



Molecular marker based genetic diversity study of wild, cultivated and endangered species of *Curcuma* from Chhattisgarh region for *in situ* conservation

Akanksha Jain, Dharmendra Kumar Parihar*

Department of Biotechnology, Guru Ghasidas Vishwavidyalaya, Bilaspur, C.G., 495009, India



ARTICLE INFO

Keywords:

EST-SSR
Genetic diversity
Curcuma aromatica
Curcuma caesia
Curcuma longa

ABSTRACT

Curcuma is a well-known genus for its extensive medicinal properties derived from several chemical constituent. Curcumin is known as a wonder drug because of its pharmaceutical importance. Curcumin content could vary based on different species and its genetic makeup and environmental factors. So in the present study, was based on the investigation of gene expression of different species of *Curcuma* from Chhattisgarh region based on EST based molecular fingerprinting. Genetic fingerprints of *Curcuma caesia*, *Curcuma longa* and *Curcuma aromatica* were determined using 17 EST-SSR (Expressed sequence tag Simple Sequence Repeats) to elucidate the genetic diversity for their utilization, genotypic conservation and large scale production of curcumin derived from these species. The primer combinations were amplified 127 loci among which 102 were found to be polymorphic in nature and the rest were monomorphic, the polymorphism percent was 81%. These markers were used to estimate genetic similarity and distance between the *Curcuma* species. Various data scoring methods were used to analyze the similarities and relationships between species of *Curcuma*. Data analysis revealed distinct genetic identity found between species of *Curcuma* at the genomic level. The present study provides a baseline data for optimization of conservation and breeding program of this medicinally important rhizome on the basis of their documented genetic diversity which would be an important step towards optimization of industrial level production of curcumin.

1. Introduction

Curcuma is a highly important genus of Zingiberaceae family. Most of the members of this genus belong to spice plants with very huge medicinal value, among them few species have been cultivated, but majority of them are wild, endemic or endangered in nature. Therefore, these species are gradually depleting from nature due to extensive collection, biopiracy and habitat destruction. *Curcuma longa* is a well explored species, however the other two non-conventional species such as *Curcuma caesia* and *Curcuma aromatica* are not explored. The tribal people have been using these species medicine as well as spice. Traditional methods based on phenotypic observations for identifying different genotypic varieties of *Curcuma* species are slow and limited. Thus, new methods based on studies of DNA fingerprinting to assess genetic relationships or diversity among genotypes can accelerate plant breeding programs. Molecular markers studies have been highly useful for breeding and cultivar development in many crops (Syamkumar and Sasikumar, 2007). Inter simple sequence repeats (ISSR) techniques have

proven to be a reliable, reproducible, easy to generate, inexpensive and versatile set of markers that relies on repeatable amplification of DNA sequences using single primers (Mohanta et al., 2015). Microsatellites or simple sequence repeats (SSR), which are tandem repeats of 1–6 nucleotide long DNA motifs, had gained considerable importance in plant genetics and breeding because of their multi-allelic nature, codominant inheritance, high abundance, extensive genome coverage, reproducibility, and discriminatory power (Kalia et al., 2011). Several researchers have been used different molecular marker techniques such as random amplified polymorphic DNA (RAPD) (Nowbuth et al., 2005; Lapitan et al., 2007; Barakat et al., 2010) and SSR (Tian et al., 2008) have been for measuring genetic variability as well as genetic similarity in plants. In addition, several molecular studies have employed various DNA markers in *Curcuma* (Apavatjirut et al., 1999). Researchers used isozyme markers to resolve the taxonomic confusion in the genus *Curcuma* (Taheri et al., 2012). The knowledge of genetic variation is a prerequisite to study the evolutionary history of a species, as well as other intraspecific variation, genetic resource conservation etc. (Islam

* Corresponding author.

E-mail addresses: akankshanakhat@gmail.com (A. Jain), parihardkp@rediffmail.com (D.K. Parihar).

Abbreviations

EDTA	Ethylene diamine tetraacetic acid
EtBr	Ethidium bromide
PCA	Principle coordinate analysis
PIC	Polymorphic information content

PCR	Polymerase chain reaction
EST	SSR- Expressed sequence tag Simple sequence repeats
TE	Tris-EDTA buffer
TAE	Tris acetate EDTA
UPGMA	Unweighted pair group method with arithmetic mean

et al., 2007). Therefore, study of genetic diversity and gene differentiation through molecular marker analysis are essential for evaluation their taxonomic relationship, conservation and sustainable utilization. It is essential to characterize the plants at the genetic level for proper conservation of germplasm. A number of molecular markers have been regularly used for studying genetic relations, population genetics, genetic characterizations in different plant groups and cultivars. The molecular markers are not influenced by the external environmental factor unlike that of morphological markers and hence accurately testify the genetic relationship among plant groups. Molecular markers like RAPD, ISSR and SSR are being used for genetic diversity assessment as through knowledge of the level and distribution of genetic variation is essential for conservation (Dreisigacker et al., 2005; Sharma et al., 2008; Naik et al., 2010; Das et al., 2011). RAPD, ISSR and SSR have proven to be very informative, and cost-effective. Interestingly PCR-based DNA fingerprinting techniques in many plant species do not require prior knowledge of a species genetics (Williams et al., 1990; Zeitzkiewicz et al., 1994; Lee et al., 2007). Many workers had reported the genetic diversity among *Curcuma* species (Das et al., 2011; Jatoi et al., 2006; Syamkumar and Sasikumar 2007) but the studied species are area specific based on their availability in that region. The scarcity of knowledge genetic relationship among cultivated and wild species is the main reason for extinction these species. Earlier studies have been attempted based on morphological, biochemical and anatomical characterization in *Curcuma* species (Jiang et al., 2006; Zhou et al., 2007; Paramasivam et al., 2009), but these studies have their own limitation as unable to define the genetic structure of these studies (Noli et al., 1997). Molecular profiling of non-conventional *Curcuma* species are still at an emerging stage. Reports were restricted to specific species and its genotypes. Studies have done on the genetic diversity analysis using RAPD, SSR, and ISSR markers on the members of genus *Curcuma*,

but no effort has been made to analyze the endemic *Curcuma* species of Chhattisgarh India. The main goal of the present study was to access the degree of genetic diversity and to analyze the genetic proximity among the selected species of *Curcuma* by using EST- SSR marker system. ESTs are short and single pass sequence reading from mRNA (cDNA) representing a snapshot of genes expressed in a given tissue and at a given developmental stage (Adamset al., 1991). EST databases have been proved to be a valuable source of polymorphic SSRs (EST-SSRs or genic SSRs) in a number of plant species. EST- SSR markers having following advantages over other markers presence in the gene rich regions of the genome, such as identification by electronic sorting, relative abundance, and easy transferability to related species (Varshney et al., 2005). The present study was based on scanning ESTs to determine the hypervariable repeats and generating a robust set of polymorphic markers for *Curcuma* species. This might lead to genetic improvement, selection of high yielding germplasm and evaluation of accessions from different geographical regions of Chhattisgarh that may increase the efficiency of selection in breeding programs.

2. Materials and methods**2.1. Plant material and DNA extraction**

The present investigation deals with three species of *Curcuma* (*C. longa* L., *C. caesia* Roxb. and *C. aromatica* salibs., collected from village Dongargaon of Rajnandgaon district, Chhattisgarh, India. The protocol for DNA extraction was performed according to Edwards et al. (1991) and Das et al. (2011) with slight modification. Genomic DNA was isolated from liquid nitrogen crushed leaf samples by grinding with a mortar, pestle in extraction buffer (100 mM Tris-HCl [pH 8.0], 1.4 M NaCl, 20 mM EDTA, 2% SDS) and incubated at 65 °C for 1 h in 500 µl of

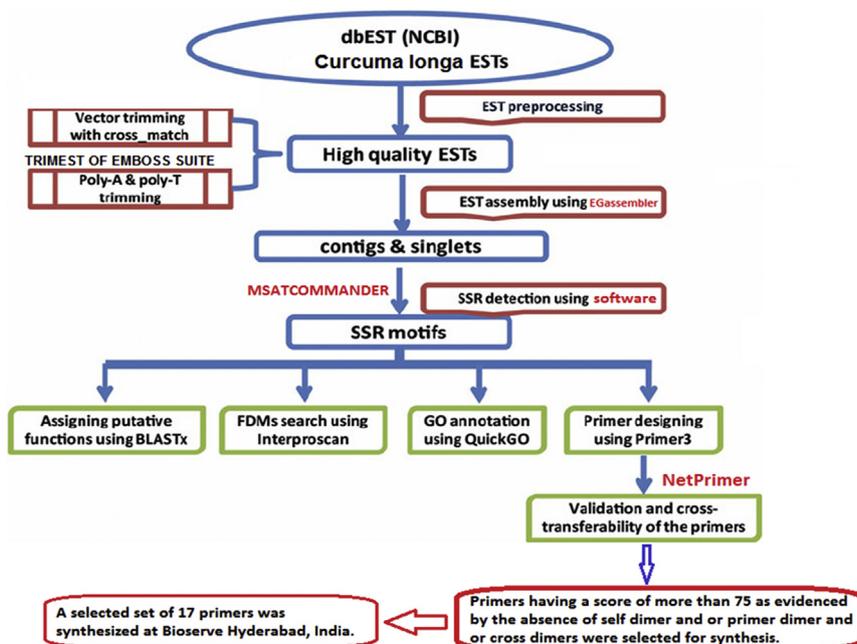


Fig. 1. Graphical representation of In silico survey and designing of EST SSR primers for gene expression study of three targeted *Curcuma* species.

Table 1
List of primers designed from EST and CDS regions of gene for expression studies.

LOCUS	EST/CDS sequence ID	GENE blast X hits
CLEST SSR-01	DY394887	Predicted protein [Populustrichocarpa]
CLEST SSR-02	DY389303	UDP-glucose 6-dehydrogenase [Zea mays]
CLEST SSR-03	DY394652	Alpha-1,4-glucan-protein synthase [Ricincommunis]
CLEST SSR-04	DY393469	Hypothetical protein [Zea mays]
CLEST SSR-05	DY394828	Hypothetical protein [Zea mays]
CLEST SSR-06	DY393861	S-adenosyl-L-methionine synthetase
CLEST SSR-07	DY393567	Hypothetical protein [Oryza sativa Japonica]
CLEST SSR-08	DY394591	Curcuminoid synthase (CURS 3)
CLEST SSR-09	DY394891	Hypothetical protein [Oryza sativa] Japonica
CLEST SSR-10	DY393238	Photosystem-1 F subunit precursor [Oryza sativa Japonica group]
CLEST SSR-11	DY390357	Germacone synthase
CLEST SSR-12	DY391910	Predicted protein [Populustrichocarpa]
CLEST SSR-13	DY393462	Hypothetical protein [Oryza sativa Indica]
CLEST SSR-14	DY391880	Hypothetical protein [Vitisvinifera]
CLEST SSR-15	HM161811	Calcone synthase (CHS1) for flavonoid synthesis
CLEST SSR-16	DY384950	Curcumin synthase (CURS)
CLEST SSR-17	DY388605	AP2/ERF domain-containing transcription factor

Table 2
Sequences and nucleotide length of primers used in the EST- SSR analysis.

Locus	Forward Primer Sequences	Reverse Primer Sequences	polymorphism
CLEST SSR-01	TTTGAGATGGCGAGTAGAAC	ATGAGGGAAGAGAGGAGAAG	100%
CLEST SSR-02	ACCGTAGCAAAGAAATAGGAC	AAGGTGGAAGGAACTCG	71.42%
CLEST SSR-03	AGGGAAAATAGAGTAGGCAAG	TGAAGGATTACAGTCAGCAA	100%
CLEST SSR-04	ACACAACATTCAGTTTAGCAC	TCCCTATTCCTTCCTCTCG	66.67%
CLEST SSR-05	TATCCTCCCTGGTCGTTT	GATTCCTTTCTTTCTTTTG	88.89%
CLEST SSR-06	TCATCGTCTGCTTTAGTTTTC	AGGCTCTGCTCCTTCAAC	83.34%
CLEST SSR-07	AGACAGAAGAAGAGGCAGAAG	AAATGATGACCACGGACTAC	87.50%
CLEST SSR-08	CTGTGAGAAGACGAAGGTGA	CITTTGATATCTCCTCCACCA	75%
CLEST SSR-09	TCGGTTCTACTGAATCTTTACTCG	AGACTGTTTTCCCATTTGTTGC	75%
CLEST SSR-10	GTGGTGGAGGAGGAAGAGAAG	TTGAGGGAACAAAAGGAAGAC	75%
CLEST SSR-11	TTCAATCGACGCAAACAGC	CGACGAATAGTCGAAGGC	100%
CLEST SSR-12	GGGATTGAGGTGAGGTAGG	GCTGGCGAAGTAGAAGAAGAA	60%
CLEST SSR-13	TGTACAAGCTCCAATAAGTCAAG	CAGGAGTGTCTAATGTTGCC	77.78%
CLEST SSR-14	CACCTCTCCTCCCAACC	GCCGTCCTCGCTTCTTCTTA	75%
CLEST SSR-15	AGGTCAACACGCTCATCTTC	CGTCTCCAAGTTGTTGGCTA	83.34%
CLEST SSR-16	ATAACACCCTCCTCCTCTC	AAGTGCTCTGCTCGTAAAGG	71.43%
CLEST SSR-17	GTGCCTGTGGACCTATCCG	GAAGCATGCGAATTCATCTAAAC	81.81%

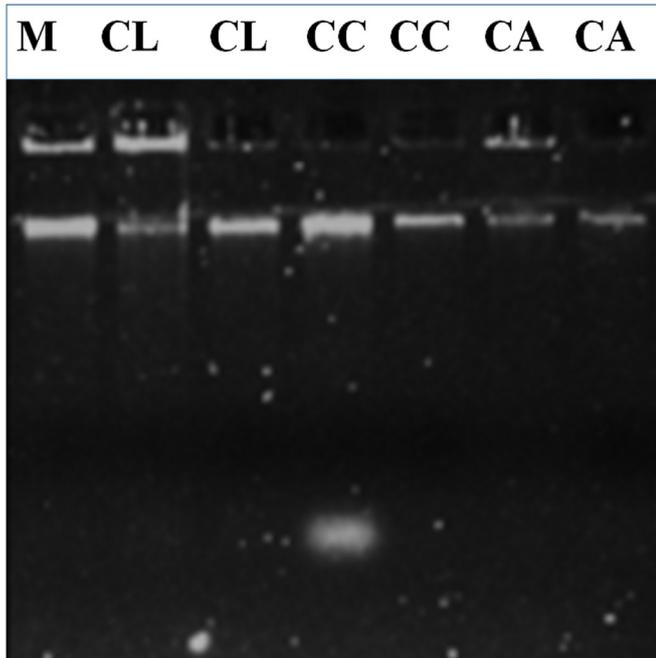


Fig. 2. Gel image of genomic DNA of Three targeted Curcuma species on 1% agarose and λ DNA (lane M).

SDS extraction buffer. It was further extracted with an equal volume of chloroform-isoamylalcohol (24:1) and centrifuged at 5000 rpm for 15 min, the upper phase containing DNA was treated with 5 μ l (10 mg/ml) RNase for samples to eliminate RNA by incubation at 37 °C for 15 min and again treated with Chloroform:Isoamylalcohol for removal of non-nucleic acid compounds. Genomic DNA was precipitated out from the upper phase by adding 0.6 vol of chilled ethanol and 100 μ l of 5M NaCl and keeping at less than 4 °C for several hours. The solution was centrifuged at incubation centrifuge at 10000 rpm for 15 min after incubation and the precipitate was washed several times with 75% ethanol, and then air-dried and dissolved in appropriate volume of TE buffer (10 mM Tris-HCl [pH 8.0], 1 mM EDTA). The quality and quantity of the extracted DNA were determined with a Thermo Scientific Nano drop and genomic DNA was also quantified by agarose gel electrophoresis with use of standard lambda DNA as a reference.

2.2. Designing of primers

A total of 12,593 *C. longa* L. EST sequences were downloaded from the dbEST database hosted in GenBank (National Centre for Biotechnology Information, <http://www.ncbi.nlm.nih.gov/dbEST/>) using the keyword “*Curcuma longa*”. In a preliminary step, poly A and poly T-stretches of the ESTs corresponding to the poly A tails of eukaryotic mRNA were removed using TRIMEST program of EMBOSS suite according to parameters described by Kumpatla and Mukhopadhyay (Kumpatla et al., 2005). The sequences were assembled

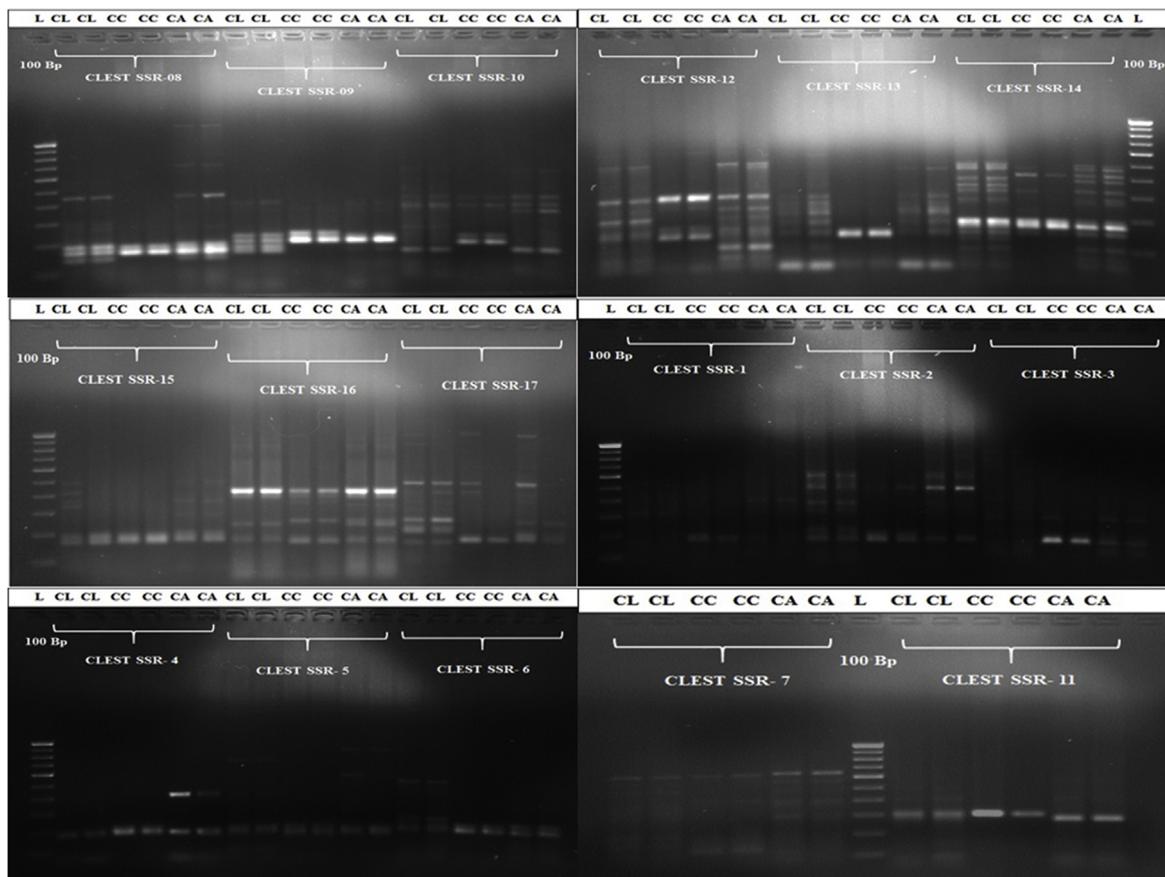


Fig. 3. Banding patterns of EST-SSR fragments of three targeted *Curcuma* species. CL - *Curcuma longa*, CC - *Curcuma caesia*, CA - *Curcuma aromatica*.

Table 3

List of selected informative EST-SSR primers with percentage of polymorphic bands (%) and their polymorphic information content (PIC).

Locus	Expected product size (bp)	Allele size (bp)	Total band	Polymorphic band	Monomorphic band	Polymorphism	PIC
CLEST SSR-01	180	189–157	4	4	0	100%	0.38
CLEST SSR-02	183	204–152	7	5	2	71.42%	0.30
CLEST SSR-03	172	173–133	8	8	0	100%	0.41
CLEST SSR-04	184	188–172	3	2	1	66.67%	0.30
CLEST SSR-05	184	185–175	9	8	1	88.89%	0.43
CLEST SSR-06	202	199–191	6	5	1	83.34%	0.33
CLEST SSR-07	152	181–104	8	7	1	87.50%	0.40
CLEST SSR-08	171	181–154	8	6	2	75%	0.41
CLEST SSR-09	188	218–184	4	3	1	75%	0.46
CLEST SSR-10	196	200–188	8	6	2	75%	0.48
CLEST SSR-11	209	305–292	7	7	0	100%	0.38
CLEST SSR-12	150	162–140	10	6	4	60%	0.43
CLEST SSR-13	154	158–136	9	7	2	77.78%	0.30
CLEST SSR-14	176	186–166	12	9	3	75%	0.25
CLEST SSR-15	172	198–162	6	5	1	83.34%	0.33
CLEST SSR-16	173	174–164	7	5	2	71.43%	0.33
CLEST SSR-17	174	192–174	11	9	2	81.81%	0.45
Total			127	102	25	81%	0.38

into contigs for creating a non-redundant dataset using EGassembler (Masoudi-Nejad et al., 2006). The identification of Class I (hypervariable) microsatellite repeats (Temnykh et al., 2001) in the generated non-redundant EST dataset was calculated using MSATCOMMANDER (Faircloth et al., 2008). When two SSRs were present close to each other in one EST, they were counted as individual SSRs rather than compound SSRs (Gupta, 2003). Primers were designed only for Class I SSR containing EST sequences using primer 3 (Rozen and Skaletsky, 2000) and also manually. The quality of the designed primers was validated using primers having a score of more than 75 as evidenced by the

absence of self-dimer or primer dimer or cross dimers were selected for synthesis. A selected set of 17 primers was synthesized at Bioserve Hyderabad, India. These primers were tested for functionality and polymorphisms against three species of *Curcuma*. The putative functions of sequences containing polymorphic EST-SSRs were detected using BLAST X (Altschulet al., 1997) by comparing against the non-redundant protein database (see Fig. 1).

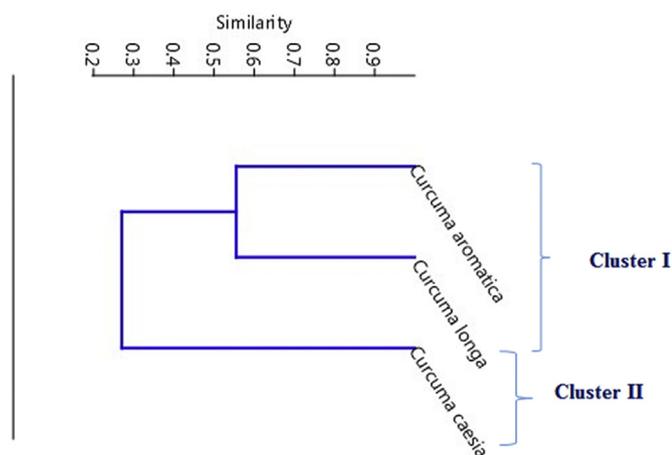


Fig. 4. Dendrogram showing clustering of *Curcuma* species constructed by using UPGMA cluster analysis of genetic similarity based on EST- SSR data by Jaccard similarity matrix (0.9868).

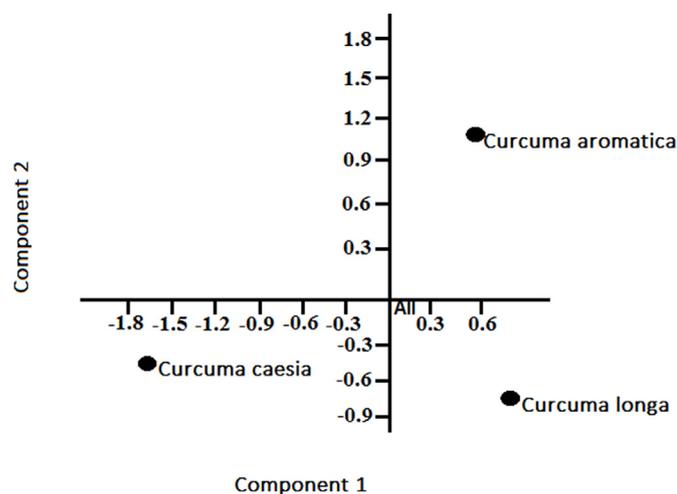


Fig. 5. Principle component analysis according to EST- SSR primers of *Curcuma* species.

2.3. Genetic diversity analysis by molecular markers

Seventeen EST-SSR primers (Oligos) (Bioserve, Hyderabad, India) were used for PCR amplification. Based on results, good resolution and reproducibility ability, EST-SSRs primers were selected for several primers utilized during screening. PCRs with a single primer were carried out in a final volume of 20 μ l containing (2.5 μ l) 50 ng template DNA, 2 μ l 2 mM of dNTP mix (Himedia), 2 μ l 10 mM of oligonucleotides synthesized primer, 2 μ l 1X Taq buffer and 0.3 μ l 2U Taq DNA polymerase (Thermo fisher). Amplification was performed in a thermal cycler (Eppendorf and Biorad). The detection of microsatellite polymorphism was performed using 17 EST- SSR markers characterized by Siju et al. (2010). The EST-SSR amplification condition was as follows: an initial hot start and denaturing step at 95 $^{\circ}$ C for 5 min followed by 40 cycles of a 1 min denaturation at 94 $^{\circ}$ C, a 1 min annealing based on primer tm, and a 1 min primer elongation at 72 $^{\circ}$ C. A final extension step was performed at 72 $^{\circ}$ C for 5 min. The PCR amplified products were resolved in a 4% agarose gel. Electrophoresis was done for about 2.5 h at 60 V. The molecular weights of amplified products were compared using 100bp ladder and visualization of the amplified bands by gel documentation system BioRad.

2.4. Data scoring and analysis

Only clear, unambiguous and reproducible bands were considered for data analysis. Each band was considered to be a single locus. Data were scored as "1" for presence and "0" for absence. To avoid taxonomic ambiguities, the intensity of the bands was not taken into considerations, only the presence of band was taken as indicative. The binary data of the ISSR and SSR fingerprints were used for population genetic analyses. The numbers of monomorphic and polymorphic bands were derived from the binary data, and their percentages were calculated. The level of similarity between species, percentage of polymorphic bands and a matrix of genetic similarity were compiled using Jaccard's similarity coefficient (JSI) (Jaccard, 1908). Similarity, coefficients were used to construct the dendrogram between species using the unweighted pair group method with arithmetic average (UPGMA) and the sequential hierarchical and clustering with the help of Past 3 software, representing genetic relationship between 3 different species of *Curcuma* (*C. longa*, *C. caesia* and *C. aromatica*). The polymorphism information content (PIC) for each marker was calculated with the formula described by Roldan-Ruiz et al. (2006); Zhi-Hui Guo et al. (2014): $PIC_i = 2fi(1 - fi)$ where PIC_i is the polymorphic information content of marker i, fi the frequency of the marker bands which were present, and $(1 - fi)$ the frequency of marker bands which were absent. Other basic parameters for genetic diversity were calculated in the POPGENE application. The number of different alleles (na), the mean number of effective alleles (ne) $No. \text{ of Effective Alleles} = 1/(p^2 + q^2)$ and the Shannon Information index was calculated by formula $I = -1 * (p * \ln(p) + q * \ln(q))$ index (I), $He = \text{Expected Heterozygosity} = 2 * p * q$ and $uHe = \text{Unbiased Expected Heterozygosity} = (2N/(2N-1)) * He$ were calculated by above mentioned formula, where for Diploid Binary data and assuming Hardy-Weinberg Equilibrium, $q = (1 - \text{Band Freq.})^{0.5}$ and $p = 1 - q$.

3. Results

3.1. EST-SSR polymorphism

Primers designed from EST and CDS regions of gene for expression studies is presented in Table 1. Table 2. Shows sequences and nucleotide length of primers used for EST- SSR analysis was showed in the integrity of isolated genomic DNA of three targeted *Curcuma* species i.e., *C. longa*, *C. caesia* and *C. aromatica* was visualized in 0.8% agarose gel (Fig. 2.). Banding pattern of all seventeen EST-SSR primers was amplified as shown in Fig. 3. Table 3 shows the diversity of EST-SSR primers with percentage of polymorphic bands (%) and their polymorphic information content (PIC). The primer combinations had amplified 127 loci among which 102 were found to be polymorphic in nature and the rest were monomorphic. Maximum number of 12 bands were resolved for the primer CLEST SSR-14 and the minimum 4 bands for CLEST SSR-4. Maximum number of polymorphism was found in primer CLEST SSR-1, CLEST SSR -3, CLEST SSR-11(100%) and minimum in primer CLEST SSR-12 (60%). The average PIC of all the primers was 0.38. The bands resolved in the range between 310 and 100 bp were consider in the present investigation. Jaccard's coefficient by UPGMA algorithm showed that the species were most closely related to a similarity value 0.9868 and Simpson similarity matrix was 0.5744. Cluster analysis, principle component analysis and genetic diversity index of three *Curcuma* species showed in Figs. 4–6 respectively. Genetic variation between *Curcuma* species calculated by different parameters using popgene software Table 4. However, Similarity matrix among species was evaluated by past 3 software Table 5.

3.2. Data scoring and cluster analysis

The dendrogram was constructed through UPGMA algorithm using Jaccard's similarity coefficient of EST-SSR through Past 3 software. In

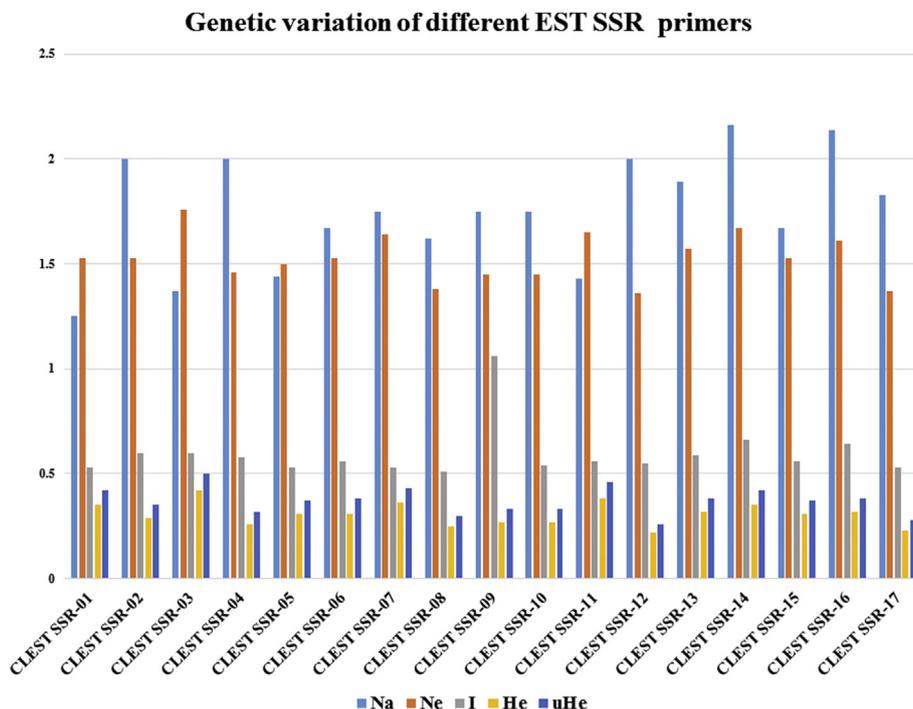


Fig. 6. Graph showing the Genetic variation of three targeted *Curcuma* species in different EST-SSR primers.

Table 4
Genetic variation of three *Curcuma* species in different EST-SSR primers.

SNO	EST-SSR primers	Band Freq.	p	q	N	Na	Ne	I	He	uHe
1	CLEST SSR-01	0.42 ± 0.17	0.24 ± 0.12	0.76 ± 0.12	3	1.25 ± 0.5	1.53 ± 0.26	0.53 ± 0.10	0.35 ± 0.09	0.42 ± 0.11
2	CLEST SSR-02	0.67 ± 0.27	0.52 ± 0.35	0.48 ± 0.35	3	2 ± 0.82	1.53 ± 0.43	0.60 ± 0.15	0.29 ± 0.22	0.35 ± 0.26
3	CLEST SSR-03	0.46 ± 0.17	0.33 ± 0.12	0.67 ± 0.12	3	1.37 ± 0.52	1.76 ± 0.27	0.60 ± 0.11	0.42 ± 0.09	0.50 ± 0.11
4	CLEST SSR-04	0.67 ± 0.27	0.54 ± 0.34	0.46 ± 0.34	3	2 ± 0.82	1.46 ± 0.39	0.58 ± 0.13	0.26 ± 0.2	0.32 ± 0.24
5	CLEST SSR-05	0.48 ± 0.24	0.32 ± 0.27	0.67 ± 0.27	3	1.44 ± 0.73	1.50 ± 0.29	0.53 ± 0.19	0.31 ± 0.14	0.37 ± 0.17
6	CLEST SSR-06	0.56 ± 0.27	0.4 ± 0.32	0.60 ± 0.32	3	1.67 ± 0.82	1.53 ± 0.37	0.56 ± 0.19	0.31 ± 0.18	0.38 ± 0.21
7	CLEST SSR-07	0.58 ± 0.18	0.41 ± 0.13	0.59 ± 0.13	3	1.75 ± 0.53	1.64 ± 0.28	0.53 ± 0.11	0.36 ± 0.10	0.43 ± 0.12
8	CLEST SSR-08	0.54 ± 0.31	0.42 ± 0.37	0.58 ± 0.37	3	1.62 ± 0.91	1.38 ± 0.30	0.51 ± 0.52	0.25 ± 0.17	0.30 ± 0.20
9	CLEST SSR-09	0.58 ± 0.32	0.45 ± 0.38	0.55 ± 0.38	3	1.75 ± 0.96	1.45 ± 0.39	1.06 ± 0.81	0.27 ± 0.20	0.33 ± 0.24
10	CLEST SSR-10	0.58 ± 0.30	0.45 ± 0.36	0.55 ± 0.36	3	1.75 ± 0.89	1.45 ± 0.36	0.54 ± 0.67	0.27 ± 0.19	0.33 ± 0.22
11	CLEST SSR-11	0.48 ± 0.18	0.29 ± 0.13	0.71 ± 0.13	3	1.43 ± 0.53	1.65 ± 0.28	0.56 ± 0.11	0.38 ± 0.1	0.46 ± 0.12
12	CLEST SSR-12	0.67 ± 0.31	0.56 ± 0.39	0.44 ± 0.39	3	2 ± 0.94	1.36 ± 0.37	0.55 ± 0.51	0.22 ± 0.2	0.26 ± 0.24
13	CLEST SSR-13	0.63 ± 0.26	0.47 ± 0.32	0.53 ± 0.32	3	1.89 ± 0.78	1.57 ± 0.40	0.59 ± 0.64	0.32 ± 0.2	0.38 ± 0.24
14	CLEST SSR-14	0.72 ± 0.19	0.55 ± 0.28	0.45 ± 0.28	3	2.16 ± 0.57	1.67 ± 0.43	0.66 ± 0.41	0.35 ± 0.22	0.42 ± 0.26
15	CLEST SSR-15	0.55 ± 0.27	0.39 ± 0.32	0.60 ± 0.32	3	1.67 ± 0.82	1.53 ± 0.37	0.56 ± 0.95	0.31 ± 0.17	0.37 ± 0.21
16	CLEST SSR-16	0.71 ± 0.23	0.55 ± 0.32	0.44 ± 0.32	3	2.14 ± 0.69	1.61 ± 0.45	0.64 ± 0.27	0.32 ± 0.23	0.38 ± 0.27
17	CLEST SSR-17	0.61 ± 0.32	0.49 ± 0.40	0.50 ± 0.40	3	1.83 ± 0.98	1.37 ± 0.35	0.53 ± 0.69	0.23 ± 0.19	0.28 ± 0.23

Na = No. of Different Alleles.
 Ne = No. of Effective Alleles = $1/(p^2 + q^2)$.
 I = Shannon's Information Index = $-1 * (p * \ln(p) + q * \ln(q))$.
 He = Expected Heterozygosity = $2 * p * q$.
 uHe = Unbiased Expected Heterozygosity = $(2N/(2N-1)) * He$.
 Where for Diploid Binary data and assuming Hardy-Weinberg Equilibrium, $q = (1 - \text{Band Freq.})^{0.5}$ and $p = 1 - q$.

Table 5
Genetic similarity matrix among three *Curcuma* species by using EST-SSR primer.

<i>Curcuma</i> Species	<i>Curcuma longa</i>	<i>Curcuma caesia</i>	<i>Curcuma aromatica</i>
<i>Curcuma longa</i>	1		
<i>Curcuma caesia</i>	0.24299065	1	
<i>Curcuma aromatica</i>	0.55555556	0.2970297	1

EST-SSR primers similarity matrix of *Curcuma* species revealed that it divided the dendrogram in two main clusters. Cluster I containing *C. longa* and *C. aromatica* however, *C. caesia* place in different cluster due

to the presence of greater polymorphism, genetic diversity, environmental, soil, agroclimatic and its genetic makeup. *C. longa* and *C. aromatica* again divided in different group due to minor changes in its genetic sequence. Principle component analysis was also performed to interpret the diversity index or to know the position of our components. (Fig. 5).

4. Discussion

Curcuma is a huge genus displaying diversity in habitat, ethno botanical use, and morphology among other genus of Zingiberaceae family (Syamkumar and Sasikumar, 2007). Detailed study about genetic

relationships among wild and cultivated species of *Curcuma* may enhance its utilization value for biotechnological innovations. Several studies have been conducted earlier based on morphological, anatomical, and biochemical characterization of *Curcuma* species and cultivars have been attempted earlier (Paisooksantivatana et al., 2001; Jiang et al., 2006; Zhou et al., 2007; Das et al., 2011; Policegoudra et al., 2008; and Paramasivam et al., 2009) there is scarcity of information regarding phylogenetic relationship among *Curcuma* of different states of India. Moreover, no attempt had been undertaken to systematically study the genetic diversity of *Curcuma* wild, endemic and cultivated varieties of Chhattisgarh. Therefore, in this study, an attempt was done to calculate the existence polymorphism in different species of *Curcuma* from Chhattisgarh and store germplasm for their cultivation and preservation. The diversity analysis efficiency and discriminatory power of the DNA based markers largely depends upon the rate of polymorphism, differential gene expression pattern was clearly detected amongst the different species of *Curcuma*. Marker system used in present studies are highly informative, cost effective, and require less time and labour for fingerprinting of genome. Results of the biochemical analysis revealed that the curcumin content was found to be highest in *C. longa* as compare to *C. caesia* and *C. aromatica* due to high expression of curcumin synthesis gene. The 17 polymorphic primer pairs used here were amplified multiple reproducible banding patterns with high intensity in all tested *Curcuma* species. Amplification of multiple bands through EST-SSR marker in the present study could be due to the genomic diversity and differential gene expression nature of turmeric.

In the present study, we have established for the first time, a unique DNA profiling of three *Curcuma* species using gene specific EST-SSR primers. ESTs accurately reflects the density of SSRs in the transcribed and coding portions of the genome (Varshney et al., 2005). Secondary metabolite and the genes involved in their pathways can be determined by studying functional genomics. Expression studies of genes involved in curcuminoid synthesis in *Curcuma* species studied by gene specific marker CURS, CURS 2 and CHS (Chalcone synthase) (Behar et al., 2016) and germacrone synthase. Chalcone synthase is key enzyme for flavonoid synthesis that possess antibacterial property. Curcuminoid synthase gene expressed in all the selected *Curcuma* species and exhibited genetic diversity in the species level. Seventeen polymorphic EST-SSR primer pairs used in the study were amplified for multiple reproducible banding patterns with high intensity. Amplification of multiple bands through EST-SSR marker may be due to the gene expression level of different turmeric species as well as environmental factors.

Multiple bands produced by SSRs in the study is in agreement with reported works (Siju et al., 2010; Singh et al., 2015). In conclusion, the present paper, with its reliable and reproducible nature results, indicated that a standard set of SSR markers (CSSR 14, CSSR 1, CSSR 3 and CSSR 11) can be efficiently used for authentic identification of the commercially important *Curcuma* species cultivars.

5. Conclusion

Molecular marker-based study of genetic variations facilitates employing in suggestive of an evolutionary pattern among *Curcuma* species exploration. Furthermore, the scientific data presented here indicated that the application of PCR-based fingerprinting using whole DNA and arbitrary primers may provide a rapid and sensitive method for detection of genetic variations among the different species of *Curcuma* from Chhattisgarh. Evaluation of *Curcuma* species diversity would be of great significance for *in situ* conservation of important *Curcuma* species especially for their long term medicinal practices and to utilize the full medical potential of high quality chemical constituents found in *Curcuma* species.

Acknowledgments

Akanksha Jain gratefully acknowledges Department of Biotechnology, Guru Ghasidas Vishwavidyalaya Bilaspur (A Central University) to provide infrastructure and financial aid. Author also thank Dr. Rakesh Kumar Meena, Dr. Pulak Das and Mr. Dev Prakash Sahu [Aditya Biotech lab & Reasearch Pvt. Ltd, Raipur (C.G.)] for providing the lab facilities and experimental guidance.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bcab.2019.101033>.

Conflicts of interest

The authors declared no conflict of interest.

References

- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., et al., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.
- Apavatjrit, P., Somboon, A., Puangpen, S., Chiara, A., 1999. Molecular markers in the identification of some early flowering *Curcuma* L. (Zingiberaceae) species. *Ann. Bot.* 84, 529–534.
- Barakat, M.N., Abdel Fattah, R.S., Badr, M., El-Torky, M.G., 2010. In vitro mutagenesis and identification of new variants via RAPD markers for improving *Chrysanthemum morifolium*. *Afr. J. Agric. Res.* 8, 748–757.
- Behar, N., Tiwari, K.L., Jadhav, S.K., 2016. Semi- quantitative expression studies of genes involved in biosynthesis of curcuminoid in *Curcuma caesia* Roxb. *IJBT* 15, 491–494.
- Das, A., Kesari, V., Satyanarayana, V.M., Parida, A., Rangan, L., 2011. Genetic relationship of *Curcuma* species from Northeast India using PCR-based markers. *Mol. Biotechnol.* 49, 65–76.
- Dreisigacker, S., Zhang, P., Warburton, M.L., Skovmand, B., Hoisington, D., Melchinger, A.E., 2005. Genetic diversity among and within CIMMYT wheat landrace accessions investigated with SSRs and implications for plant genetic resources management. *Theor. Appl. Genet.* 101 (1–2), 653.
- Edwards, K., Johnstone, C., Thompson, C., 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.* 19, 1349.
- Faircloth, B.C., 2008. MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* 8, 92–94.
- Guo, Zhi-Hui, Fu, Kai-Xin, Zhang, Xin-Quan, Bai, Shi-Qie, Fan, Yan, Peng, Yan, Huang, Lin-Kai, Yan, Yan-Hong, Liu, Wei, Xiao, Ma, 2014. Molecular insights into the genetic diversity of *Hemarthria compressa* germplasm. *Collections Native to Southwest China Molecules* 19, 21541–21559 2014.
- Gupta, P.K., Rustgi, S., Sharma, S., Singh, R., Kumar, N., Balyan, H.S., 2003. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Mol. Genet. Genom.* 270, 315–323.
- Islam, M.A., Meister, A., Schubert, V., Kloppstech, K., Esch, E., et al., 2007. Genetic diversity and cytogenetic analysis in *Curcuma zedoaria* (Christm.) Roscoe. from Bangladesh. *Genet. Resour. Crop Evol.* 54, 149–156.
- Jaccard, P., 1908. Nouvelles recherches sur la distribution florale. *Bulletin de la Societe Vaudoise des Science Naturelles* 44, 223–270.
- Jatoi, S.A., Kikuchi, A., Yi, S.S., Naing, K.W., Yamanaka, S., Junko, A., et al., 2006. Use of SSR markers as RAPD markers for genetic diversity analysis in Zingiberaceae. *Breed Sci.* 56, 107–111.
- Jiang, H., Xie, Z., Koo, H.J., McLaughlin, S.P., Timmermann, B.N., Gang, D.R., et al., 2006. Metabolic profiling and phylogenetic analysis of medicinal Zingiber species tools for authentication of gingers (Zingiberofficinale). *Phytochemistry* 67, 1673–1685.
- Kalia, R.K., Rai, M.K., Kali, S., Singh, R., et al., 2011. Microsatellite markers: an overview of the recent progress in plants. *Euphytica* 177, 309–334.
- Kumpatla, S.P., Mukhopadhyay, S., 2005. Mining and survey of simple sequence repeats in expressed sequence tags of dicotyledonous species. *Genome* 48, 985–998.
- Lapitan, V.C., Brar, D.S., Abe, T., Redona, E.D., 2007. Assessment of genetic diversity of Philippine rice cultivars carrying good quality traits using SSR markers. *Breed Sci.* 57, 263–270.
- Lee, S.Y., Fai, W.K., Zakaria, M., Ibrahim, H., Othman, R.Y., Gwag, J.G., Rao, V.R., Park, Y.J., et al., 2007. Characterization of polymorphic microsatellite markers, isolated from ginger (ZingiberofficinaleRosc.). *Mol. Ecol. Notes* 7, 1009–1011.
- Masoudi-Nejad, A., Tonomura, K., Kawashima, S., Moriya, Y., Suzuki, M., Itoh, M., et al., 2006. E-Gassembler: online bioinformatics service for large-scale processing, clustering and assembling ESTs and genomic DNA fragments. *Nucleic Acids Res.* 34, 459–462.
- Mohanta, S., Pk, S., Sial, P., Gr, R., 2015. Morphological and molecular screening of turmeric (*Curcuma longa* L.) cultivars for resistance against parasitic Nematode *Meloidogyne incognita*. *Plant Pathology & Microbiology* 6.
- Naik, P.K., Alam, M.A., Singh, H., Goyal, V., Parida, S., Kalia, S., Mohapatra, T., et al., 2010. Assessment of genetic diversity through RAPD, ISSR and AFLP markers in *Podophyllumhexandrum*: a medicinal herb from northeastern Himalayan region.

- Physiol. Plantarum 16 (2), 145–148.
- Noli, E., Cont, S., Maccaferri, M., Sanguinetti, M.C., et al., 1997. Molecular characterization of Tomato cultivars. Seed Sci. Technol. 27, 1–10.
- Nowbuth, P., Khittoo, G., Bahorun, T., Venkatasamy, S., 2005. Assessing genetic diversity of some Anthuriumandraeanum Hort. cut-flower cultivars using RAPD Markers. Afr. J. Biotechnol. 4, 1189–1194.
- Paisooksantivatana, Y., Kako, S., Seko, H., 2001. Genetic diversity of *Curcuma alismatifolia* Gagnep (Zingiberaceae) in Thailand as revealed by allozyme polymorphism. Genet. Resour. Crop Evol. 48, 459–465.
- Paramasivam, M., Poi, R., Banerjee, H., Bandyopadhyay, A., 2009. High-performance thin layer chromatographic method for quantitative determination of curcuminoids in *Curcuma longa* germplasm. Food Chem. 113, 640–644.
- Policegoudra, R.S., Aradhya, S.M., 2008. Biochemical changes and antioxidant activity of mango ginger (*Curcuma amada* Roxb.) rhizomes during post harvest storage at different temperatures. Postharvest Biol. Technol. 46, 189–194.
- Roldán-Ruiz, I., Dendauw, J., Van Bockstaele, E., Depicker, A., de Loose, M., 2000. AFLP markers reveal high polymorphic rates in ryegrasses (*Lolium* spp.). Mol. Breed. 6, 125–134.
- Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. In: Misener, S. Krawetz S. (Ed.), Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, USA, pp. 365–386.
- Sharma, A., Naredo, A.G., Mahadik, K.R., 2008. Molecular markers: net prospects in plant genome analysis. Phcog. Rev. 2 (3), 23–34.
- Siju, K., Dhanya, S., Syamkumar, B., Sasikumar, T.E., Sheeja, A.I., Bhat, V., Parthasarathy, A., 2010. Development, characterization and cross species amplification of polymorphic microsatellite markers from expressed sequence tags of turmeric (*Curcuma longa* L.). Mol. Biotechnol. 44, 140–147.
- Singh, A.K., Nanda, P., Singh, A., Singh, B., 2015. Genetic diversity analysis in turmeric (*Curcuma longa* L.) based on SSR markers. Journal of Biological Engineering Research and Review 2 (1), 20–24.
- Syamkumar, S., Sasikumar, B., 2007. Molecular marker based genetic diversity analysis of *Curcuma* species from India. Sci. Hortic. 112, 235–241.
- Taheri, S., Abdullah, T.L., Abdullah, N.A., Ahmad, Z., 2012. Genetic relationships among five varieties of *Curcuma alismatifolia* (Zingiberaceae) based on ISSR markers. Mol. Res. 11, 3069–3076.
- Temnykh, S., DeClerck, G., Lukashova, A., Lipovich, L., Cartinhour, S., McCouch, S., 2001. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res. 11, 1441–1452.
- Tian, H.L., Chen, X.Q., Wang, J.X., Xue, J.H., 2008. Development and characterization of microsatellite loci for lotus (*Nelumbonucifera*). Conserv. Genet. 9, 1385–1388.
- Varshney, R.K., Graner, A., Sorrells, M.E., 2005. Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 23, 48–55.
- Williams, J.K.F., Kubelik, A.R., Livak, K.G., Rafalki, J.A., Tingey, S.V., et al., 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18, 6531–6535.
- Zeitkiewicz, E., Rafalski, A., Labuda, D., et al., 1994. Genome finger printing by simple sequence repeat (SSR)-anchored PCR amplification. Genomics 20, 176–183.
- Zhou, X., Zhangwan, L., Liang, G., Zhuh, J., Wang, D., Cai, Z., 2007. Analysis of volatile components of *Curcuma sichuanensis* X. X. Chen by gas chromatography–mass spectrometry. J. Pharmaceut. Biomed. Anal. 43, 440–444.