through SSR markers either genomic or EST-SSRs up till now. This was the first study revealing genetic diversity and high polymorphisms in *G. sylvestre* with the help of EST-SSR markers.

Furthermore, a dendrogram encompassing all 25 genotypes of *G.Sylvestre* collected from different locations across India used in this study was also built based on EST-SSR markers. The results showed all the varieties of *G. sylvestre* were clustered in two main groups. Genotypes DGS 16 and DGS 34 were the most dissimilar genotypes in which DGS 16 was collected from Tamilnadu and DGS 34 was collected from Jharkhand. DGS 15 and DGS 8 both were collected from the same state (Karnataka) but were falling under different clusters. DGS 1, 2, 3 were collected from Gujarat in which DGS 1 and 3 both are having similarity and falling under the same cluster but DGS 2 is falling under different cluster. Whereas Genotypes collected from Andhra Pradesh (DGS 19, DGS 20, DGS 22 and DGS 23) and Tamilnadu (DGS 16, DGS 17, and DGS 18) were covered under cluster B.Thus, disorderly distribution of genotypes of same state in different clusters was revealed from this study. Central location of the genotype collected from Rajasthan in the minimum span tree of the PCA may be because that the *G. sylvestre* may not be monophyletic. At the same time, it is pertinent to mention here that *G. sylvestre* is an important medicinal plant people have been using from ancient time and therefore, it is also presumed that it would have been migrated to different places which may be the reason behind messy distribution of genotypes in this study.

Conclusion

The information on genetic variation within populations of plants based on EST-SSR was lacking in *G. sylvestre.* This was the first attempts at *G. sylvestre* genetic and genotypic diversity analysis using EST-SSR markers developed from transcriptome sequencing. Through transcriptome analysis, we have generated huge data (9.14 Gb) and predicted 71,676 CDSs with an N50 length of 896 bp and identified total of 5276 SSRs loci. The AAG/CTT repeats were nearly ten times higher than the CCG/CGG repeats. We have designed 40 pairs of primers and reported 27 primers to be polymorphic in 25 genotypes of *G. sylvestre* collected from different parts of India. This was the first study revealing genetic diversity and high polymorphisms in *G. sylvestre* with the help of EST-SSR markers having a higher transfer rate of 67.5%. A dendrogram showed that 25 genotypes of *G. sylvestre* were clustered in two main groups. Genotypes DGS 16 and DGS 34 were the most dissimilar genotypes.

Availability of Data and Materials

The data generated or analysed during this study are included in this published article, its supplementary information files, and publicly available repositories. The transcriptome raw data are deposited at NCBI under Project SUB2977090 with SRR 5965323.

Competing Financial Interests

The authors declare no competing financial interests.

Acknowledgments

We acknowledge the funding through FAP Scheme from Gujarat State Biotechnology Mission (GSBTM 4867, 2016-17,23.09.2016) Govt. of Gujarat, and the ICAR-DMAPR, Anand and the ICAR, New Delhi for providing the basic facilities for this research work, the germplasm explorer and all the curators of the genotypes used in this study.

References

1. Stocklin W (1969) Chemistry and physiological properties of gymnemic acid: the antisaccharine principle of the leaves of *Gymnema sylvestre*. J Agric Food Chem 17: 704-708.
2. Kalariya KA, Minipara DB, Manivel P (2018) De novo transcriptome analysis deciphered polyoxypregnane glycoside biosynthesis pathway in Gymnemasylvestre. Biotech 8:381.
3. Kalariya KA, Minipara D, Saran PL, Poojara, L, Polireddy AC, et al. (2019) Photosynthetic efficiency in Gymnemasylvestre (Retz.) R. Br. genotypes. Vegetos 1-13.
4. Kalariya KA, Gajbhiye N, Minipara D, Meena RP, Kumar S, et al. (2019) Deep sequencing-based de novo transcriptome analysis reveals biosynthesis of gymnemic acid in Gymnema sylvestre (Retz.) Schult. Eco Gen Genom 13:100047
5. Kalariya KA, Meena RP, Lipi P, Deepa S, Sandip P (2021) Characterization of squalene synthase gene from Gymnemasylvestre R. B 10:6.
6. Kalariya KA, Meena RP, Saran PL,Manivel P (2019) Identification of microRNAs from transcriptome data in gurmar (Gymnema sylvestre). Hort Env Biotech 60:383-397
7. Nair S, Keshavachandran R (2007) Molecular diversity in chakkarakolli (*Gymnemasylvestre* R. Br.) assessed through isozyme and RAPD analysis. J Trop Agric 44:31-7.
8. Shahnawaz M, Zanan RL, Wakte KV, Mathure SV, Kad TD, et al. (2012) Genetic diversity assessment of *Gymnema sylvestre* (Retz.) R. Br. ex Sm. populations from Western Ghats of Maharashtra, India. Genet Resour Crop Evol 59:125-34.
9. Osman MA, Dhawan SS, Bahl JR, Darokar MP (2013) Genetic diversity analysis in *Gymnema sylvestre* R. Br. by RAPD. Int J Int Sci Inn Tech 2:50-54.
10. Khatak S, Chauha N, Malik DK (2014) DNA isolation and optimization of PCR conditions in *Gymnema sylvestre* by using inter simple sequence repeat (ISSR) markers. Int J Chem Sci 12:385-92.
11. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatic 30: 2114-2120.
12. Haas M, Sis B, RacusenLC, Solez K, Glotz D, et al. (2014) Banff 2013 meeting report: inclusion of c4d-negative antibody-mediated rejection and antibody-associated arterial lesions. Am J Transplant 14:272-283.
13. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST a new generation of protein database search programs. Nucleic Acids Res 25:3389-402.
14. 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.
15. Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Bioinform. Methods Protoc 132:365-386.
16. Shiva V (1994) Agriculture and food production. UNESCO/ Environmental Education Dossiers No. 9May2-3.
17. Rao VR, Hodgkin T (2002) Genetic diversity and conservation and utilization of plant genetic resources. Plant cell Tiss Organ 68:1-9.
18. Ellstrand NC and Elam DR (1993) Population genetic consequences of small population size: implications for plant conservation. Ann Rev Ecol Syst 4:217-242.
19. Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. Mol Ecol 7:465-474.
20. Gorman GC, Renzi Jr J (1979) Genetic distance and heterozygosity estimates in electrophoretic studies: effects of sample size. Copeia 18:242-9.
21. Luo R, Wu W, Zhang Y, Li Y (2010) SSR marker and its application to crop genetics and breeding. Genomics Appl Biol 29:137-143.
22. Feng B, Soojin VY, Zhang M, Zhou X (2018) Development of novel EST-SSR

markers for ploidy identification based on de novo transcriptome assembly for Misgurnusanguillicaudatus. PloS one 13:e0195829.

23. Shi X, Wu H, Li W, Guo W, Zheng Y, et al. (2018) Development of polymorphic EST-SSR markers in Iteachinensis (Iteaceae) and cross-amplification in related species. Appl Plant Sci 6:e1013.
24. Wang J, Ge Q, Li J, Gao H, Li J, et al. (2018) Identification of novel E st-ssr markers by transcriptome sequencing in ridgetail white prawn Exopalaemoncarinicauda. Genes Genom 40:207-215.
25. Tang DQ, Lu JJ, Fang W, Zhang S, Zhou MB (2010) Development characterization and utilization of GenBank microsatellite markers in Phyllostachys pubescens and related species. Mol Breed 25:299-311.
26. Cardle L, Ramsay L, Milbourne D, Macaulay M, Marshall D, et al. (2000) Computational and experimental characterization of physically clustered simple sequence repeats in plants. Genetics 156:847-854.
27. Peng JH, Lapitan NL (2005) Characterization of EST-derived microsatellites in the wheat genome and development of eSSR markers. Funct Integr Genomic 5:80-96.
28. Cai K, Zhu L, Zhang K, Li L, Zhao Z, et al. (2019) Development and characterization of EST-SSR markers from RNA-Seq data in Phyllostachys violascens. Front Plant Sci 10:50.
29. Zhou Q, Luo D, Ma L, Xie W, Wang Y, et al. (2016) Development and cross-species transferability of EST-SSR markers in Siberian wildrye (Elymus sibiricus L.) using Illumina sequencing. Sci Rep 6:20549.
30. Hu Z, Zhang T, Gao XX, Wang Y, Zhang Q, et al. (2016) De novo assembly and characterization of the leaf, bud, and fruit transcriptome from the vulnerable tree Juglans mandshurica for the development of 20 new microsatellite markers using Illumina sequencing. Mol Genet Genom 291:849-862.
31. Chen LY, Cao YN, Yuan N, Nakamura K, Wang GM, et al. (2015) Characterization of transcriptome and development of novel EST-SSR makers based on next-generation sequencing technology in Neolitsea sericea (Lauraceae) endemic to East Asian land-bridge islands. Mol Breed 35:187.
32. You Y, Liu D, Liu H, Zheng X, Diao Y, et al. (2015) Development and characterisation of EST-SSR markers by transcriptome sequencing in taro (Colocasia esculenta (L.) Schoot). Mol Breed 35:134.
33. Varshney RK, Thiel T, Stein N, Langridge P, Graner A (2002) In silico analysis on frequency and distribution of microsatellites in ESTs of some cereal species. Cell Mol Biol Lett 7:537-46.
34. Bosamia TC, Mishra GP, Thankappan R, Dobaria JR (2015) Novel and stress relevant EST derived SSR markers developed and validated in peanut. PLoS One 10:e0129127.

Author Affiliations

Top

ICAR-Directorate of Medicinal and Aromatic Plants Research, Boriavi, Anand-387310