Biocatalysis and Agricultural Biotechnology 21 (2019) 101319

Contents lists available at ScienceDirect

Biocatalysis and Agricultural Biotechnology

journal homepage: http://www.elsevier.com/locate/bab

DNA based molecular markers discriminate genders of commercially important dioecious tree Kokum, *Garcinia indica* (choicy)

Reshma V. Patil, Kiran D. Pawar *

School of Nanoscience and Biotechnology, Shivaji University, Kolhapur, Maharashtra, India

ARTICLE INFO

Keywords:
Garcinia indica
Molecular markers
Gender discrimination
Gender polymorphism

ABSTRACT

Garcinia indica choicy (Thaurus-Dupettite) is commercially important, endemic, polygamadioecious tree species of the Western Ghats of India. Genetic improvement and marker-assisted breeding of *G. indica* has been slow due to the lack of efficient molecular markers. In the present study, we mined and identified simple sequence repeat (SSRs) loci from male, female and bisexual flower transcriptome data and used along with ISSR and RAPD to assess their efficacy to discriminate gender of Kokum plants.'AG/TC' was the most common SSR motif observed within the transcriptome. The analyses of profiles generated from these DNA markers revealed a high degree of inter-gender polymorphism while the level was considerably low among the accessions of the same gender. The UPGMA clustering based on ISSR showed better gender-wise resolution and discrimination than SSR, RAPD, and combined data. Further, assessments based on UPGMA dendrogram and principal component analysis (PCA) demonstrated that the performances for gender discrimination by tested markers were in the order of merit as ISSR > RAPD > SSR. In addition, ISSR amplified using primer UBC-818 produced specific band for male and female while primer UBC-864 produced specific band for bisexual demonstrating that they could be used for discrimination and identification of the gender of *G. indica*.

1. Introduction

Economically important, underexploited fruit bearing tree *Garcinia indica* choicy (Thaurus-Dupettite) is mainly found in the Western Ghats of India. The fruit rind of Kokum, as it is commonly known, is a rich source of an important active metabolite hydroxycitric acid (HCA) that shows the anti-obesity property (Heymsfield et al., 2014). The major constraint in popularizing this tree species as a potential horticultural crop is the polygamadioecious nature. Differentiation between male and female trees can be done only at the flowering stage which occurs at almost 7–9 years of age. Raising seedlings throughout the year is difficult as seeds are recalcitrant and have a short shelf life (Malik et al., 2005). Though softwood grafting can be practiced for clonal propagation (Haldankar et al., 1993; Nawala and Karmarkar, 1997), it is seldom practiced for reasons such as being season dependent, the requirement of large space and scarcity of rootstocks for grafting, etc. In addition, complex genetic systems together with extended juvenility of this plant have greatly hampered conventional breeding for improvement of this fruit-bearing tree. The determination of sex at an early stage in Kokum is of utmost importance for the commercial plantation of Kokum orchard. In agriculture, knowing the gender of the plant at the early stage prior to floral initiation, and then selection of the appropriate gender types of the progeny for commercial planting would be of commercial importance as only female and hermaphrodite plants bear fruits (Magdalita and Mercado, 2003). Current available morphological markers that can be employed for the purpose of plant gender identification at an early stage are insufficient (Sawardekar et al., 2011). In the view of this unavailability of effective methods for gender identification of *G.indica* at an early stage, suitable, alternate and effective marker system needs to be developed. With a great deal of advancement in the field of genomics tools, considerable progress has been made in understanding the evolution of sex-determining genes and chromosomes in model plant species. The establishment of modern genomics research in strawberry (*Fragaria*), melon (*Cucumismelo*), asparagus (*Asparagus officinalis*), persimmon (*Diospyros lotus*), papaya (*Carica papaya*), white campion (*S. latifolia*) as well as the, sorrel (*Rumex)*, poplar (*Populus*) and willow (*Salix*) genera have shown important findings about the early and progressive evolution of sex chromosomes in angiosperms (Harkess and Leebens-Mack, 2017). All dioecious species have sex chromosomes, which carry genes that control the development of males and females as

* Corresponding author.
E-mail addresses: kdp.snst@unishivaji.ac.in, pawarkiran1912@gmail.com (K.D. Pawar).

https://doi.org/10.1016/j.bcab.2019.101319
Received 21 May 2019; Received in revised form 26 August 2019; Accepted 27 August 2019
Available online 5 September 2019
1878-8181/

Table 1
Details of accessions and locations of male, female and bisexual plants of *G. indica.*

No	Code	Places	Coordinates Latitude	Longitude	Altitude (m)
ML					
1	gi_ML_SS	Shiroda, Sindhudurg, Maharashtra	16°60′ N	73°19′ E	63
2	gi_ML_OS	Otavane, Sindhudurg, Maharashtra	16°60′ N	73°19′ E	40
3	gi_ML_KS	Kasal, Sindhudurg, Maharashtra	16°60′ N	73°19′ E	35
4	gi_ML_VS	Vajarat, Sindhudurg, Maharashtra	16°80′ N	73°06′ E	11
5	gi_ML_WS	Wayangani, Sindhudurg, Maharashtra	16°80′ N	73°06′ E	11
6	gi_ML_VeS	Vetore, Sindhudurg, Maharashtra	16°80′ N	73°06′ E	11
7	gi_ML_DS	Dabholi, Sindhudurg, Maharashtra	16°80′ N	73°06′ E	11
8	gi_ML_JS	Jambhavade, Sindhudurg, Maharashtra	16°60′ N	73°19′ E	26
9	gi_ML_UR	Dr. Balasaheb Sawant Agriculture University, Ratnagiri, Maharashtra	17°48′ N	73°12′ E	11
10	gi_ML_LR	Lote, Ratnagiri, Maharashtra	17°30′ N	73°36′ E	137
11	gi_ML_CR	Chiplun, Ratnagiri, Maharashtra	17°30′ N	73°36′ E	7
12	gi_ML_PS	Parab Farm, Sindhudurg, Maharashtra	16°34′ N	73°75′ E	42
13	gi_ML_KaS	Karul, Sindhudurg, Maharashtra	16°30′ N	73°74′ E	80
14	gi_ML_SaS	Sawantwadi, Sindhudurg, Maharashtra	15°54′ N	73°49′ E	112
FL					
15	gi_FL_SS	Shiroda, Sindhudurg, Maharashtra	16°60′ N	73°19′ E	63
16	gi_FL_OS	Otavane, Sindhudurg, Maharashtra	16°60′ N	73°19′ E	40
17	gi_FL_KS	Kasal, Sindhudurg, Maharashtra	16°60′ N	73°19′ E	35
18	gi_FL_VS	Vajarat, Sindhudurg, Maharashtra	16°80′ N	73°06′ E	11
19	gi_FL_WS	Wayangani, Sindhudurg, Maharashtra	16°80′ N	73°06′ E	11
20	gi_FL_VeS	Vetore, Sindhudurg, Maharashtra	16°80′ N	73°06′ E	11
21	gi_FL_DS	Dabholi, Sindhudurg, Maharashtra	16°80′ N	73°06′ E	11
22	gi_FL_TS	Tarkarli, Sindhudurg, Maharashtra	16°03′ N	73°48′ E	5
23	gi_FL_AS	Achra, Sindhudurg, Maharashtra	16°03′ N	73°47′ E	2
24	gi_FL_JS	Jambhavade, Sindhudurg, Maharashtra	16°60′ N	73°19′ E	26
25	gi_FL_UR	Dr. Balasaheb Sawant Agriculture University, Ratnagiri, Maharashtra	17°48′ N	73°12′ E	11
26	gi_FL_LR	Lote, Ratnagiri, Maharshtra	17°30′ N	73°36′ E	137
27	gi_FL_PS	Parab Farm, Sindhudurg, Maharashtra	16°34′ N	73°75′ E	42
BS					
28	gi_BS_K$_1$S	Kasal1, Sindhudurg, Maharashtra	16°60′ N	73°19′ E	35
29	gi_BS_K$_2$S	Kasal2, Sindhudurg, Maharashtra	16°60′ N	73°19′ E	35
30	gi_BS_TS	Tarkarli, Sindhudurg, Maharashtra	16°03′ N	73°48′ E	5
31	gi_BS_CR	Chiplun, Ratnagiri, Maharashtra	17°30′ N	73°36′ E	7
32	gi_BS_UR	Dr. Balasaheb Sawant Agriculture University, Ratnagiri, Maharashtra	17°48′ N	73°12′ E	11
33	gi_BS_PS	Parab Farm, Sindhudurg, Maharashtra	16°34′ N	73°75′ E	42
34	gi_BS_Sa$_1$S	Sawantwadi1, Sindhudurg, Maharashtra	15°54′ N	73°49′ E	112
35	gi_BS_Sa$_2$S	Sawantwadi2, Sindhudurg, Maharashtra	15°54′ N	73°49′ E	112
36	gi_BS_KuS	Kudal, Sindhudurg, Maharashtra	16°00′ N	73°68′ E	18
37	gi_BS_OrS	Oros, Sindhudurg, Maharashtra	16°34′ N	73°75′ E	35
38	gi_BS_JaS	Janvale, Sindhudurg, Maharashtra	16°34′ N	73°75′ E	35
39	gi_BS_Sa$_3$S	Sawantwadi3, Sindhudurg, Maharashtra	15°54′ N	73°49′ E	112

separate individuals. The DNA markers responsible for the genetic discrimination of females, males and hermaphrodites were mapped to the male-specific region on Y chromosome (Urasaki et al., 2002). However, hybrid breeding and improved yield are still constrained in plants like castor by genetic instability of female and unknown mechanism of sex expression (Tan et al., 2016).

DNA markers have proved valuable in crop breeding, especially in studies on genetic diversity and gene mapping. It is well established that two molecular typing approaches viz. inter-simple sequence repeat (ISSR) markers (Zietkiezicz et al., 1994) and random amplification of polymorphic DNA (RAPD) (Williams et al., 1990) can be employed for identification of gender of various dioecious taxa (Urasaki et al., 2002; Xu et al., 2004; Danilova and Karlov, 2006; Prakash and Van Staden, 2006; Chaves-Bedoya and Nuñez, 2007). ISSRs have been reported to work well for ascertaining gender in *Carica papaya* (Parasnis et al., 1999; Gangopadhyay et al., 2007), *Humulus lupulus* (Danilova and Karlov, 2006) and *Cycas circinalis* (Gangopadhyay et al., 2007). In addition to ISSR and RAPD, SSR markers are also very useful for a variety of applications in plant breeding and genetics (Powell et al., 1996), whereas, locus-specific primers flanking EST- or genic SSRs can be designed to amplify the microsatellite loci present in the genes. In comparison, genic SSRs which are present in expressed regions of the genome are quickly obtained by electronic sorting and show the advantage over genomic SSRs. Therefore, for an accurate estimation of genetic diversity and identity of genotypes these molecular markers are currently being used (Smith and Helentjaris, 1996).

In *G.indica,* there is a lack of information on the utility of molecular markers for plant gender discrimination and identification. To the best of our knowledge, no study has been conducted to evaluate and utilize the DNA based molecular marker system for gender identification and discrimination in *G.indica*. In the present study, three PCR based DNA marker systems namely, Random Amplified Polymorphic DNA (RAPD), Inter-Simple Sequence Repeat (ISSR) and Simple Sequence Repeat (SSR) were evaluated and compared for their utility to discriminate and identify gender of *G.indica*. Previously developed and analyzed high throughput, next-generation sequence data for male, female and bisexual flowers were used to develop SSRs. These SSRs were then amplified and compared with RAPD and ISSR for their possible utility to discriminate and identify the gender of the *G.indica*.

2. Material and methods

2.1. Collection of plant material

Leaves from confirmed and identified individual male (ML), female (FL) and bisexual (BS) plants were collected in the flowering season during the months of December to Februaryfrom different locations in the Kokan region of Maharashtra (Table 1). Four to six apical and mature leaves from each individual tree were collected in plastic zip bags, labeled, brought to the laboratory and stored immediately at −80 °C until further use in genomic DNA (gDNA) extraction.

2.2. SSR identification, design of SSR primers, SSR amplification and analysis

Recently, for the comparative analysis and understanding of the molecular mechanism of sex determination in *G.indica,* ML, FL and BS flower transcriptomes were sequenced. To this end, we performed high throughput next-generation sequencing of flower transcriptome using Illumina next-generation sequencing platform. The sequence data (>10 GB data per sample) were deposited to NCBI under following accession and SRA project accession numbers as NCBI PRJNA360660 (SRP096940), ML flower (SRS1926202), FL flower (SRS1926203), BS flower (SRS1926204)] (Data under review). The raw data (SRA project accession: SRP096940) generated from three flower were assembled into transcripts, unigenes were generated from the transcripts using TGICL and used to identify SSRs using MicroSAtellite (MISA-http://pgrc.ipk-gatersleben.de/misa, Thiel et al., 2003). Then, using in-house perl script, the flanking region of 150 bp were extracted with parameters for identifying SSRs with di-, tri-, tetra-, penta- and hexa-nucleotide motifs, For the design of SSR primers, Primer3 (Thiel et al., 2003) with following criteria was used: PCR product size range of 100 to 400 bp; primer length of 18–21 nucleotides; GC content of 40%–70% with 50% as optimum and annealing temperature between 50 °C and 70 °C. All SSR primers were tested for their amplification potential using gDNA from *G. indica*. PCR amplifications were performed in a 25 μL reaction containing 50 ng template DNA, 1X assay buffer, 1.5 mM $MgCl_2$, 0.68 mM dNTP's, 0.4 μM each forward and reverse SSR primer and 1U *Taq* polymerase (Bangalore, Genei, India). The thermal cycling conditions for PCR amplification were set on Eppendorf's Nexus GX2 Master Cycler (Eppendorf, Hamburg, Germany) and consisted of denaturation at 95 °C for 7 min; followed by 45 cycles of denaturation at 95 °C for 1 min, annealing at 58.8 °C for 45 s and extension at 72 °C for 1 min. The final extension was done for 7 min at 72 °C.

2.3. Genomic DNA extraction

The gDNAs were extracted by a modified CTAB method (Khanuja et al., 1999). The quality and purity of extracted gDNAs were checked by agarose gel electrophoresis using 0.8% agarose gel with 1x TAE (Tris Acetate EDTA) buffer (Tris base 40 mM, pH 8.0, Acetate 20 mM, EDTA 1 mM pH 8.0). The electrophoretic separation was done at 50 V for 3 h, then stained with ethidium bromide (0.1 μg/ml of gel solution) and imaged on Gel-Doc^{TM}XR+ (Bio-Rad, California, USA). The concentrations of gDNAs were estimated spectrophotometrically, and working stock solutions of 10 ng/μl were prepared and used for amplifications of SSR, ISSR, and RAPDs.

2.4. ISSR amplification and analysis

Eighty ISSR primers from UBC set # 9 were synthesized and supplied by XclerisLabs Pvt. Ltd, (Ahmedabad, India). The amplification efficiency and annealing temperatures of these ISSR primers were initially tested and optimized by setting up PCR reactions in a gradient mode (45–60 °C) for annealing. After optimizing annealing temperature, ISSR amplifications were performed as indicated above for amplification of SSR with few modifications (0.80 mM dNTP's and 0.68 μM ISSR primer).

Table 2
Summary statistics of unigene sequence data mined for SSR identification.

statistics	ML	FL	BS
Total number of sequences examined	49028	45944	49414
Total no of SSR identified	8213	6380	10996
No. of SSR containing sequences	6544	5035	8498
No. of sequences containing more than 1 SSR	1307	1014	1871
Number of SSRs present in compound formation	996	911	1244
Dinucleotide	1799	1393	525
Trinucleotide	1719	1115	834
Tetra, Penta and Hexanucleotide loci	81	85	42

2.5. RAPD amplification and analysis

For the amplification of RAPDs and assessment of their utility to discriminate the gender of the plant by grouping tested accessions into gender-specific groups, 60 RAPD primers from Kit A, B, and C supplied by Operon Technologies (Alameda, CA, USA) were used. The RAPDs were amplified using the PCR reaction composition and thermal cycling conditions which were optimized and used for amplification of SSR and ISSR with few modifications (0.72 mM dNTP's, 1.2 μM RAPD primer and primer annealing at 32.7 °C for 30 sec).

2.6. Agarose gel electrophoresis and data analysis

In order to ensure the reproducibility of PCR amplifications, SSRs, ISSRs, and RAPDs were PCR amplified three times and analyzed for reproducibility and repeatability. Amplified SSR, ISSR, and RAPD fragments were separated using 2.5% and 1.8% agarose gel electrophoresis, respectively as indicated above. The band size was calculated using a 100 bp DNA ladder loaded alongside the amplified PCR products. The amplified bands were scored using Image Lab^{TM} software version 5.2.1 (Bio-Rad, California, USA). The presence or absence of band was recorded as 1 or 0 respectively to generate a binary data matrix. These binary data were then used to generate similarity coefficients (SE) which were further used to build a dendrogram by UPGMA (Unweighted Pair Group Method of Arithmetic average) and cluster analysis using the NTSYS pc 2.02e (Numerical System, Applied Biostatistics, Inc., New York, USA) computer program according to the method proposed by Nei and Li (1979). For estimating the diversity and similarity among and within the genders of *G.indica*, non-parametric analysis of molecular variance (AMOVA) (Excoffier et al., 1992) was performed using GenAlEx (Peakall and Smouse, 2006). Assessment of the genetic relatedness within and among the tested accessions of ML, FL and BS plants and Principal Component Analysis (PCA) were also performed using GenAlEx. For PCA, covariance matrixof pairwise species PhiPT values with data standardization was used as input to generate two-dimensional PCA plot.

3. Results

In the present study, different DNA based molecular markers such as genic SSR, ISSR, and RAPD were amplified and assessed for their efficacy to discriminate, cluster and identify genders of the plants of *G.indica.*

3.1. Identification and amplification of SSRs

The ML, FL, and BS flower transcriptome sequencing analysis, assembly into unigenes and use of MISA identified a total of 8213, 6380, 10996 SSR loci in ML, FL and BS libraries, respectively. Of these, 1307, 1014, 1871 unigene sequences respectively contained more than one SSR. The analysis of the number of repeat size distribution indicated that mono-nucleotide repeat SSRs were more abundant whereas SSRs with penta- and hexa-nucleotides were less abundant (Table 2). The

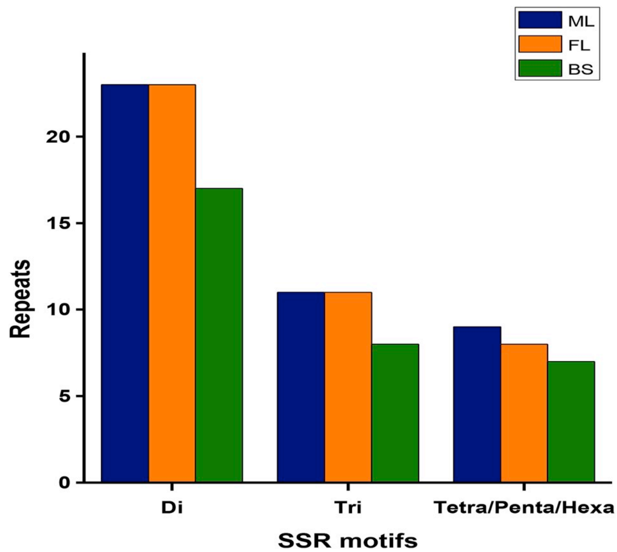

Fig. 1. Frequency distribution of classified di, tri, tetra, penta and hexa-nucleotide repeats in *G. indica* unigenes.

distribution pattern of different SSR repeat units (Table S1) revealed that the most abundant di repeat type with 1799(21.9%) in ML, 1393 (21.8%) in FL and 525(4.7%) in BS were followed by trinucleotide 1719 (21.9%) in ML, 1115(21.8%) in FL and 834(4.7%) in BS. Tetra-, penta-, and hexa-nucleotide repeats were present in low numbers, totaling 81 (0.9%) in ML, 85(1.3%) in FL and 42(0.3%) in BS. The maximum frequency of 775(9.4%) of dinucleotide (AG/TC) SSR repeat motifs was found in ML followed by 587(9.2%) of (AG/GA) in FL and 200(1.8%) of (AG/TC) in BS while 77(0.93%), 60(0.94%), and 44(0.40%) tri-nucleotides (GGA/GAG/GAA) repeat motifs were observed in ML, FL and BS plants respectively. With respect to the number of iterations of SSR motifs, the maximum of 23 iterations were present in identified unigenes (Fig. 1). Based on number of iterations, SSR loci were placed into two classes: class I containing ≥20 nucleotides, also known as hypervariable SSRs, whereas class II, containing ≥12 but <20 nucleotides, contained potentially variable SSRs. Of identified SSRs, 0.28% in ML and 0.36% in FL were assigned to class I and rest to class II.

3.2. Efficacy of SSRs for gender discrimination

In order to validate SSR loci detected from non-redundant unigenes, a total of 50 primer pairs were designed using Primer3. A total of 20 primers were randomly selected and validated in ML, FL and BS accessions of *G.indica.* Of the 20 primers, 15 did not amplify clear and reproducible fragments, one primer showed monomorphic banding pattern while four primers showed the polymorphic banding pattern with reproducible profiles. These five SSR primers produced a total of 43 reproducible and scorable bands out of which 11 (25.5%) were polymorphic with an average of 2 polymorphic bands per primer. The size range of the amplified fragments was 50–400 bps. Total 4 bands each in

Table 3
Characteristics of amplification products obtained from SSR primers.

No	Primer	ML				FL				BS			
		No of bands	Band Size (bp)	Poly morphic band	% polymorphism	No of bands	Band Size (bp)	Poly morphic band	% polymor phism	No of bands	Band Size (bp)	Polymorphic band	% polymorphism
1	Pri12	4	100–400	0	0	4	100–400	0	0	4	100–400	0	0
2	Pri5	2	100–190	1	50	3	100–190	1	33.3	2	100–190	1	50
3	Pri4	2	50–120	1	50	2	50–120	0	0	1	50–120	0	0
4	Pri9	4	50–300	1	25	4	50–300	1	25	4	50–300	2	50
5	Pri11	3	150–280	2	66.6	2	150–280	1	50	2	150–280	0	0

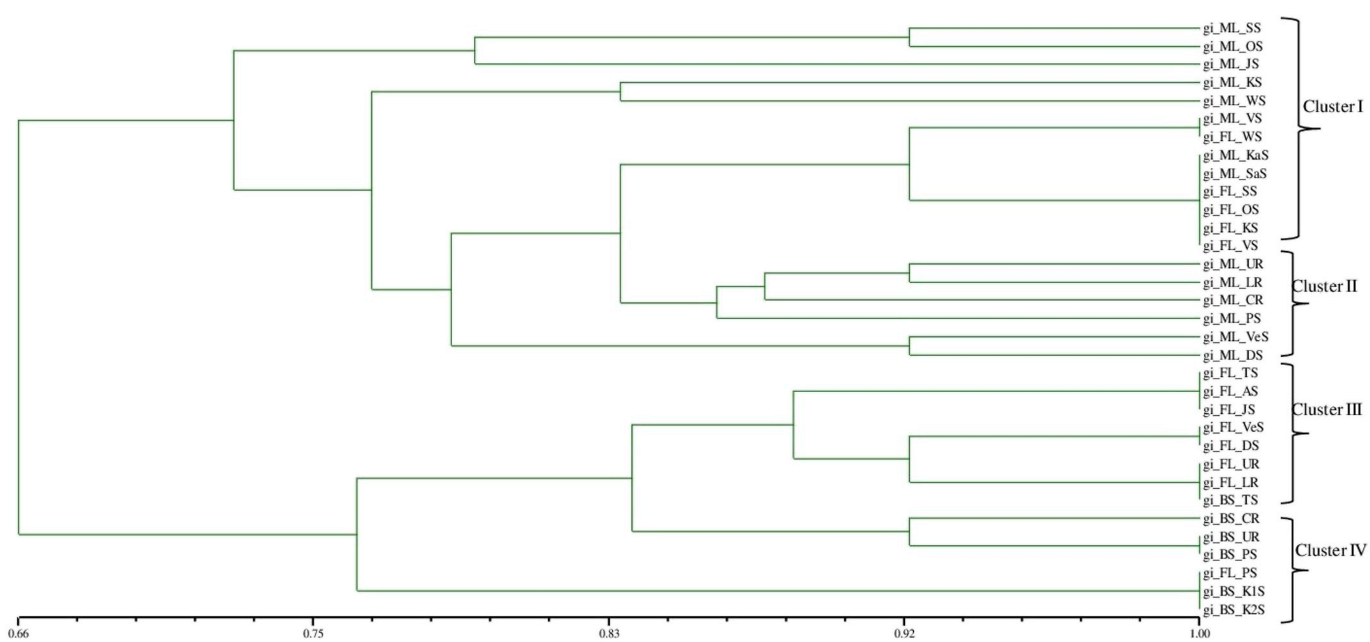

Fig. 2. UPGMA dendrogram based on SSR profiles generated from *G. indica* accessions. The codes in the figure correspond to the codes listed in Table 1.

Table 4
Characteristics of amplification products obtained from ISSR primers.

No	Primer Name	ML				FL				BS			
		No of bands	Band Size (bp)	Poly morphic band	% polymorphism	No of bands	Band Size (bp)	Poly morphic band	% polymor phism	No of bands	Band Size (bp)	Polymorphic band	% polymorphism
1	UBC803	6	100–1600	4	66.6	4	100–1600	0	0	3	100–1600	0	0
2	UBC859	11	100–1100	1	9	10	100–1100	1	10	9	100–1100	1	11.1
3	UBC864	9	200–800	5	55.5	8	200–1500	4	50	8	200–1500	0	0
4	UBC866	6	200–900	0	0	6	200–900	0	0	6	200–900	0	0
5	UBC818	5	500–1600	0	0	6	500–1000	0	0	1	500	0	0
6	UBC868	2	100–200	1	50	7	200–1500	5	71.4	3	200–1500	1	33.3
7	UBC876	9	100–500	5	55.5	7	100–500	4	57.1	10	200–900	5	50
8	UBC802	3	450–1100	2	66.6	3	300–1500	1	33.3	5	220–900	3	60
9	UBC878	6	300–1300	6	100	7	490–1100	6	85.7	7	290–1900	6	85.7
10	UBC805	3	300–700	1	33.3	7	200–1400	5	71.4	5	200–950	3	60

ML, FL, and BS were amplified by Pri12 and Pri9 while, 2 polymorphic bands were amplified in ML by Pri11, 1 polymorphic band in FL by Pri5, 9 and 11 and 2 polymorphic bands in BS by Pri9. The percentage of polymorphism revealed by SSRs ranged from 0% (Pri12) - 66.6% (Pri11) in ML, 0% (Pri12) – 50%(Pri11) in FL and 0% (Pri12)- 50%(Pri5) in BS (Table 3). The genetic similarity based on Jaccard coefficients among 39 accessions of *G.indica* was used to construct a dendrogram (Fig. 2) by the UPGMA method. The genetic distance among the accession ranged from 0.170 to 0.452. The dendrogram based on SSR resolved 33 accessions into four clusters. Cluster comprised 8 ML and 5 FL, cluster II comprised 6 ML and cluster III comprised 7 FL and 1 BS while cluster IV comprised of 5 BS and 1 FL. Thus, SSRs were not found to resolve, cluster and discriminate completely the tested accessions into gender-specific groups.

3.3. Efficacy of ISSRs for gender discrimination

After the initial screening of 80 ISSR primers, 50 failed to amplify, while 20 did not amplify clear and reproducible banding pattern and only 10 primers were found to amplify clear and reproducible ISSR profiles of the tested accessions. All 10 primers produced a highly informative polymorphic banding pattern (Table 4). It was observed that ISSR profiles were greatly dependent on annealing temperatures and also affected the efficiency of amplification. The difference between the optimum annealing temperatures and the theoretical melting temperatures (Tm) ranged from 18 °C to 4 °C. These 10 ISSR primers produced a total of 182 reproducible and scorable bands. Of these, 70 (38.4%) were polymorphic with an average of 7 polymorphic bands per primer. The amplified ISSRs ranged in size from 100–1800 bps. In general, a high number of band per primer was detected, although a low number of polymorphic bands was found. Total 11 bands from ML were amplified by UBC 859, 10 bands in FL by UBC 859 and 10 bands in BS by UBC 876, whereas, 6 polymorphic bandseachin ML, FL, and BSwere amplified by UBC-878. Using ISSR, the percentage of polymorphism ranged from 0% (UBC 866, 818) – 100% (UBC 878) in ML, 0% (UBC 803, 866, 818) – 85.7% (UBC 878) in FL and 0% (UBC 803, 866, 818, 864) 85.7% (UBC 878) in BS (Table 4). The genetic similarity based on Jaccard coefficients among 39 accessions was used to construct a dendrogram (Fig. 3) by the UPGMA method. The genetic distance among the accession ranged from 0.275 to 0.530. The UPGMA dendrogram based on ISSR grouped 39 accessions into three gender-specific clusters (Fig. 3). The major cluster I comprised of 14 ML, cluster II comprised 13 FL whereas cluster III

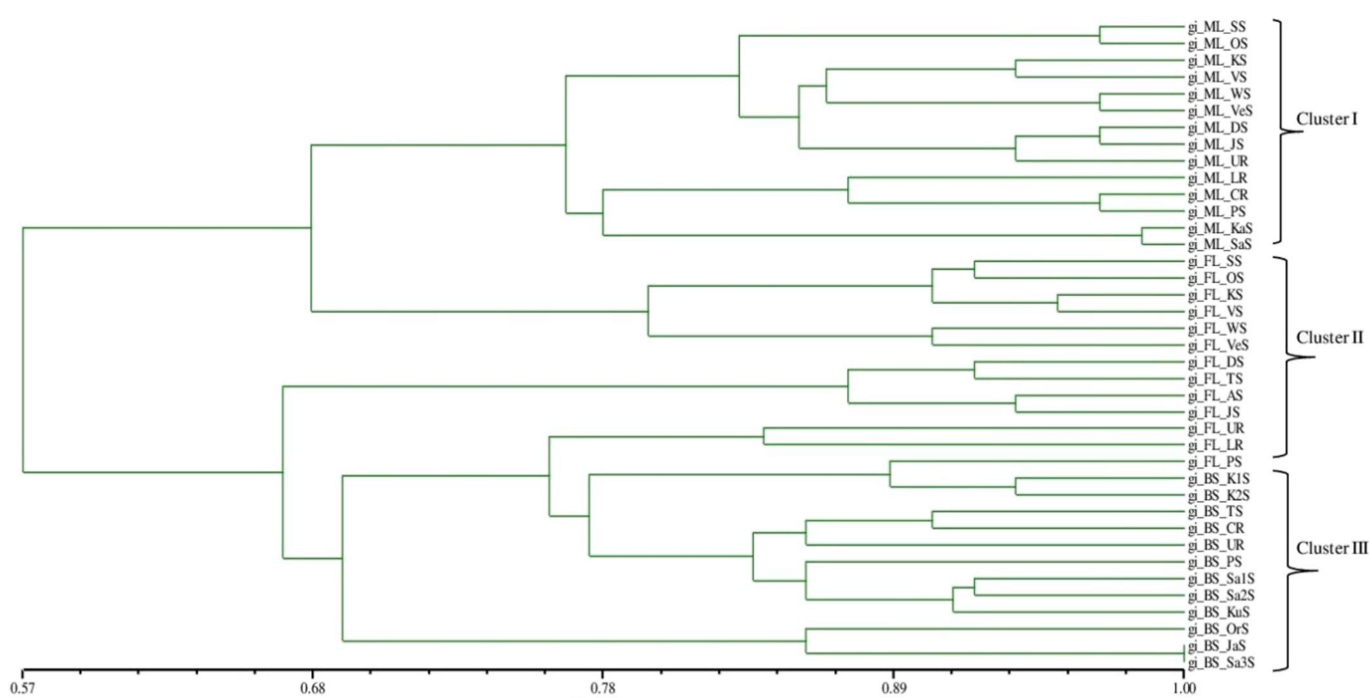

Fig. 3. UPGMA dendrogram based on ISSR profiles generated from *G. indica* accessions. The codes in the figure correspond to the codes listed in Table 1.

Table 5
Characteristics of amplification products obtained from RAPD primers.

No	Primer Name	ML				FL				BS			
		No of bands	Band Size (bp)	Poly morphic band	% polymorphism	No of bands	Band Size (bp)	Poly morphic band	% polymor phism	No of bands	Band Size (bp)	Polymorphic band	% polymorphism
1	OPB6	8	400–1300	2	25	8	400–1300	1	12.5	7	400–1300	5	71.4
2	OPB11	6	200–900	2	33.3	11	200–1900	9	81.8	9	200–1900	6	66.6
3	OPA8	7	200–1250	2	28.5	8	200–1250	3	37.5	8	200–1250	3	37.5
4	OPB15	7	200–1100	6	85.7	8	400–1100	4	50	6	500–1100	0	0
5	OPC1	12	200–2000	10	83.3	9	250–1900	6	66.6	6	250–1900	2	33.3
6	OPB8	9	300–1200	3	33.3	7	300–1200	1	14.2	7	300–1200	4	57.1
7	OPC14	10	400–1200	7	70	7	400–1200	4	57.1	6	500–1200	5	83.3

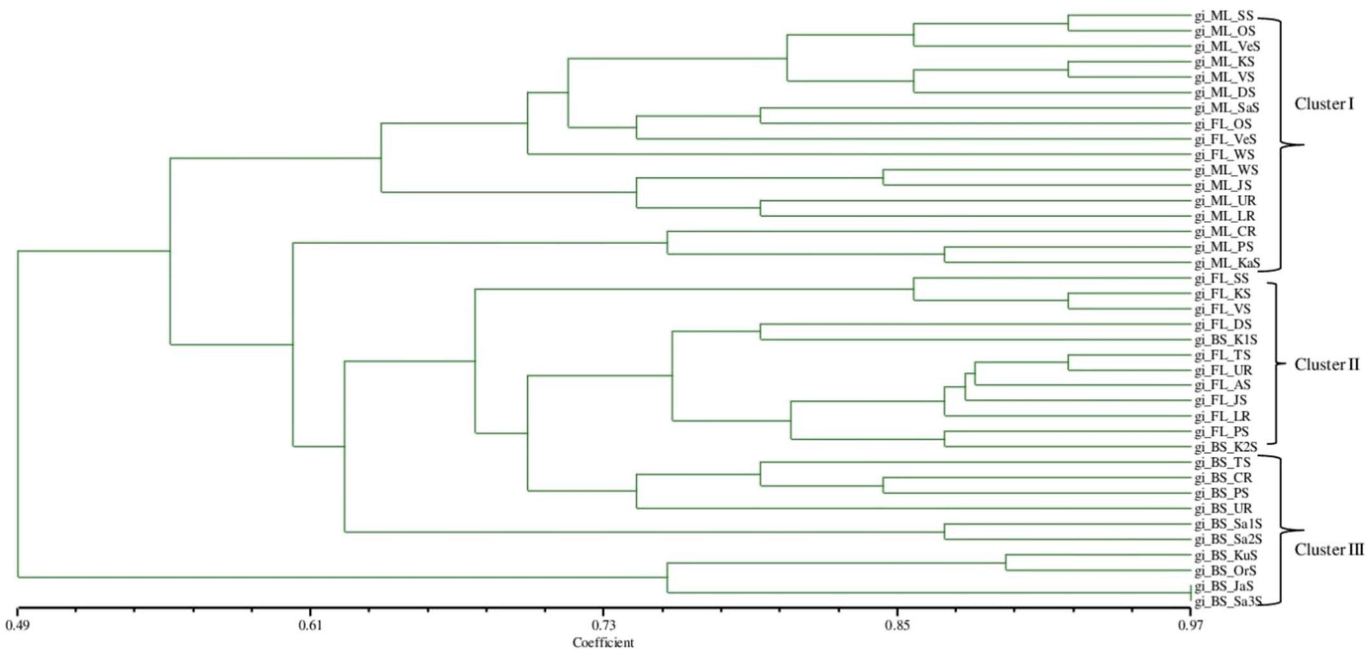

Fig. 4. UPGMA dendrogram based on RAPD profiles generated from *G. indica* accessions. The codes in the figure correspond to the codes listed in Table 1.

comprised of 12 BS accessions. This pattern of clustering based ISSR signifies that ISSR could potentially be used for resolving, clustering and discriminating the gender of *G.indica*. This efficacy of ISSR could further be extended for identification of gender of *G.indica*.

3.4. Efficacy of RAPD for gender discrimination

Initially, the preliminary screening of 60 RAPD primers was done to obtain reproducible, polymorphic and distinguishable banding patterns from ML, FL, and BS accession. Of these 60 primers, 43 primers did not amplify and 10 primers did not amplify clear and reproducible while 7 primers amplified clear, reproducible and highly informative polymorphic banding patterns (Table 5). These seven RAPD primers produced a total of 166 reproducible and scorable bands of which 85 (51.2%) were polymorphic with an average of 4 polymorphic bands per primer. The size range of the amplified fragments ranged from 100–1700 bps. Total,12 bands in ML were amplified by OPC1, 11 bands in FL and 9 bands in BS by OPB11 were amplified. Similarly, 10 polymorphic bands in ML by OPC1, 9 bands in FL and 6 bands in BS by OPB11 were produced. Using RAPD, the percentage of polymorphism ranged from 25% (OPB6) - 85.7% (OPB15) in ML, 12.5% (OPB6) – 81.8% (OPB11) in FL and 0% (OPB15) – 83.3% (OPC14) in BS (Table 5). The genetic distance among the accession ranged from 0.182 to 0.266. The UPGMA dendrogram based on the RAPD profile resolved tested 39 accessions into three clusters (Fig. 4). The major cluster I comprised of 14 ML and 3 of FL; the cluster II comprised of 10 FL and 1 BS whereas cluster III comprised all 11BS. Thus, in comparison, RAPD was better suited than SSR, but less suited than ISSR to resolve, cluster and discriminate the tested accessions into gender-specific groups.

3.5. Efficacy of combined SSRs, ISSR and RAPD data for gender discrimination

It is known that the genetic diversity is better studied if several DNA based molecular markers are employed and analysis is performed based on the combined data of all molecular markers used. The species diversity data analysis based on combined, RAPD and ISSR molecular markers was studied by Gholave et al. (2017). Similarly, Singh et al. (2014) studied an effective resolving polymorphism in *Vigna* species by RAPD, SSR and combined data analysis. In light of this possible implication and advantage of combined data analysis, we also combined SSR, ISSR and RAPD data for the construction of UPGMA dendrogram. The UPGMA dendrogram based on combined data analysis is given in Fig. 5. The combined data resolved 39 accessions into 3 clusters; cluster I of 14 ML accessions, cluster II of 13 FL accessions while cluster III of 12 BS accessions. The clustering is similar to ISSR but still the resolution between FL and BS was incomplete. However, the genetic distance among the accessions ranged from 0.336 to 0.73. Though, recent few studies have indicated the better performance of combined data analysis (Gholave et al., 2017; Patil et al., 2016), the pattern of grouping based

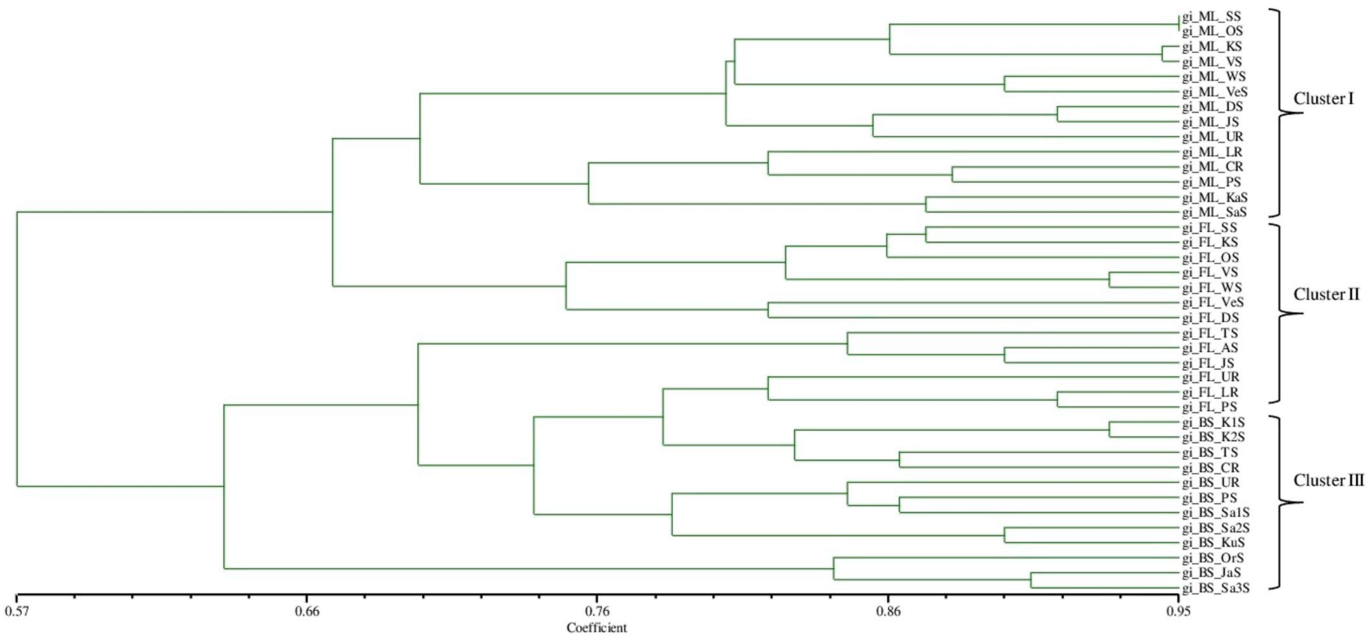

Fig. 5. UPGMA dendrogram based on combined data of SSR, ISSR, RAPD profiles generated from *G. indica* accessions. The codes in the figure correspond to the codes listed in Table 1.

Table 6
Analysis of molecular variance (**AMOVA**) for the accessions of *G.indica* based on RAPD, ISSR and SSR.

Marker	Within gender					Among gender					PhiPT
	DF	SS	MS	variation	% of total variation	DF	SS	MS	variation	% of total variation	
SSR	2	13.18	6.59	0.51	30	30	36.87	1.22	1.22	70	0.295
ISSR	2	145.7	72.86	5.03	40	36	273.2	7.58	7.58	60	0.399
RAPD	2	63.5	31.75	s1.94	23	36	235.35	6.53	6.53	77	0.299
Combine	2	234.05	117.02	7.83	34	36	555.17	15.42	15.42	66	0.460

DF degree of freedom (n - 1), SS sum of squares, MS mean of square, PhiPT = VAP/(VAP ? VWP) (where, VAP variance among the species and VWP variance within the species).

on combined data obtained in the present study demonstrated that use of a multi-marker system may not be much promising for discrimination and identification of gender of *G. indica.* Among all the tested markers and their combined data, ISSR alone was found best suited for gender discrimination and identification in *G. indica.*

3.6. Analysis of molecular variance and PCA

The individual and combined data matrix generated by SSR, ISSR and RAPD were used to generate nonparametric AMOVA (Table 6). The variation percentage among the gender was 70%, 60%, 77% and 66% while within the gender, it was 30%, 40%, 23% and 34% based on SSR, ISSR, RAPD, and combined data, respectively. Our results demonstrate that the inter-gender variation is higher than intra gender, however, the resolution based on ISSR was better than RAPD, genic SSR alone and combined data. Further, PCA analysis based on ISSR explained that more than 50% of the total variation was among the gender. In addition, the two-dimensional plot of PCA clusterstested accession into gender-specific groups which closely resembled to the clustering obtained in UPGMA clustering. It was observed all accessions of respective genders grouped together to form respective gender-specific groups and that ML, FL and BS specific groups were well separately from each other (Fig. 6). In comparison, variance among gender was more than within gender in all the marker systems (Table 6). This PCA based analysis further signifies and demonstrates that in comparison to RAPD and SSR; the ISSR is a best-suited marker system that can be employed for discrimination and identification of gender of *G. indica*.

4. Discussion

It is known that 6% of angiosperms are known to be dioecious which produce separate female and male floral organs, i.e. pistils and stamens (Renner and Ricklefs, 1995). In such plants, the development of molecular strategies for the early identification of sex has been a priority in breeding programs for increasing their economic potential and a better understanding of the developmental as well as evolutionary pathways of dimorphism. The economic utilization of polygamodioecious *G.indica* is greatly hampered by the unavailability of information on the differentiation of ML, FL and BS plants and difficulty in identification of the gender of the plants in the non-flowering and juvenile stage. The slow-growing and cross pollinated nature of *G.indica* further make it difficult to differentiate and identify the gender of the plant. Based on morphometric characteristics, a lot of ambiguities arise for proper identification of *G. indica*. Although few research groups have studied molecular markers in *G. indica,* none could completely link marker with a specific gender. RAPD based diversity study in *G. mangostana* conducted by Sando et al. (2004) reported 34% of polymorphism among the accessions. Similarly, the study by Sahasrabudhe and Deodhar (2010) showed a very low level of polymorphism in *G.indica.* Likewise, Thatte et al. (2011) also conducted RAPD and ISSR analysis in *G.indica* which revealed low genetic diversity in different populations of *G.indica.* In the present study, we collected 39 accessions with confirmed gender identity from the Western Ghats of Maharashtra and used for PCR amplification of DNA based molecular markers. Further, the performance of these markers to resolve the collected accession into gender-specific

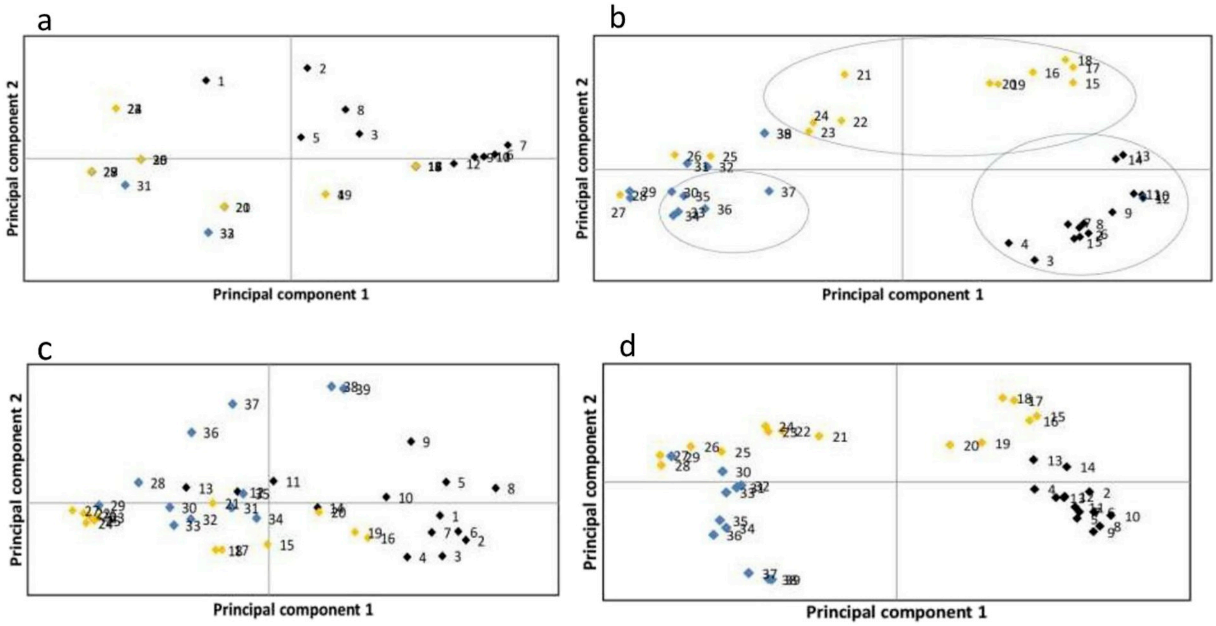

Fig. 6. Principal component analysis showing the plot of the first two principal components for the three marker systems used in the study. a: PCA based on SSR, b: PCA based on ISSR, c: PCA based on RAPD, d: PCA based on combined data.
Color represents ◆ ML ◆ FL ◆ BS (Table 6), The accession numbers correspond to Table 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

clusters was also tested. To this end, DNA-based markers such as genic SSR, ISSR, and RAPD which provide several advantages over traditional morpho-biochemical markers were used in the present study. Moreover, DNA markers provide additional benefits such as low cost per assay, are not affected by the environmental factors, require a lower level of skill, and readily available primers allow the scanning of the entire genome and efficient genotype characterization. In addition, these markers show characteristics higher efficiency for intra or inter-specific polymorphism detection. Because of these desired characteristics of RAPD and ISSR, they were used to calculate the inter or intra-specific genetic diversity in different domestic and wild species like *Jacaranda mimosifolia* and rice bean (Escandón et al., 2005; Muthusamy et al., 2008).

Genome-wide analysis of SSRs can provide abundant markers for genetic, genomic, and evolutionary studies. The distribution and frequency of EST-SSRs are reported to be highly variable, depending upon their search criteria, the size of a dataset, and the database mining tools (Varshney et al., 2005). In the present investigation, a total of 49028, 45944 and 49414 unique sequences were used for SSR search in ML, FL and BS plants which identified 8213(16.7%), 6380(13.8%) and 10996 (22.2%) SSRs respectively. This is relatively high abundance of SSRs in comparison to the earlier reports from maize (1.4%), barley (3.4%), wheat (3.6%), sorghum (3.6%), and rice (4.7%) (Kantety et al., 2002). In *G. indica,* the highest proportion of repeat motif comprised dinucleotides followed by trinucleotides. The earlier findings in wheat, rice, corn, soybean (Cardle et al., 2000), cotton (Han et al., 2006), sugarcane (Cordeiro et al., 2001), barley (Thiel et al., 2003), peanut (Liang et al., 2009) and citrus (Chen et al., 2006) reported trinucleotide repeat motifs as the most common which is not in agreement with present findings in *G.indica*. On the contrary, in agreement with our findings, the dinucleotide repeat motifs were reported as common in Picea (Rungis et al., 2004), apricot and peach (Jung et al., 2005), coffee (Aggarwal et al., 2007), lotus (Pan et al., 2010), rubber (Feng et al., 2009), cassava (Raji et al., 2009) and Jatropha (Kumar Yadav et al., 2011). The SSR search criteria and data analysis can be associated with the relative difference of trinucleotide and dinucleotide. Among the dinucleotides, the AG/TC in ML and BS while AG/GA in FL were most common representing a frequency of 21.50%, 38.09%, and 21.03% respectively. In accordance to rare presence of GC repeats in plants, very low frequency (<1%) of CG/GC was found in *G.indica* which is in congruence with similar observation reported in coffee (Aggarwal et al., 2007) and Jatropha (Kumar Yadav et al., 2011). No GC repeat motifs were reported in rice, corn, soybean (Gao et al., 2003) and wheat (Nicot et al., 2004). Among trinucleotide motifs, GGA/CTC/GAA occurred most commonly in ML, FL and BS while CCG, CGC, CGG was rare in respective samples. However, otherwise distribution was observed in the monocots rice and sorghum, in which CCG was the most abundant trinucleotide motif (Kumar Yadav et al., 2011). Cavagnaro et al. (2010) suggested that these differences between dicots and monocots are related to differences in the average GC content of their genomes (34.6% for dicots and 43.7% for monocots). In addition, GC bias in next-generation sequencing may also play a role in detecting the frequency of CG and CCG repeat motifs, since the potential effect of GC bias in sequencing is the detection of fewer GC reads (Chen et al., 2013).

The outcome of our present study successfully demonstrates the utility of genic SSR, ISSR and RAPD markers to assess genetic variation among the *G.indica* accessions, and therefore, their applications can be extended for discrimination and identification of gender of *G. indica.* In all tested accessions, the marker system yielded more than 50% polymorphism, however, in terms of variation within accessions, ISSR was found better suited than RAPD and genic SSR. Of 78 ISSR primers used to amplify the DNA from ML, FL and BS individuals, 10 primers amplified reproducible ISSR pattern. However, the number of amplification products varied from 1 to 11 and the fragment sizes ranged from 100 to 1800 bp. Of all 78 primers, only one primer UBC 818 was found to show sex specificity in bulk analysis. Primer UBC 818 produced a unique ~1600 bp fragment in ML which was absent in FL and BS whereas, UBC 818 also produced a unique ~300 bp fragment in FL, that was absent in ML and BS. Similarly, primer UBC 864 produced a unique ~1000 bp fragment in BS which was absent in ML and FL. For confirming this observation, this primer was retested with individual ML, FL and BS samples from different accessions. Interestingly, the unique ~1600 bp fragment was amplified from only ML individuals from all 14 accessions (Fig. 7) and was completely absent in the profiles generated from respective FL and BS individuals tested. Similarly 300 bp fragment was

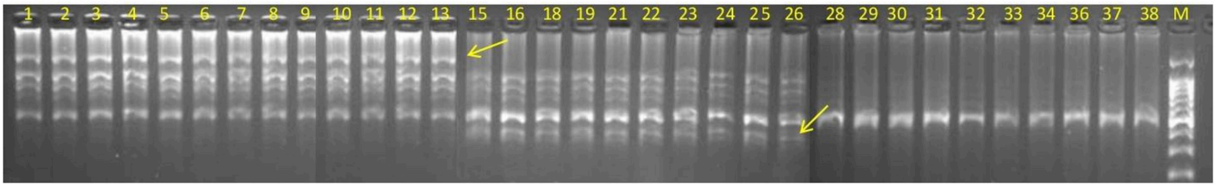

Fig. 7. ISSR profile generated from the accessions of *G.indica* amplified using primer UBC-818 in ISSR. Lane 1–13: ML, lane 14–23: FL, lane 24–33: BS, lane 34: 100bp ladder.

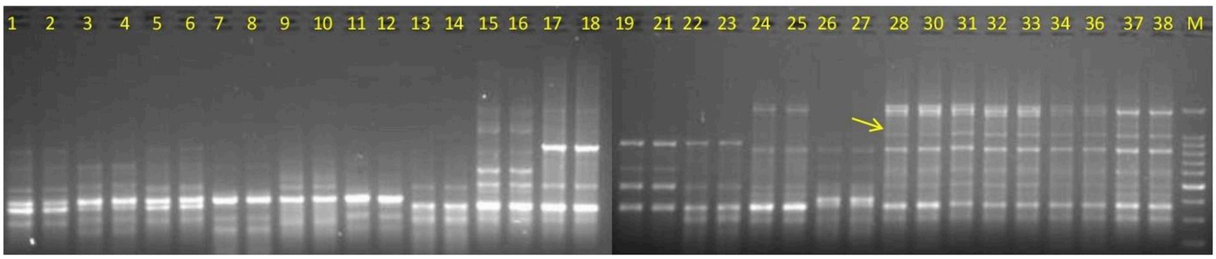

Fig. 8. ISSR profile generated from the accessions of *G. indica* amplified using primer UBC-864. Lane 1–14: ML, lane 15–26: FL, lane 27–35: BS, lane 36: 100bp ladder.

amplified from 10 FL accessions (Fig. 7) and was completely absent in ML and BS while ~1000 bp fragment was amplified from 9 BS accessions (Fig. 8). Thus, here we proposing that ISSR markers amplified by UBC-818_{1600}, UBC- 818_{300} and UBC- 864_{1000} could be recognized as a putative sex-linked marker and used for discrimination and identification of the gender of the *G. indica*. To the best of our knowledge, the present study is the first report that has evaluated the performances of DNA based marker for discrimination and identification in *G. indica* (Figs. 7, 8).

The first two principal components in PCA based on ISSR profile of 39 accessions explained 77% variation within accessions and grouped accessions into the gender-specific groups. On the contrary, SSR analysis could not completely resolve the accessions into gender-specific groups. This grouping pattern based on ISSR clearly indicated that ISSR can be used for resolution of *G.indica* accession into gender-specific groups and, therefore, be utilized for identification of the gender of the plants. Similarly, the performance of RAPD was better that SSR but not as good as ISSR. In addition, the groups formed based on combined data analysis of all tested markers were similar to that based on ISSR profile. In comparison, the resolution based on ISSR and combined data was better than RAPD and genic SSR alone. Overall,the efficiencies of the tested markers for gender discrimination and identification in order of merit was observed as ISSR > RAPD > SSR. Also, the results of PCA analysis largely matched to those obtained through UPGMA based cluster analysis.

The present study significantly differs from the previous attempts by Sawardekar et al. (2011), and Thatte and Deodhar (2012) to identify genders of *G. indica* using molecular markers, in that Sawardekar et al. used 9 RAPD primers for ML, FL and hermaphrodite plants which yielded characteristic band only in ML while Thatte and Deodhar screened 92 RAPD and 28 ISSR primersin ML and FL plants which showed 2 RAPD and 1 ISSR marker specific for FL and ML respectively (Thatte and Deodhar, 2012). On the contrary present study conducted and employed a large number of accessions and employed multiple markers such as RAPD, ISSR, and SSR.

5. Conclusion

This study described the identification of microsatellites mined from transcriptome sequence data of *G.indica* and evaluated them along with ISSR and RAPD for their ability to discriminate and identify the gender of the *G. indica*. In conclusion, it was found that total dinucleotide repeats were the most frequently observed SSRs within the transcriptome, with 'AG/TC' being the most common motif. It provided valuable information about the abundant genic SSRs which can be used for further genetic studies of *G.indica*.The success of our study in identifying the polymorphism was due to the use of randomly selected pre-screened highly informative markers. The assessment of performances of tested DNA based markers based on UPGMA dendrogram and PCA demonstrated that the performances for gender discrimination were in the order of merit as ISSR > RAPD > SSR. The proposed ISSR markers amplified by selected ISSR primers UBC-818 and UBC-864 marker could be used to discriminate between ML, FL, and BS plants in *G. indica*. These markers may facilitate genetic and genomic studies, leading to improvement in the selection and commercial planting of Kokum.

Conflicts of interest

The authors declare that there is no conflict of interests regarding the publication of this article.

Acknowledgments

This work was supported by DST-SERB project sanctioned to Dr. K. D. Pawar under Young Scientist Scheme (YSS/2015/000319).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bcab.2019.101319.

References

Aggarwal, R.K., Hendre, P.S., Varshney, R.K., Bhat, P.R., Krishnakumar, V., Singh, L., 2007. Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. Theor. Appl. Genet. 114, 359–372.

Cardle, L., Ramsay, L., Milbourne, D., Macaulay, M., Marshall, D., Waugh, R., 2000. Computational and experimental characterization of physically clustered simple sequence repeats in plants. Genetics 156, 847–854.

Cavagnaro, P.F., Senalik, D.A., Yang, L., Simon, P.W., Harkins, T.T., Kodira, C.D., Huang, S., Weng, Y., 2010. Genome-wide characterization of simple sequence repeats in cucumber (Cucumissativus L.). BMC Genomics 11, 569.

Chaves-Bedoya, G., Nuñez, V., 2007. A SCAR marker for the sex types determination in Colombian genotypes of Carica papaya. Euphytica 153, 215–220.

Chen, C., Zhou, P., Choi, Y.A., Huang, S., Gmitter, F.G., 2006. Mining and characterizing microsatellites from citrus ESTs. Theor. Appl. Genet. 112, 1248–1257.
Chen, Y.C., Liu, T., Yu, C.H., Chiang, T.Y., Hwang, C.C., 2013. Effects of GC bias in next-generation-sequencing data on de novo genome assembly. PLoS One 8.
Cordeiro, G.M., Casu, R., McIntyre, C.L., Manners, J.M., Henry, R.J., 2001. Microsatellite markers from sugarcane (Saccharum spp.) ESTs cross transferable to erianthus and sorghum. Plant Sci. 160, 1115–1123.
Danilova, T.V., Karlov, G.I., 2006. Application of inter simple sequence repeat (ISSR) polymorphism for detection of sex-specific molecular markers in hop (*Humulus lupulus* L.). Euphytica 151, 15–21.
Escandón, A.S., Pérez De La Torre, M., Acevedo, A., Marcucci-Poltri, S., Miyajima, I., 2005. Anchored ISSR as molecular marker to characterize accessions of jacaranda mimosifolia L. Don. Acta Hortic. (Wagening.) 683.
Excoffier, L., Smouse, P.E., Quattro, J.M., 1992. Analysis of Molecular Variance Inferred from Metric Distance, pp. 479–491.
Feng, S.P., Li, W.G., Huang, H.S., Wang, J.Y., Wu, Y.T., 2009. Development, characterization and cross-species/genera transferability of EST SSR markers for rubber tree (*Hevea brasiliensis*). Mol. Breed. 23, 5–97.
Gangopadhyay, G., Roy, S.K., Ghose, K., Poddar, R., Bandyopadhyay, T., Basu, D., Mukherjee, K.K., 2007. Sex detection of *Carica papaya* and *Cyca scircinalis* in pre-flowering stage by ISSR and RAPD. Curr. Sci. 92, 524–526.
Gao, L., Tang, J., Li, H., Jia, J., 2003. Analysis of microsatellites in major crops assessed by computational and experimental approaches. Mol. Breed. 12, 245–261.
Gholave, A.R., Pawar, K.D., Yadav, S.R., Bapat, V.A., Jadhav, J.P., 2017. Reconstruction of molecular phylogeny of closely related Amorphophallus species of India using plastid DNA marker and fingerprinting approaches. Physiol. Mol. Biol. Plants 23, 155–167.
Haldankar, P.M., Salvi, M.J., Joshi, G.D., Patil, J.L., 1993. Factors influencing softwood grafting in kokam (*Garcinia indica* Choisy). Indian Cocoa Arecanut Spices J 17, 15–18.
Han, Z., Wang, C., Song, X., Guo, W., Gou, J., Li, C., Chen, X., Zhang, T., 2006. Characteristics, development and mapping of *Gossypium hirsutum* derived EST-SSRs in allotetraploid cotton. Theor. Appl. Genet. 112, 430–439.
Harkess, A., Leebens-Mack, J., 2017. A century of sex determination in flowering plant. J. Hered. 69–77.
Heymsfield, S.B., Allison, D.B., Vasselli, J.R., Pietrobelli, A., Greenfield, D., Nunez, C., 2014. *Garcinia cambogia* (hydroxycitric acid) as a potential antiobesity agent. Jama 280, 1596.
Jung, S., Abbott, A., Jesudurai, C., Tomkins, J., Main, D., 2005. Frequency, type, distribution and annotation of simple sequence repeats in Rosaceae ESTs. Funct. Integr. Genom. 5, 136–143.
Kantety, R.V., Rota, M.L., Matthews, D.E., Sorrells, M.E., 2002. Data mining for simple sequence repeats in expressed sequence tags from barely, maize, rice, sorghum and wheat. Plant Mol. Biol. 48, 501–510.
Khanuja, S.P.S., Shasany, A.K., Darokar, M.P., Kumar, S., 1999. Rapid isolation of DNA from dry and fresh samples of plants producing large amounts of secondary metabolites and essential oils. Plant Mol. Bio.l Rep. 17, 74-74.
Kumar Yadav, H., Ranjan, A., Asif, M.H., Mantri, S., Sawant, S.V., Tuli, R., 2011. EST-derived SSR markers in Jatrophacurcas L.: development, characterization, polymorphism, and transferability across the species/genera. Tree Genet. Genomes 7, 207–219.
Liang, X., Chen, X., Hong, Y., Liu, H., Zhou, G., Li, S., Guo, B., 2009. Utility of EST-derived SSR in cultivated peanut (*Arachis hypogaea* L.) and Arachis wild species. BMC Plant Biol. 9, 35.
Magdalita, P.M., Mercado, C.P., 2003. Determining the sex of papaya for improved production. Bull. Food Feri. Tech. Centre, Philippines 1–6.
Malik, S.K., Chaudhury, R., Abraham, Z., Culture, T., Unit, C., Campus, P., Station, R., 2005. Desiccation - freezing sensitivity and longevity in seeds of. Culture 723–732.
Muthusamy, S., Kanagarajan, S., Ponnusamy, S., 2008. Efficiency of RAPD and ISSR markers system in accessing genetic variation of rice bean (*Vigna umbellata*) landraces. Electron. J. Biotechnol. 11.
Nawala, R.N., Karmarkar, M.S., 1997. Softwood grafting technique for rapid multiplication of high yielding clones in kokum (*Garcinia indica* Choisy). Indian Cocoa Arecanut Spices J 21, 76–77.
Nei, M., Li, W.H., 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. U.S.A. 76, 5269–5273.
Nicot, N., Chiquet, V., Gandon, B., Amilhat, L., Legeai, F., Leroy, P., Bernard, M., Sourdille, P., 2004. Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (ESTs). Theor. Appl. Genet. 109, 800–805.
Pan, L., Xia, Q., Quan, Z., Liu, H., Ke, W., Ding, Y., 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, 71–82.
Parasnis, A.S., Ramakrishna, W., Chowdari, K.V., Gupta, V.S., Ranjekar, P.K., 1999. Microsatellite (GATA)(n) reveals sex-specific differences in papaya. Theor. Appl. Genet. 99, 1047–1052.
Patil, R.R., Pawar, K.D., Rane, M.R., Yadav, S.R., Bapat, V.A., Jadhav, J.P., 2016. Assessment of genetic diversity in *Mucuna* species of India using randomly amplified polymorphic DNA and inter simple sequence repeat markers. Physiol. Mol. Biol. Plants 22, 207–217.
Peakall, R., Smouse, P.E., 2006. Genalex 6: genetic analysis in excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6, 288–295.
Powell, W., et al., 1996. Polymorphism revealed by simple sequence repeats. Trends Plant Sci. 1, 215–222.
Prakash, S., Van Staden, J., 2006. Sex identification in encephalartosnatalensis (dyer and Verdoorn) using RAPD markers. Euphytica 152, 197–200.
Raji, A.A.J., Anderson, J.V., Kolade, O.A., Ugwu, C.D., Dixon, A.G.O., Ingelbrecht, I.L., 2009. Gene-based microsatellites for cassava (*Manihot esculenta* Crantz): prevalence, polymorphisms, and cross-taxa utility. BMC Plant Biol. 9, 118. https://doi.org/10.1186/1471-2229-9-118.
Renner, S.S., Ricklefs, R.E., 1995. Dioecy and its correlates in the flowering plants. Am. J. Bot. 82, 596–606.
Rungis, D., Bérubé, Y., Zhang, J., Ralph, S., Ritland, C.E., Ellis, B.E., Douglas, C., Bohlmann, J., Ritland, K., 2004. Robust simple sequence repeat markers for spruce (Picea spp.) from expressed sequence tags. Theor. Appl. Genet. 109, 1283–1294.
Sahasrabudhe, A., Deodhar, M., 2010. Standardization of DNA extraction and optimization of RAPD-PCR conditions in *Garcinia indica*. Int. J. Bot. 6, 293–298.
Sando, L.C., Ramage, B.C., Drew, R., 2004. Assessment of genetic diversity in Australian grown mangosteen (*Garcinia mangostana*) and its wild relatives. Euphytica 136, 1–2.
Sawardekar, S.V., Gokhale, N.B., Mote, M.S., Joshi, S.N., Sawant, S.S., 2011. Sex detection of kokum (*Garcinia indica* choisy) by RAPD markers. Indian J. Genet. Plant Breed. 71, 82–83.
Singh, A., Dikshit, H.K., Jain, N., Singh, D., Yadav, R.N., 2014. Efficiency of SSR, ISSR and RAPD markers in molecular characterization of mungbean and other Vigna species. Indian J. Biotechnol. 13, 81–88.
Smith, S., Helentjaris, T., 1996. In: Paterson, A.H. (Ed.), Genome Mapping in Plants. Academic Press, RG Landes Company, Texas, p. 95.
Tan, M., Xue, J., Wang, L., Huang, J., Fu, C., Yan, X., 2016. Transcriptomic Analysis for Different Sex Types of *Ricinus communis* L .during Development from Apical Buds to Inflorescences by Digital Gene Expression Profiling 6.
Thatte, K.S., Deodhar, M.A., 2012. Study of flowering behavior and sex determination in *Garcinia indica* (Thomas-Du Pettite) Choisy by means of molecular markers. Biotechnology 11, 232–237.
Thatte, K.S., Khandekar, R.G., Deodhar, M.A., 2011. Assessment of diversity in *Garcinia indica* (Dupetit-Thouars.) Choisy.using morphological and molecular markers. J. Trop. Agric. 50, 30–36.
Thiel, T., Michalek, W., Varshney, R.K., 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.
Urasaki, N., Tokumoto, M., Tarora, K., Ban, Y., Kayano, T., Tanaka, H., Oku, H., Chinen, I., Terauchi, R., 2002. A male and hermaphrodite specific RAPD marker for papaya (*Carica papaya* L.). Theor. Appl. Genet. 104, 281–285.
Varshney, R.K., Graner, A., Sorrells, M.E., 2005. Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 23, 48–55.
Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A., Tingey, S.V., 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18, 6531–6535.
Xu, W.J., Wang, B.W., Cui, K.M., 2004. RAPD and SCAR markers linked to sex determination in *Eucommia ulmoides Oliv*. Euphytica 136, 233–238.
Zietkiezicz, E., Rafalski, A., Labuda, D., 1994. Genome fingerprinting by simple sequence repeat (SSR) - anchored PCR ampification. Genomics 20, 176–183.