Physiol Mol Biol Plants (October–December 2015) 21(4):519–529
DOI 10.1007/s12298-015-0322-2

RESEARCH ARTICLE

Morphological and genetic diversity assessment of sesame (*Sesamum indicum* L.) accessions differing in origin

Sarita K. Pandey[1,2] · **Arna Das**[1] · **Pooja Rai**[1] · **Tapash Dasgupta**[1]

Received: 5 June 2015 / Revised: 13 September 2015 / Accepted: 21 September 2015 / Published online: 6 October 2015

Abstract Sesame is an important ancient oilseed crop of high medicinal value. In the present study, 37 characters including both quantitative and qualitative traits of sixty genotypes were characterized following IPGRI morphological descriptors for sesame. Multivariate analysis was computed to distinguish the varieties into different groups. Though thirty six microsatellite markers including genomic and Est-SSR markers were initially selected, but, finally, the accessions were genotyped by eight polymorphic primers. Altogether, 27 alleles were detected among the 60 genotypes, with an average of 3.37 alleles per locus. The number of alleles ranged from 2 to 6 alleles. From data of microsatellite markers, dissimilarity coefficients between varieties were computed following Jaccard's coefficient method. Principal co-ordinate analysis was used to represent the varieties in bi-directional space. Dendrogram was constructed using NJ method based on dissimilarity matrix. Cluster analysis based on morphological and molecular marker classified sesame genotypes into two major groups. Mantel test showed an insignificant correlation between phenotypic and molecular marker information. The genotypes belonging to the same geographical area did not always occupy the same cluster. The results confirmed that both genetic and phenotypic diversity in a combined way could efficiently evaluate the variation present in different sesame accessions in any breeding program.

✉ Sarita K. Pandey
isarita06@gmail.com

[1] Department of Genetics and Plant Breeding, Institute of Agricultural Science, Calcutta University, Kolkata, India

[2] International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Telangana, India

Keywords Sesame · Genetic diversity · Phenotypic · SSR · EST-SSR · Clustering

Introduction

Sesame (*Sesamum indicum* L.) is a popular oilseed crop in tropical and subtropical countries, including India. Sesame production in India is encouraging, as more than 30 % of world production of sesame is contributed by India alone and India ranks top in total production. Despite being largely self-sufficient in production, sesame productivity is in decline (Anthony et al. 2015; FAOSTAT 2015) and it is of great concern and that is caused mainly due to unavailability of high yielding varieties (Duhoon 2004; Ram et al. 2006). Sesame is considered as a nutritious oilseed crop being rich source of protein (18–25 %), carbohydrate (13.5 %), minerals and healthy polyunsaturated fatty acid (Bedigian et al. 1985). Sesame oil is favored as a media of cooking by Indians and Africans. Presence of sesamol, a unique anti-oxidant and more poly-unsaturated fatty acid, have made it to 'queen of oilseed crop' (Ashri 1989; Fukuda et al. 1986). Sesame contributes in multifarious ways to human being like culinary preparation, cosmetic industry, decorative elements etc. but the farmers often do not get attracted for expansion of cultivation as the crop suffers from low productivity and hence its cultivation leads to low economic return. Presence of genetic diversity is a desirable requirement for any breeding program. Large variability of sesame genotypes in India is not a surprising proposition, as India along with China, central Asia, Near East and Abysinia has been identified as sesame diversity centers in classical studies (Hawkes 1983; Laurentin and Karlovsky 2006; Zeven and Zhukovsky 1975). Presently, the main technique of analyzing plant genetic resources, centers on modern DNA marker technology, since morphological study disguises

the genetic diversity because of environmental effect. The markers like RAPD (Bhat et al. 1999; Ercan et al. 2004; Zhang et al. 2004), AFLP (Ali et al. 2007; Laurentin and Karlovsky 2006, 2007) and ISSR (Kim et al. 2002) have already been reported in sesame. Microsatellites or simple sequence repeat has proved its unequivocal supremacy because of high abundance and reproducibility, easy scoring, extensive coverage and co dominant nature. Despite the great importance of SSR, its application is limited in sesame. Again, SSR can be partitioned into genomic SSR and Est-SSRs. Information about genomic SSR is limited in sesame (Dixit et al. 2005). On the contrary, information on Est-SSR information is now available to a considerable extent (Wei et al. 2008).

Though sesame genotypes in Indian subcontinent have been reported to represent a wide diversity but still no systematic efforts have been made to characterize and document the indigenous and exotic collections along with the local germplasm to reveal the genetic diversity (Kim et al. 2002). The knowledge of genetic variation present between indigenous and exotic collection of sesame will provide critical information for better management strategy for crop improvement in adverse conditions. Diversity analyses based only on morphological characters are prone to environmental bias due to environmental influences and complex genetic structure of different morphological traits (Banerjee and Kole 2009; Tabatabaei et al. 2011). So combination of both morphological and molecular study would be more preferable. The aims of the present study were to elucidate genetic relationships among sixty sesame accessions specifically chosen from different countries representing diverse centers of diversity such as India, USA, Bangladesh and Bulgaria and also to determine association between geographical origin and genetic diversity using both morphological and SSR markers.

Materials and methods

Plant material

Sixty genotypes (Table 1) collected from different parts of the world were grown in randomized block design with three replications at Calcutta University experimental farm, Baruipur (22° 22′ N, 88° 26′E) in 2011, 2012 and 2013 during summer season. Out of these sixty genotypes, thirty cultivars were collected from eight major sesame growing states of India. A few landraces were collected from different parts of West Bengal, India. Altogether, 13 quantitative and 48 qualitative traits of sixty genotypes were characterized on the basis of field data as outlined in IPGRI morphological descriptors for Sesame (IPGRI 2004). Many traits showed single and binary type of character state. For each qualitative traits numeric was assigned as suggested in the descriptor and finally, 37 phenotypic traits which varied extensively were taken for data analysis.

DNA extraction

DNA extraction was mainly carried out from apical young leaves of sesame plants around 10–12 days old. Leaves were grounded in liquid nitrogen and then treated with CTAB buffer as described by Saghai-maroof et al. (1984) and RNase A was used to purify the extracted DNA followed by phenol: chloroform. The purified DNA was quantified with Nanodrop Lite (Thermo Scientific, USA).

SSR primers

Thirty six microsatellite markers including genomic SSR primers reported earlier by Dixit et al. (2005) and Est-SSR were selected for the present study. Additionally, from dbEST/Gene Bank/(http://www.ncbi.nlm.nih), 3328 expressed sequence tag (Est) sequences deposited for sesame were collected. The sequences were downloaded in Fasta format. Est-SSR primers were designed following primer3 (http://fokker.wi.mit.edu/primer3/) from the screened SSRs as described by Bhattacharyya et al. (2014).

PCR amplification

DNA amplification was done in 25 μl reaction mixture containing 0.2 $\mu molL^{-1}$ SSR primers, 0.2 mM of each dNTPs, 2 $mmolL^{-1}$ $MgCl_2$, 1 X PCR buffer and 0.5 unit Taq polymerase (Life Technologies, New York, USA) and 50 ng sample DNA. The program for SSR and Est-SSR was carried out in a DNA thermocycler (model-Pro, Eppendorf AG 6321, Germany). The program was set up as follows: denaturation at 95 °C for 5 min, 35 cycles 1 min at 95 °C, 50s at 50 to 62 °C depending on apropos annealing temperature of the primer, and finally extension at 72 °C for 5 min. Amplified PCR products were separated on a 3 % agarose gel (Sigma USA) stained with ethidium bromide (Biorad, USA). A 50 base pair ladder marker (GeneRuler 50 bp DNA Ladder, Thermo Scientific, USA) was used to estimate PCR fragment size.

Genetic diversity analysis using ssr profiles and bootstrap analysis

Only clear and reproducible bands were selected for data analysis. Presence or absence of bands related to each primer was scored as '1' or '0' respectively. Effective allele per locus (Aep) was computed as suggested by Weir (1990) by 1/(1-Hep), where Hep designates the genetic diversity for each locus and Hep is equal to $1-\sum P^2i$, Pi means the frequency of ^{i}th allele at the locus. All the data were analyzed for multivariate

Table 1 The list of genotypes of sesame used in the present study

S.No.	Cultivar Name	Originating Institute	Year of Release	Pedigree
1.	TKG-22	ZARS, JNKVV,	1995	HT-6 x JLT-3
2.	CUMS-04	West Bengal	–	Rama (400Gy)
3.	V12	West Bengal	–	–
4.	RT-348	Local Rajasthan	–	Local landrace
5.	OSC-593	Bhubaneshwar	–	–
6.	RT-54	ARS, SKRAU	1992	A6-5 x BS 6-1
7.	CUMS-11	West Bengal	–	SI 1666 (200Gy)
8.	SWB-32–10–1	Behrampur	2008	Selection from SWB-32
9.	RAMA	Behrampur	1989	Selection from 'Khosla' local
10.	CUMS-20	West Bengal	–	IC 21706 (0.5 % EMS)
11.	NIRMALA	OUAT, Bhubaneshwar	2003	Mutant of B-67
12.	CUMS-09	West Bengal	–	SI 1666 (200Gy)
13.	Gujarat TIL-2	ARS, GAU, (Gujarat)	1995	Gujarat Til-1× TC-25
14.	TKG-352	Local Tikamgarh (MP)	–	Local landrace
15.	UMA	OUAT, Bhubaneshwar	1999	Mutant of Kanak
16.	V10	West Bengal		Local landrace
17.	SAHEB	West Bengal	–	Local landrace
18.	B-14	West Bengal	–	Local landrace
19.	TMV-6	TNAU, Coimbatore	1978	Selection from local material of AP
20.	TMV-4	TNAU, Coimbatore	1984	Selection from variety sattur
21.	OSC-207	Local Orissa		Local landrace
22.	TILOTTAMA	Behrampur	1984	Selection from Jinardi Ducca-2
23.	V-1	West Bengal	–	Local landrace
24.	V-15	West Bengal	–	Local landrace
25.	B76	West Bengal	–	Local landrace
26.	NIC-8316	Indigenous Collection	–	–
27.	VRI (SV)-1	TNAU,Coimbatore	1997	Pureline selection from thirukattupalli
28.	CST-2001–12	Kanpur, UP	–	–
29.	AMRIT	OUAT, Bhubaneshwar	2007	Selection from XU-2 X Krishna
30.	DSS-9	UAS, Dharwad	–	Mutant of Phule Til-1 x E-8
31.	EC-79/EC-334988-(3)	Bangladesh	–	Exotic collections
32.	EC-76/EC-335004 (25)	Bangladesh	–	Exotic collections
33.	EC-70/EC-334973 (38)	Bangladesh	–	Exotic collections
34.	EC-83/EC-334971 (23)	Bangladesh	–	Exotic collections
35.	EC-84/EC-335004 (34)	Bangladesh	–	Exotic collections
36.	EC-72/EC-334991	Bangladesh	–	Exotic collections
37.	EC-90/EC-310448 (36)	Bulgaria	–	Exotic collections
38.	EC-77/EC-182832 (26)	Bulgaria	–	Exotic collections
39.	EC-112/EC-310421	Bulgaria	–	Exotic collections
40.	EC-91/EC-204704 (44)	Bulgaria	–	Exotic collections
41.	EC-92/EC-334962 (28)	Bulgaria	–	Exotic collections
42.	EC-67/EC-138835 (21)	Bulgaria	–	Exotic collections
43.	EC-87/EC-303435 (4)	USA	–	Exotic collections
44.	EC-106/EC 303433(17)	USA	–	Exotic collections
45.	EC-97/EC-164966 (50)	USA	–	Exotic collections
46.	EC-69/EC-41923-B(49)	USA	–	Exotic collections
47.	EC-107/EC-100043-A	USA	–	Exotic collections
48.	EC-108/EC-303442(32)	USA	–	Exotic collections
49.	IC-41/IC-152485	India	–	Indigenous collection

Table 1 (continued)

S.No.	Cultivar Name	Originating Institute	Year of Release	Pedigree
50.	IC-42/IC-204159	India	–	Indigenous collection
51.	IC-43/IC-20477	India	–	Indigenous collection
52.	IC-49/IC-17477	India	–	Indigenous collection
53.	IC-51/IC-2621694	India	–	Indigenous collection
54.	IC-54/IC-96230	India	–	Indigenous collection
55.	IC-56/IC-43033	India	–	Indigenous collection
56.	IC59/IC-204063	India	–	Indigenous collection
57.	IC-60/IC-131490	India	–	Indigenous collection
58.	IC-62/IC-14331	India	–	Indigenous collection
59.	IC-63/IC-141464	India	–	Indigenous collection
60.	IC-64/IC-14053	India	–	Indigenous collection

analysis using the software NTSYS Pc Ver. 2.20 (Sneath and sokol 1973). The data were first standardized to eliminate the effects of different measurement following STAND option. The distance coefficient utilizing DICE similarity index was then worked out by utilizing the transformed data and the information was epitomized in dendrogram following unweighted pair group method with arithmetic average (UPGMA), and Shan clustering program in NTSYS Pc. 2.20.

The morphological data of all the characters were put in multivariate analysis in order to distinguish the varieties. Principal Coordinate Analysis (PCOA) was calculated using DECENTRE, EIGEN and GRAPHICS as described by Rohlf (2000) to complement cluster analysis. From data of microsatellite markers, dissimilarity coefficients between varieties were computed following Jaccard's coefficient method. Principal co-ordinate analysis was used to represent the varieties in bi-directional space. Dendrogram was constructed using NJ method based on dissimilarity matrix and to get better reliability of the construction 1000 bootstrapping was used following DARwin 5.0.148 software program (http://darwin.cirad.fr/darwin/Home.php). The dissimilarity matrix based on morphological traits and molecular data was compared following Mantel (1967) test using NTSYS pc Ver. 2.20 program (Sneath and Sokol 1973).

Result

Morphological diversity

The morphological traits varied widely and the highest polymorphism was observed for seed coat color ranging from white to black through all intermediate colors (Pandey et al. 2013; Prasad and Gangopadhyay 2011). Few other traits like flower pigmentation, leaf shape and the size of extra floral nectaries varied extensively among sixty genotypes (Fig. 1). Thirty seven morphological characters were used for correlation analysis between traits using Microsoft® Excel, (2007). Correlations among traits varied from −0.48 to 1.0. The dissimilarity distance matrix based on morphological data suggested that among 1750 comparison, the highest distance value was estimated between the genotypes UMA and V1 (14.12). On the contrary, EC- 107 (USA) and EC-108 (USA) were very close to each other with lowest Euclidean distance value of 1.58. Sixty sesame genotypes were grouped into two major Clusters I and II based on morphological traits, each subdivided into smaller sub clusters (Fig.2). Overall most of the local Indian developed genotypes were placed in Cluster-I and majority of exotic collections were in Cluster II. In general distribution of genotypes did not reflect their geographical origin. The cluster IIC consisted of maximum number of genotypes (12 accessions), while CL-ID and Cl-IG comprised of single genotype. The first ten Eigen vectors accounted for 65 % of the variation present in the sixty genotypes (Table 3). A data matrix plot based on the morphological descriptors was subjected to PCA for estimating genetic differentiation among 60 genotypes of sesame. The scatter plot based on these components disclosed a pattern of eleven groups. The plot showed that the genotypes, Tillotama and V-12 were distinctly different from each other. Most of the genotypes were scattered between these two varieties with extremely low factor scores along the two dimensions. Distribution of genotypes according to geographical origin was lacking in the matrix plot of sixty genotypes (Fig. 4). The genetic dissimilarity estimates for sixty sesame genotypes were employed to generate a two-dimensional distribution plot by using factorial analysis with DARwin V5.0.148 software (Perrier et al. 2003). The genotypes could be classified into four main clusters. Cluster I consist of 34 genotypes, CL-II had single genotype, CL-III, the second largest cluster with 23 genotypes and CL-IV contained only two indigenous genotypes (Figs. 3 and 4). This analysis could easily distinguish exotic and Indian genotypes into two distinct groups.

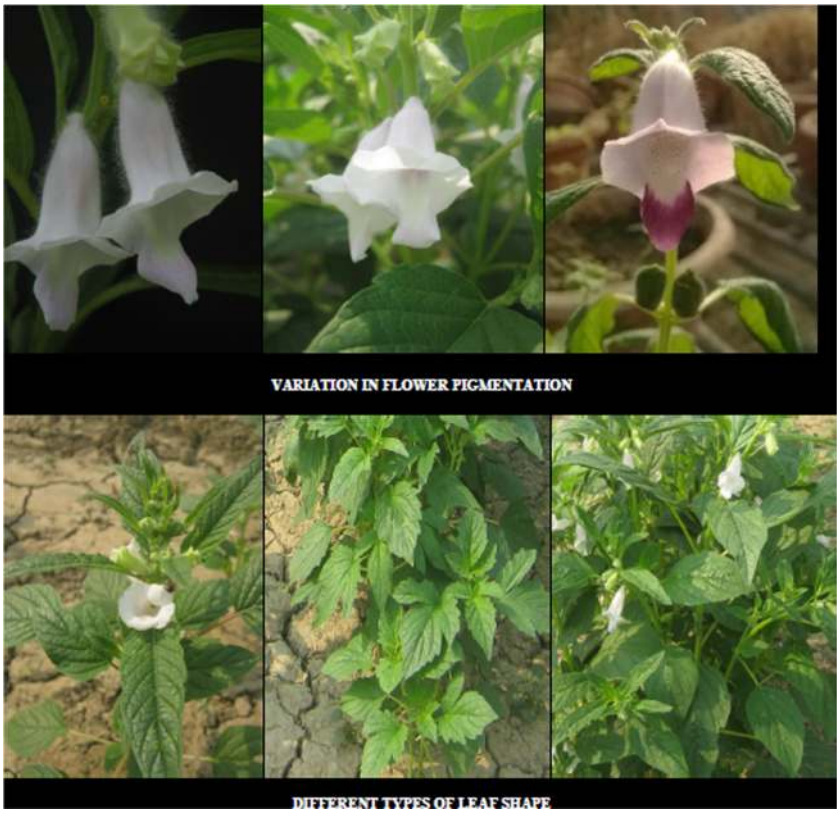

Fig. 1 Representative image showing variation of morphological traits in sesame a. Flower pigmentation (Showing variation in pigmentation along the lip region of corolla tube) and b. Leaf shape of sesame plant at Calcutta University experimental farm, Baruipur

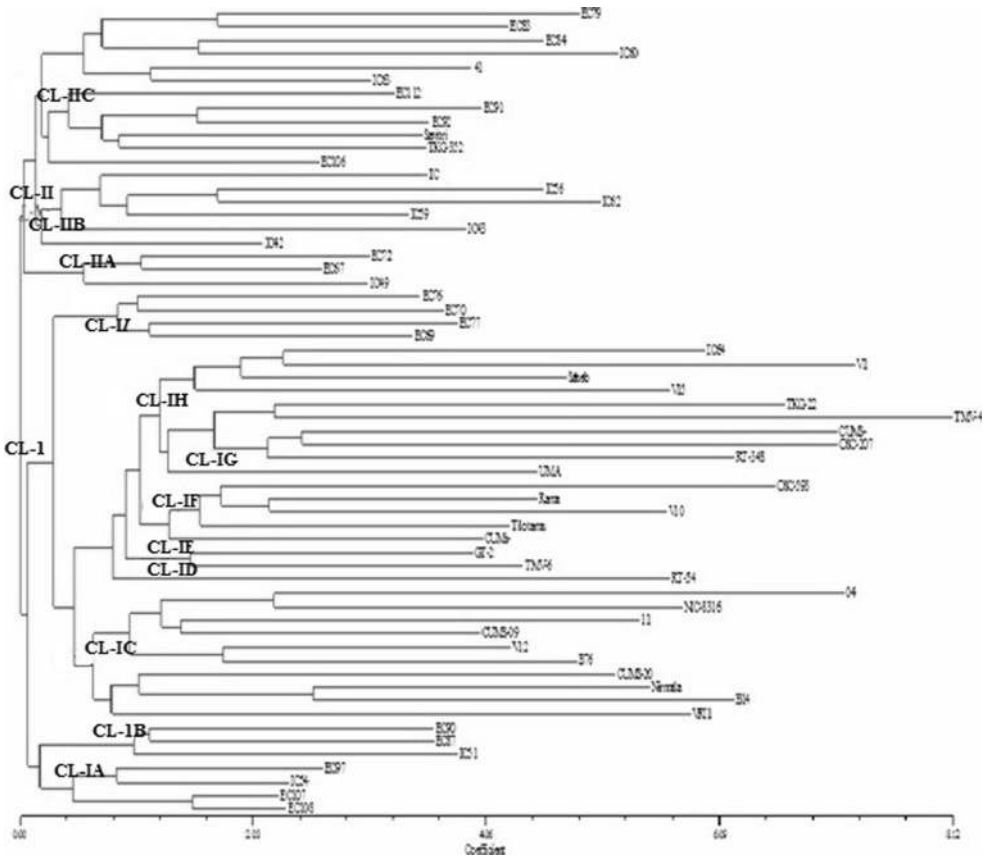

Fig. 2 Dendrogram based on the cluster analysis of sixty genotypes through distance matrix based on thirty seven morphological traits. All the genotypes grouped into major clusters – CL-I and CL-II with sub clusters. Cluster I has eight sub-clusters (CL-IA- CL-I*I*) and Cluster II has three sub-clusters (CL-IIA- CL-IIC)

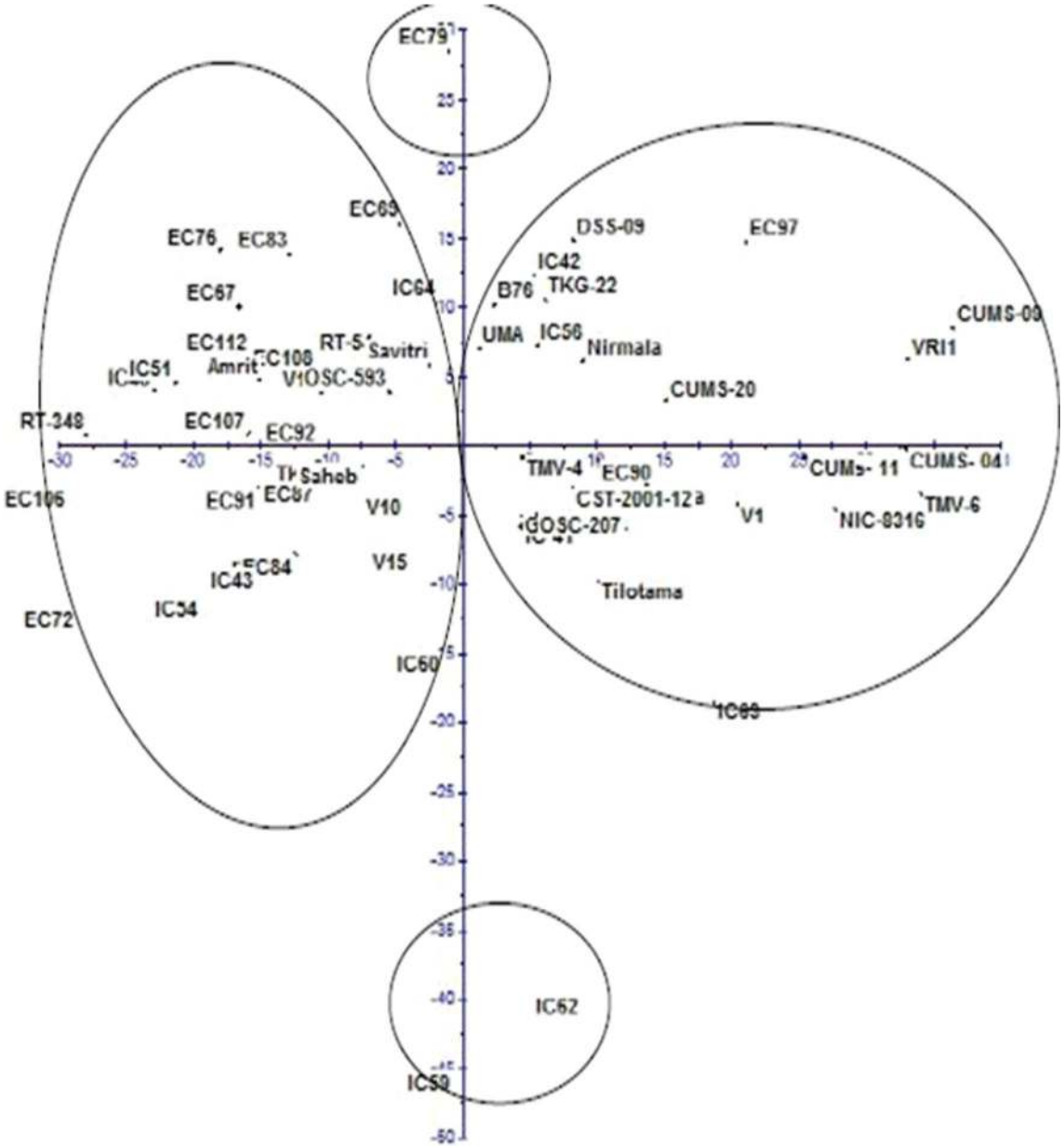

Fig. 3 Representation of the 1–2 plane of factorial analysis based on thirty seven morphological traits for sixty sesame genotypes. Two indigenous collections namely IC-62 and IC59 occupied extreme place through factorial analysis

Most of the exotic and indigenous genotypes were grouped in CL-I with seven Indian cultivars. It was interesting to note that exotic collections were mostly placed in CL-I on the other hand, the CL-III contained the developed Indian genotypes.

Molecular diversity

Initially 36 SSR markers including 29 Est-SSR and 7 genomic SSR were selected for preliminary screening and the microsatellite markers displaying non-specific bands, without polymorphism or PCR products were discarded. Finally, eight primers were selected and the primers produced multiple alleles in each genotype (Fig. 5). Altogether 27 alleles were detected among the 60 genotypes, with an average of 3.37 alleles per locus. The number of alleles ranged from 2 to 6 alleles. The data showed that the SSR primers, SSR-Es-12, followed by SSR-Es-15, and SSR-Gn-03 respectively, was more polymorphic, while, the primers SSR-Es-14 and SSR-Gn-10 exhibited two alleles (Table 2). Genetic diversity per locus in general, was found to be high for most of the SSR loci and ranged from 0.37 to 0.74, average being 0.57 (Table 2). In this study the low average value i.e. 0.57 can be partly explained by the fact that three out of eight microsatellites were developed from ESTs, which are generally less variable than unexpressed regions of the genome. The dissimilarity coefficients between varieties based on the 1–0 matrix of

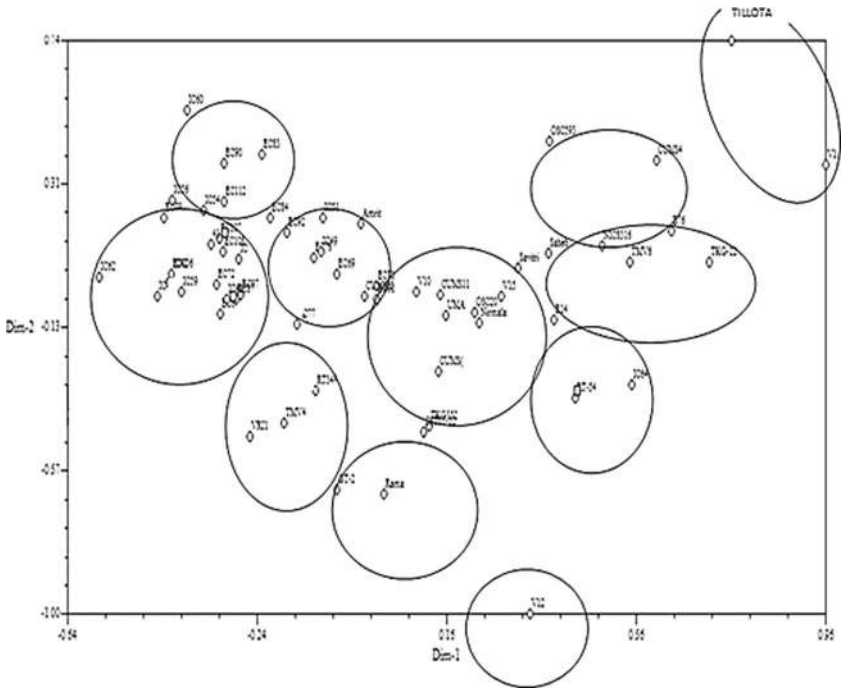

Fig. 4 Two dimensional matrix plot of sixty genotypes was computed through principal component analysis based on morphological data. All the sixty genotypes divided into eleven groups based on 37 traits including both quantitative and qualitative traits

microsatellite markers were taken into account to study the relationship between the diverse sesame germplasm accessions. The estimates of dissimilarity coefficient for 60 genotypes ranged from 2.05 to 13.51. The pair of accession namely, Savitri and EC-112 showed highest distance of 13.51 with few other pairs like Savitri and EC-91, EC-108 and Savitri. When dendrogram was constructed by NTSYS pc Ver2.20, the genotypes were grouped into two clusters I and II (Fig. 6). The clustering pattern showed that two clusters namely CL-IIB and CL-IIC contained four genotypes each. The clusters consisted of genotypes from different geographical origin. Genotypes in CL-IID were from USA and Bangladesh along with two developed Indian genotype, while, Cluster IIB consisted of three Indian genotypes with one indigenous collection. Cluster IB consisted of two Indian genotypes namely NIC-8316 and B14 while, CL-V contained one genotype each namely, from USA, West Bengal and Odisha. The CL-IA and CL-IIC comprised of 11 Indian genotypes each.

Three Indian genotypes namely TMV-6, GT-2, Rama and V-15 along with five exotic collections from Bulgaria, Bangladesh and USA and two indigenous collection formed in cluster IA. Cluster IIC consisted of nine Indian developed genotypes from Rajasthan, Karnataka, West Bengal and Tamil Nadu along with two indigenous and two exotic collections. CL-IC was the largest cluster, with 12 genotypes. Out of 12 genotypes, four were local indigenous collections, three exotic collection, and rest five were Indian genotypes from Kanpur, Rajasthan, Odisha and West Bengal. Cluster ID and IIA consisted of both Indian and exotic collections respectively, indicating the idea that in general, geographical origin did not play role in cluster composition. In the present study, the two-dimensional distribution plot in factorial analysis revealed that the sixty genotypes were classified into five broad clusters. CL-I, the largest cluster consisted of twenty six genotypes, mostly developed Indian genotype followed by indigenous collection and three exotic collections. While,

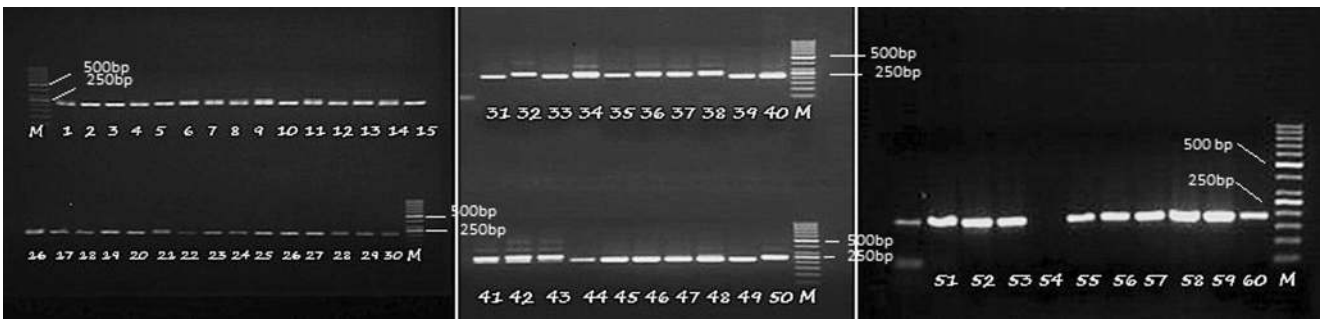

Fig. 5 Representative gel images showing SSR pattern of sesame genotypes amplified by SSR-ES-12 in sixty sesame accessions. Allele size ranges from 200 to 242 bp with PIC value of 0.74. The average PIC value was 0.57

Table 2 Characteristics of polymorphic microsatellite markers in sesame (*Sesamum indicum* L)

Name	Locus	GenBank Accession Number	Primer Sequence	Repeat Motif	^{T}A (^{0}c)	Number of Allele (A)	Size Range of Alleles (bp)	Effective Allele (Aep)	PIC Value
SSR-GN-03	SESAME-09	AY838907	F:5′-CCCAACTCTTCGTCTATCTC-3′ R: 5′-TAGAGGTAATTGTGGGGGA-3′	$(CT)_{18}$	58	3	179–220	2.90	0.65
SSR-GN-05	GBSSR-SA-123	AY838916	F: 5′-GCAAACACATGCATCCCT-3′ R: 5′-GCCCTGATGATAAAGCCA-3′	$(TC)_{21}$, $(TC)_{15}$	58	4	259–282	1.58	0.37
SSR-GN-06	GBSSR-SA-182	AY838921	F: 5′-CCATTGAAAACTGCACACAA-3′ R: 5′-TCCACACACAGAGAGCCC-3′	$(AT)_{11}$,$(TC)_{18}$,$(TG)_{12}$	60	2	200–223	2.33	0.57
SSR-GN-07	GBSSR-SA-184	AY838922	F: 5′-TCTTGCAATGGGGATCAG-3′ R:5′CGAACTATAGATAATCACTTGGAA-3′	$(TC)_{20}$	60	4	160–196	2.77	0.63
SSR-GN-10	GBSSR-SA-72	AY838913	F: 5-′GCAGCAGTTCCGTTCTTG -3′ R: 5′-AGTGCTGAATTTAGTCTGCATAG-3′	$(CT)_{9}$	52	2	268–300	1.66	0.40
SSR-ES-12	GBSSR-SA-135	AY838917	F :5′-GCTGAGGAGTCTTGAAGCAGA-3′ R: 5′-CAAAATCCCCCAACTCGATA -3′	–	60	6	200–242	3.86	0.74
SSR-ES-14	GBSSR-SA-178	AY838920	F: 5′-AAACCCGCTAAGGGACTCAT-3′ R:5′-CATGGCTTCTGGCTTTCTTC-3′	–	60	2	243–260	1.99	0.49
SSR-ES-15	GBSSR-SA-83	AY838914	F: 5′-TGCAGGAATGAACTCAAGGA-3′ R:5′-ACCTTATTCCCAGCCCACTT-3′	–	60	4	208–236	3.69	0.72
MEAN	–	–	–	–	–	3.37		2.59	0.57

Cluster II, consist of fifteen genotypes six developed Indian genotypes along with four indigenous and five exotic collections and CL- III consisted of nine genotypes, seven exotic and two developed Indian genotypes. CL-IV contained seven genotypes namely, CUMS-20, RT-348, CUMS-09, TKG-22, EC-90, CST-2001–12 and CUMS-04. CL-V the smallest cluster consists of three exotic collections, namely EC-108, EC-69 and EC-87 respectively (Fig. 7). Phenotypic and genetic diversity did not corroborate with each other especially in cluster composition. Furthermore, the correlation coefficient (r-0.0026) for the two clustering matrices tested by a Mantel test showed an insignificant correlation between phenotypic and molecular marker information (Table 3).

Discussion

Sesame genotypes possess a large innate genetic variability, which should be taken into consideration when planning conservation strategies or when sesame variability is used in breeding programs. Out of these thirty seven traits under study, few quantitative traits like number of capsules per plant, number of seeds per capsule and 1000 seed weight directly contributes to yield per plant in sesame and these are governed by multiple genes and hence for their better utilization in plant breeding program QTL-mapping is required to dissect the genetics of these traits (Morrell et al. 2012). Wu et al. (2014a) reported the first QTL mapping of yield-attributing traits with a high-density genetic map using a RIL population in sesame and the findings of their study solidified the base for studying the main agricultural traits and employing MAS toward genetic improvement in sesame. The correlation study suggests that the traits related to trichomes of different reproductive and vegetative plant parts showed high correlation coefficients. High correlation coefficient was also observed for the pigmentation traits and leaf profile; similar findings have been reported by Prasad and Gangopadhyay (2011). Based on the results of Euclidean distance data, hybridization between EC-108 and Savitri genotypes with moderate diversity both at molecular and phenotypic level would like to produce desirable segregants in segregating generation with great promise of developing promising breeding materials in advance generation. Likewise, another pair could be EC-90, the black seeded and Savitri, the light brown seeded genotypes occupying different cluster with moderate genetic divergence value. The clustering pattern based on morphological data showed that the clusters consisted of genotypes from different geographical origin but, in some cases geographical origin had also a role in cluster composition. Similar finding was also reported by earlier researchers like Tabatabaei et al. (2011) and they reported that no relationship between genetic diversity and origin of accessions in composition of clusters, but some Iranian genotypes have the tendency to group

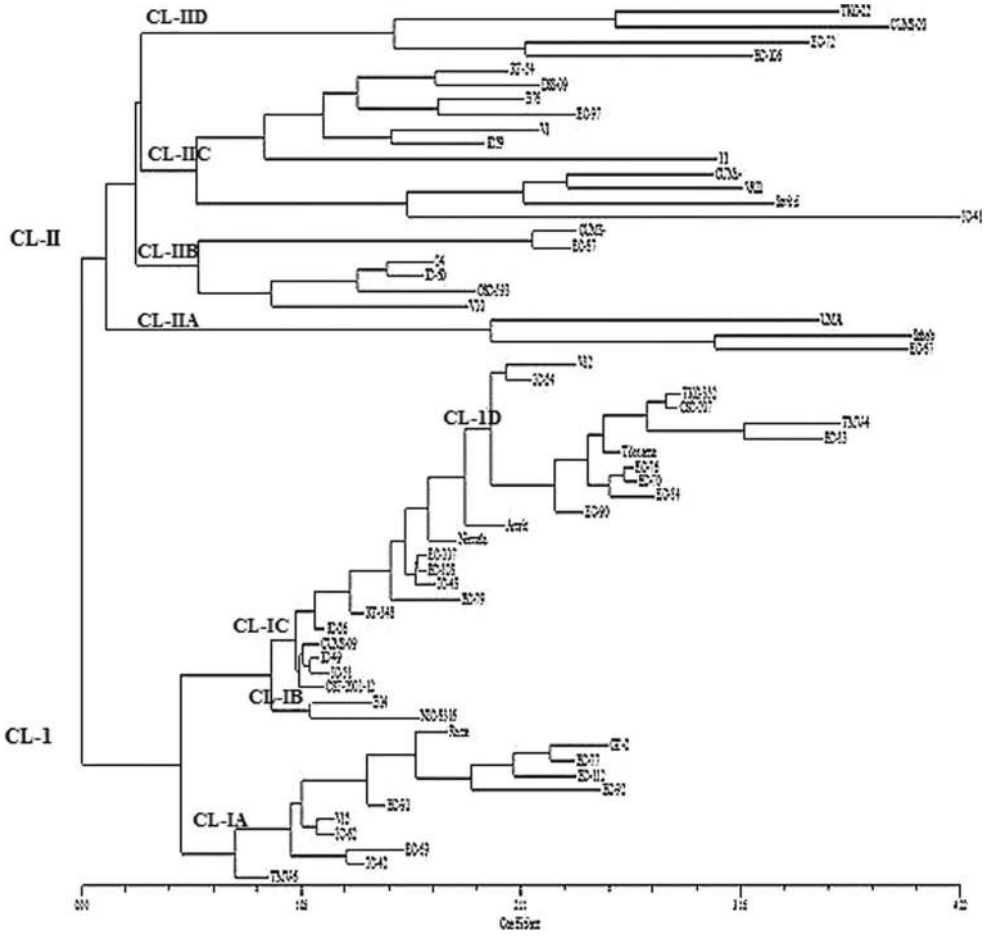

Fig. 6 Dendrogram based on the cluster analysis of 60 genotypes through distance matrix based on molecular markers. The sixty genotypes grouped into two main groups- CL-I and CL-II, each subdivided into smaller sub-clusters. CL-I has four sub-clusters (CL-IA- CL-ID) and cluster II also has four sub-clusters (CL-IIA- CL-IID)

together. Ercan et al. (2004) used RAPD for testing a wide range of germplasm collection of Turkey, showing low genetic variation among genotypes with similar geographic origin and large variation among genotypes within a region. Zhang et al. (2007) reported lack of association between genotypes and geographical origin. Our finding also finds support from

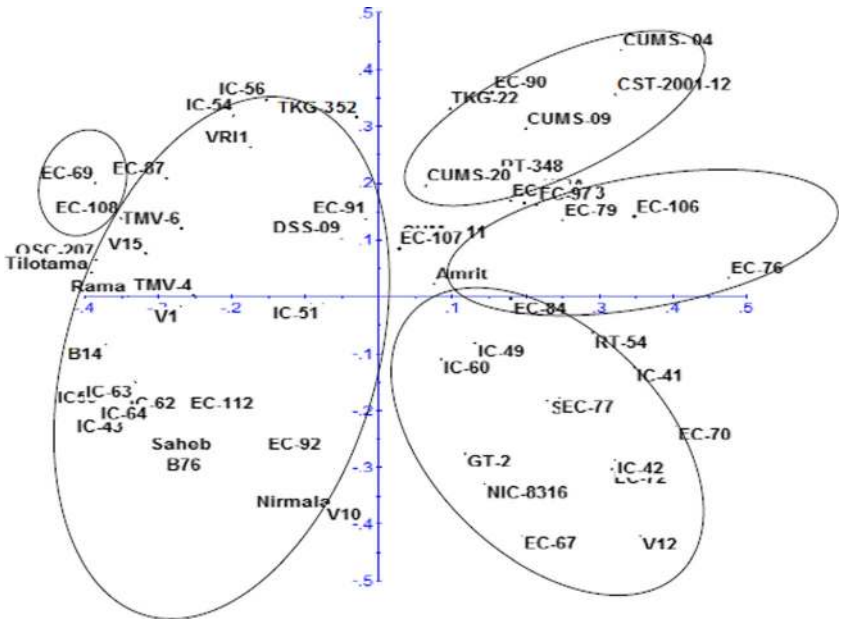

Fig. 7 Representation of the 1–2 plane of factorial analysis based on molecular markers for sixty sesame genotypes. CL-I, the largest cluster with twenty six genotypes and cluster-V is the smallest cluster with three genotypes respectively

Table 3 Eigen values for morphological descriptors, corresponding to matrix plot figure % of variance explained

Characters	Eigen Value	Percentage	Cumulative
Seed Coat Colour	4.87878630	13.1859	13.1859
Stem Shape	3.22677742	8.7210	21.9069
Stem Pubescence	2.80003621	7.5677	29.4746
Leaf Arrangement	2.59968127	7.0262	36.5008
Leaf Shape	2.16190336	5.8430	42.3437
Leaf Angle	1.97873115	5.3479	47.6917
Number Of Branches	1.97445906	5.3364	53.0280
Petal Colour	1.71233349	4.6279	57.6560
Petal Hairiness	1.55478717	4.2021	61.8581
Flowers/Axel	1.34437810	3.6335	65.4916
Capsule Shape	1.31889933	3.5646	69.0561
Capsule Hairiness	1.20360932	3.2530	72.3091
Branching Habit	1.12218717	3.0329	75.3421
Shape Of Stem Hair	1.05413493	2.8490	78.1911
Stem Branching	0.96020227	2.5951	80.7862
Leaf Hairiness	0.89239229	2.4119	83.1981
Basal Leaf Profile	0.83258512	2.2502	85.4483
Basal Leaf Margin	0.68676696	1.8561	87.3045
Lobe Incision Of Basal Leaf	0.58436054	1.5794	88.8838
Petiole Colour	0.49722499	1.3439	90.2277
Petiole Hairiness	0.45471846	1.2290	91.4566
Extra Floral Nectaries	0.41762274	1.1287	92.5853
Extra Floral Nectaries Colour	0.40252018	1.0879	93.6732
Calyx Hairs	0.35259078	0.9529	94.6262
Interior Corolla Colour	0.33209440	0.8976	95.5237
Interior Corolla Pigment	0.28426226	0.7683	96.2920
Lower Lip Colour	0.26694246	0.7215	97.0135
Capsule Arrangement	0.21754238	0.5880	97.6014
Capsule Beak Shape	0.20457550	0.5529	98.1543
Plant Height	0.17538522	0.4740	98.6284
Days To 50 % Flowering	0.13134863	0.3550	98.9833
Days To Maturity	0.11357390	0.3070	99.2903
Capsule Length	0.10644317	0.2877	99.5780
Capsules Per Plant	0.06837385	0.1848	99.7628
Seeds Per Capsules	0.04594096	0.1242	99.8869
1000 Seed Weight	0.02217329	0.0599	99.946
Seed Yield/Plant	0.01965538	0.0531	100.00

the study by Kim et al. (2002) where clustering of genotypes did not indicate any clear division based on their geographical origin. In a similar study Surapaneni et al. (2014) did not observ any differentiation of sesame genotypes according to geographical origin.

Powell et al. (1996) reported that SSR markers gave a good discrimination between closely related individuals in some cases even when only a few loci were applied. Therefore, SSR markers would be one of the useful molecular markers in sesame genetic diversity analyses and in marker assisted breeding programs. Our study showed congruence with the previous studies by Surapaneni et al. (2014) and Zhang et al. (2010). Zhang et al. (2010) reported that SSR markers are appropriate for evaluation of genetic diversities in sesame and also concluded that extensive genetic divergence existed among indigenous and exotic collections of sesame. Park et al. (2013) reported that 41genotype-specific alleles were identified for 12 of 14 SSR markers. Sun et al. (2009) used seven SSR markers to study genetic diversity and discriminate among recently distributed rice accessions and Lopes et al. (2006) used 11 SSR markers to discriminate among 46 Portuguese grapevine accessions. Euclidean distances based on the 1–0 matrix of microsatellite markers were taken into account to study the relationship between the diverse sesame germplasm accessions. The clustering pattern indicated that geographical origin did not play role in cluster composition even at molecular level. Wu et al. (2014b) suggested that domestication along with advanced plant breeding techniques have likely narrowed the genetic basis of cultivated sesame. Many newly developed sesame varieties were bred with a few number of landrace in their pedigree. The genetic variation in sesame was consequently reduced by genetic drift and selection. Characterization of genetic diversity of available landraces especially the indigenous and exotic collection by molecular markers is of great value to assist parental line selection and breeding strategy design (Wu et al. 2014b). The understanding of these landraces can provide a better foundation for further conservation and utilization of these resources. The weak correlation coefficient for the two clustering matrices by a Mantel test showed an insignificant correlation between phenotypic and molecular marker information. Zhang et al. (2012) concluded that molecular measures of genetic diversity may not fully explain quantitative genetic variability. Therefore, the combination of phenotypic and molecular based analysis in genetic diversity assessment of the sesame appears to be momentous in developing any breeding program. The SSR markers under study can be utilized as a useful tool for genetic linkage map construction, genetic diversity detection, and marker-assisted selective breeding in future sesame breeding program.

Acknowledgments The authors are grateful to Dr. A.R.G. Ranganatha, Project Co-coordinator (Sesame and Niger), Indian Council of Agricultural Research for providing us seeds of sesame genotypes.

References

Ali GM, Yasumoto S, Seki-Katsuta M (2007) Assessment of genetic diversity in sesame (*Sesamum indicum* L.) detected by amplified fragment length polymorphism markers. Electron J Biotechnol 10: 12–23

Anthony O, Ogunshakin R, Vaghela S, Patel B (2015). Towards sustainable intensification of sesame-based cropping systems diversification in northwestern India. J Food Secur 3(1):1–5.

Ashri A (1989) In: Robbelen G, Downey RK, Ashri A (eds) Sesame- Oil Crops of the World. McGraw-Hill Publishing Company, New York, pp. 375–387

Banerjee PP, Kole PC (2009) Analysis of genetic architecture for some physiological characters in sesame (*Sesamum indicum* L.). Euphytica 168:11–22

Bedigian D, Seigler DS, Harlan JR (1985) Sesamin, sesamolin and the origin of sesame. Biochem Syst Ecol 13:133–139

Bhat KV, Babrekar PP, Lakhanpaul S (1999) Study of genetic diversity in Indian and exotic sesame (*Sesamum indicum* L.) germplasm using random amplified polymorphic DNA (RAPD) markers. Euphytica 110:21–34

Bhattacharyya U, Pandey SK, Dasgupta T (2014) Identification of Est-SSRs and FDM in sesame (*Sesamum indicum* L.) through data mining. Sc0068 J Agric Sci 4:60–69

Dixit A, Jin M.H, Chung JW, Yu JW, Chung HK, Ma KH, Park YJ, Cho EG (2005) Development of polymorphic microsatellite markers in sesame (*Sesamum indicum* L.). Mol Ecol Notes 5: 736–738

Duhoon SS (2004) Exploitation of heterosis for raising productivity in sesame in new directions for a diverse planet: proceedings of the 4th international crop science congress, Brisbane, Australia Vol 26.

Ercan AG, Taskin M, Turgut K (2004) Analysis of genetic diversity in Turkish sesame (*Sesamum indicum* L.) populations using RAPD markers. Genet. Resour. Crop. Evolution 51:599–3607

FAOSTAT (2015) Food and agricultural organization of the united nations. Statistical Database

Fukuda Y, Nagata M, Osawa T, Namiki M (1986) Contribution of lignan analogues to antioxidative activity of refined unroasted sesame seed oil. J Am Oil Chem Soc 63:1027–1031

Hawkes J (1983) The diversity of crop plants. Harvard Unirversity Press, Cambridge

IPGRI and NBPGR (2004) Descriptors for Sesame (*Sesamum* spp.). International Plant Genetic Resources Institute, Rome and National Bureau of Plant Genetic Resources, New Delhi, India.

Kim DH, Zur G, Danin-Poleg Y, Lee S, Shim K, Kang C, Kashi Y (2002) Genetic relationships of sesame germplasm collection as revealed by inter-simple sequence repeats. Plant Breed 121:259–262

Laurentin HE, Karlovsky P (2006) Genetic relationship and diversity in a sesame (*Sesamum indicum* L.) germplasm collection using amplified fragment length polymorphism (AFLP). BMC Genet 7:10

Laurentin HE, Karlovsky P (2007) AFLP fingerprinting of sesame (*Sesamum indicum* L.) cultivars identification, genetic relationship and comparison of AFLP informativeness parameters. Genet Resour Crop Ev 54:1437–1446

Lopes MS, Santos MR, Dias JE, Mendonça D, Machado ADC (2006) Discrimination of Portuguese ape vines based on microsatellite markers. J Biotechnol 127:193444

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220

Morrell PL, Buckler ES, Ross-Ibarra J (2012) Crop genomics: advances and applications. Nat Rev Genet 13(2):85–96

Pandey SK, Das A, Dasgupta T (2013) Genetics of seed coat color in sesame (*Sesamum indicum* L.). Afr J Biotechnol 12:6061–6067

Park JH, Suresh S, Cho GT, Choi NG, Baek HJ, Lee CW, Chung JW (2013) Assessment of molecular genetic diversity and population structure of sesame (*Sesamum indicum* L.) core collection accessions using simple sequence repeat markers. Plant Genet Resour Charact Util:1–8. doi:10.1017/S1479262113000373

Perrier X, Flori A, Bonnot F (2003) Methods for data analysis. In: Hamon P, Seguin M, Perrier X, Glazmann JC (eds) Genetic diversity of cultivated tropical plants. Science Publishers, Inc and Cirad, Montpellier, pp. 31–63

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

Prasad R, Gangopadhyay G (2011) Phenomic analyses of Indian and exotic accession of sesame (*Sesamum indicum* L.). J. Plant breed. Crop Sci 3:336–352

Ram SG, Sundaravelpandian K, Kumar M, Vinod KK, Bapu JK, Raveendran TS (2006) Pollen–pistil interaction in the inter-specific crosses of *Sesamum* sp. Euphytica 152:379–385

Rohlf FJ (2000) NTSYS-PC Numerical taxonomy and multivariate analysis system In FJ Rohlf, NTSYS-PC- Numerical Taxonomy and Multivariate Analysis System. Applied Biostatistics, Exerter Publishing Ltd., New York, Software, ISBN 0–925031–18–6.

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci 81:8014–8018

Sneath PHA, Sokol RR (1973) Numerical taxonomy, the principles and practice of numerical classification. W. H. Freeman and Co., San Francisco

Sun MM, Choi KJ, Kim HS, Song BH, Woo SH, Lee CW, Jong SK, Cho YG (2009) Genetic diversity and discrimination of recently distributed rice varieties in Korea by SSR markers. Kor J Breed Sci 41: 164–125

Surapaneni M, Yepuri V, Vemireddy RL, Ghanta A, Siddiq EA (2014) Development and characterization of microsatellite markers in Indian sesame (*Sesamum indicum* L.). Mol Breed 34(3):1185–1200

Tabatabaei I, Pazouki L, Bihamta MR, Mansoori S, Javaran MJ, Niinemets Ü (2011) Genetic variation among Iranian sesame (*Sesamum indicum* L.) accessions vis-à-vis exotic genotypes on the basis of morpho-physiological traits and RAPD markers. Austr. J. Crop Sci 5:1396–1407

Wei LB, Zhang HY, Zheng YZ, Guo WZ, Zhang TZ (2008) Developing Est-derived microsatellites in sesame (*Sesamum indicum* L.). Acta Agron Sin 34(12):2077–2084

Weir BS (1990) Genetic data analysis. Methods for discrete population genetic data. Sinauer Associates, Inc. Publishers.

Wu K, Liu H, Yang M, Tao Y, Ma H, Wu W, Zuo Y, Zhao Y (2014a) High-density genetic map construction and QTLs analysis of grain yield-related traits in sesame (*Sesamum indicum* L.) based On. RAD-Seq Techonol; BMC Plant Biolog 14(274):1471–2229

Wu K, Yang M, Liu H, Tao Y, Mei J, Zhao Y (2014b) Genetic analysis and molecular characterization of Chinese sesame [*Sesamum indicum* L.] cultivars using insertion-deletion [InDel] and simple sequence repeat [SSR] markers. BMC Genet 15(1):35

Zeven A, Zhukovsky P (1975) Dictionary of cultivated plants and their centres of diversity. Wageningen, PUDOC

Zhang XR, Chen KR, Peng J, Xu ZY (2004) The RAPD analysis and genetic diversity of selected sesame germplasm. Chinese J Oil Crop Sci 26:34–37

Zhang P, Zhang H, Guo W, Zheng Y, Wei L, Zhang T (2007) Genetic diversity analysis of *Sesamum indicum* L. germplasms using SRAP and Est-SSR markers. Acta Agron Sin:1696–1702

Zhang YX, Zhang XR, Hua W, Wang LH, Che Z (2010) Analysis of genetic diversity among indigenous landraces from sesame (*Sesamum indicum* L.) core collection in China as Revealed by SRAP and SSR Markers. Genes and Genomics 32(3):207–215

Zhang H, Wei L, Miao H, Zhang T, Wang C (2012) Development and validation of genic-SSR markers in sesame by RNA-seq. BMC Genomics 13:316