Molecular Diversity in Castor Germplasm Collection Originated from North-Eastern Hill Province of India

Kanti Meena*[1], Anjani K[2], VenkatRamya K[2]

[1]*Central Research Institute for Jute and Allied Fibres, Barrackpore, India*
[2]*Directorate of Oilseeds Research, Hyderabad, India*

***Abstract*- A total of fifty one accessions of castor (*Ricinus communis* L.) germplasm collected from North-eastern hill region of India were employed for current investigation involving 35 castor EST-SSR marker for polymorphism at genetic level. Seventy six alleles with an average frequency of 2.14 band per primer was generated among which 29 were polymorphic with a polymorphic percentage of 82.8. The number of allele ranged between 3 and 2 based on which PIC value was generated between 0.04 - 0.73 with an average PIC value of 0.38. The genetic similarity index used for dendrogram, resulted into 3 major and 6 minor clusters with a Dice similarity coefficient ranging from 0.65 to 0.90. Confusion probabilities and limit of discriminating power ranged from 0.021 to 0.623 with a mean of 0.402 and 0.348 to 0.979 with a mean of 0.598 respectively. Highly diversity genotypes can be used for crop improvement programme like tagging of germplasm, identification or elimination of duplicates in the gene stock, establishment of core collections, sorting of populations for genome mapping and in breeding program to obtain more variability in the castor crops.**

***Keywords*- *Ricinus communis*, Castor, Genetic diversity, EST-SSRs, Polymorphism.**

I. INTRODUCTION

Castor (*Ricinus communis* L., 2*n*=2x = 20, Euphorbiaceae), is industrially important non-edible oilseed crop widely cultivated in the arid and semi-arid regions of the world [1]. It appears to have originated in tropical Africa suggested by high diversity of castor plants in Ethiopia. The plants can be self or cross pollinated by wind, with out-crossing as a predominant mode of reproduction ([2], [3]). Historically it has been cultivated as an agricultural crop for the oil derived from its seeds. The seed of castor contain more than 45% oil and this oil is rich (80–90%) in an unusual hydroxyl fatty acid, ricinoleic acid [4]. Unique chemical and physical properties, of castor oil makes it a raw material for numerous and varied industrial applications, such as manufacture of polymers, coatings, lubricants for aircrafts, cosmetics, hydraulic fluids, plastics, artificial leather, manufacture of fiber optics, bulletproof glass and bone prostheses and as antifreeze for fuels and lubricants utilized in aircraft and space rockets and for the production of biodiesel ([4], [5]), farming and medicines [6].

India is the largest producer of castor in the world and also one of the centers of diversity of castor, where great diversity exists across the country in varied eco-system [7]. North-east hill region of India has a different ecosystem. Castor is not a cultivated crop of this region, yet limited type of castor plants have been found growing wild, as castor is not the cultivated crop of these regions conservation and maintenance of castor growing in these areas is not possible. Thus, wild type castor accessions were collected during the exploration [8] and conserved *ex-situ* at the Directorate of Oilseeds Research, Hyderabad, India (17.366^{o}N and 78.478^{o}E).

In the current investigation, castor accessions collected from North-east hill province of India, were evaluated to study the diversity among them at genetic level. Castor EST - SSR (Expressed Sequence Tag- Simple Sequence Repeat) markers were used to study the genetic diversity.

II. MATERIALS AND METHODS

A. Plant material

Seed source of 51 castor accessions collected from different parts of Norths-east hill region of India were procured from Germplasm Management Unit at the Directorate of Oilseed Research (DOR), Hyderabad. Two castor cultivar *viz.*,'Aruna' and 48-1 were employed study to know any genetic relationship existing with NEH castor germplasm accessions (Table 1).

Table 1: List of North-East accessions collection and 2 castor cultivars

Code No.	Germplasm	Collection site	Code No.	Germplasm	Collection site
1	RG1921	Assam	57	RG1980	Mizoram
2	RG1922	Assam	58	RG1981	Mizoram
4	RG1924	Assam	59	RG1982	Manipur
7	RG1927	Assam	64	RG1987	Manipur
8	RG1928	Assam	66	RG1990	Manipur
9	RG1929	Assam	67	RG1991	Manipur
11	RG1931	Assam	68	RG1992	Manipur
12	RG1932	Assam	74	RG1998	Manipur
14	RG1934	Assam	79	RG2004	Manipur
17	RG1937	Meghalaya	80	RG2005	Manipur
19	RG1939	Meghalaya	85	RG2011	Nagaland
20	RG1940	Meghalaya	86	RG2012	Nagaland
21	RG1941	Meghalaya	88	RG2014	Nagaland
23	RG1943	Meghalaya	89	RG2015	Nagaland
25	RG1945	Meghalaya	92	RG2018	Assam
26	RG1946	Meghalaya	95	RG2025	Assam
28	RG1948	Meghalaya	98	RG2028	Assam
29	RG1949	Meghalaya	99	RG2029	Assam
32	RG1953	Meghalaya	108	RG2038	Assam
35	RG1956	Meghalaya	112	RG2042	Assam
37	RG1958	Meghalaya	115	RG2045	Assam
38	RG1959	Meghalaya	116	RG2046	Assam
39	RG1961	Meghalaya	124	RG1936	Meghalaya
46	RG1969	Meghalaya	126	RG2023	Assam
53	RG1976	Meghalaya	Cultivars		
55	RG1978	Mizoram	52	48-1	
56	RG1979	Mizoram	53	Aruna	

B. Genomic DNA isolation

Ten fresh young leaves from each accession were bulked for DNA extraction. Bulked DNA samples were used to ensure sufficient tissue and the main aim was to study the gross genetic differences prevalent among these collections. The genomic DNA was isolated by Cetyl-trimethyl ammonium bromide (CTAB) method [9] with minor modifications as per the requirement. Sample (3g) were grounded to fine powder in liquid nitrogen and suspended in 15 ml of pre-warmed extraction buffer (100 mM Tris-HCl, 20 mM EDTA pH-8, 1.4 mM NaCl, 2 % CTAB and 0.2 % ß-mercaptoethanol). The homogenate was extracted twice with chloroform: isoamyl alcohol (24:1) and centrifuged at 6,000 rpm for 15 min at 4° C. The nucleic acid was precipitated from the aqueous phase by adding 0.6 volume of isopropanol followed by centrifugation at 10,000 rpm for 10 min. The pellet was washed in 70% ethanol, air dried and dissolved in TE buffer (10 mM Tris– HCl (pH 8.0), 1 mM EDTA). The pellet was washed with wash Buffer I (absolute ethanol and 3M sodium acetate) and wash Buffer II (absolute ethanol and 1M ammonium acetate), air dried and dissolved in appropriate quantity of TE buffer.

C. PCR amplification

The Polymerase Chain Reaction (PCR) was performed out in total volume of 20 µl reaction involved ~ 25 ng of template DNA, 0.08 mM dNTPs, 1X reaction buffer (containing 1.5mM $MgCl_2$) (Bangalore, Genei), 5 Pmoles of each forward and reverse primer and 1U of Taq DNA polymerase (Bangalore, Genei)). The amplification was performed using Master-Cycler gradient Eppendorf Version 2.1 (Eppendorf, USA), programmed according [10] with minor modifications. The PCR cycle profile followed as: initial denaturation at 94ºC for 5 min and then 30 cycles of 30 sec denaturation at 94ºC, 30 sec primer annealing at 56ºC and 30 Sec extension at 72ºC and 10 min at 72ºC for final product extension.

Amplification products were stored at 4ºC until further use. PCR reaction mixture was size fractionated on 3.0% agarose (Banglore Genei) gel containing ethidium bromide in 1 X TAE buffer at 120 volts using horizontal gel electrophoresis system. The 100 bp DNA ladder runs along the sides of the amplified product to determine their approximate size. The amplified fragments were visualized under ultraviolet light and photographed with gel documentation system (Gene Flash Syngene Bioimaging).

D. Staining of PAGE- gels

The amplified PCR products were further subjected onto 6% polyacrylamide gel on a Sequi- Gen (Bio-Rad, USA) sequencing cell in 1X TBE buffer at 100 W, 50mA at 55ºC for 2 hrs. The gel was stained with silver stain for 30 minutes and the visualized bands were photographed.

E. Statistical analysis of banding pattern

Statistical analysis of banding pattern polymorphic bands were scored visually for their presence or absence with each primer. The scores were obtained in the form of matrix with '1' and '0', which indicate the presence and absence of bands in each accession respectively. Molecular weights of the bands were estimated by using 100 bp DNA ladder as standards. Polymorphism information content (PIC) values were calculated for eSSR markers, in order to characterize the capacity of each primer to reveal or detect polymorphic loci among the genotypes. The binary data scoring was used to construct a dendrogram using NTSYS software [11]. The genetic associations between accessions were evaluated by calculating the Dice similarity coefficient for pairwise comparisons based on the proportions of shared bands produced by the primers. The similarity coefficients were used for cluster analysis and dendrogram was constructed by the unweighted pair-group method (UPGMA).

For individual primer/primer combination, confusion probability (*C*), and limit of discriminating power (*D*L) were calculated [12] and polymorphic information content (PIC) was calculated [13]. Effective number of

patterns/assay unit *(P)*, Average number of alleles per locus (*nav*), fraction of polymorphic loci (*β)*, effective number of alleles per locus (*ne)*, total number of effective alleles (*Ne*). The EST-SSR markers were characterized by assay efficiency index (*A*i), marker index (*MI*) and effective multiplex ratio (*nβ)*.

III. RESULT AND DISCUSSION

Previous studies in castor germplasm have been made using protein based markers isozymes ([14], [15], [16], [7]). Limitation of these markers has shifted the castor molecular research towards the use of DNA based markers such as AFLP [6], gSSR ([6], [18]), RAPD [19], RAPD and ISSR [20], SNPs [21] and recently eSSRs [22]. Microsatellite markers are becoming the markers of choice due to the level of polymorphism, as well as higher reliability ([23], [24]). In castor only few gSSR markers have been exploited for various studies.

To assess the genetic variation among 51 castor accessions collected from Northeast region of India along with two commercial cultivars, initial analysis was carried based on **t**hirty-five different castor eSSR to see whether they could generate suitable polymorphism among different samples. It was found that twenty nine primer pairs produced polymorphic products with a polymorphic percentage of 82.8 and 6 primer pairs produced monomorphic products. A total of 76 alleles were amplified using 29 primer pairs among 51 genotypes and two commercial cultivars analyzed with an average frequency of 2.41 bands per primer. Eleven eSSRs amplified maximum number of three alleles, while 18 markers amplified minimum number of two alleles (Table 2).

PIC value is indicative of the effectiveness of SSR loci information and measure the information of a given marker locus for the pool of genotypes [25]. The PIC value ranged from 0.04 (mRcDOR04) to 0.73 (mRcDOR24) with an average PIC value of 0.38 (Table 2), indicating that the locus were potent enough. Increase in the allele number did not affect the PIC value much. The average PIC value in the current investigation was higher compared to 0.21 reported with SNPs [21]. In comparison, the value was lower than 0.40 with gSSR and AFLP in worldwide germplasm genotyping [6] and 0.82, 0.88 with RAPD & ISSR respectively with the breeding material [20]. Genetic diversity of 72 castor accessions of north east based on quantitative and qualitative method was studied and reported seven most promising accessions based on the performance of yield attributing traits [26].

Molecular size of the bands ranged from 110 bp with primer mRcDOR203 to 350 bp with primer mRcDOR07. Confusion probabilities for EST-SSR primers ranged from 0.021 to 0.623 with a mean of 0.402. The limit of discriminating power (*D*p) was used to measure the efficiency of SSRs, *Dp* value ranged from 0.348 to 0.979 with a mean of 0.598 (Table 3). The Effective number of patterns per assay unit (P), assay efficiency index (Ai), marker index (MI) were calculated (Table 3). The extent of polymorphism obtained among the 51 accessions along two cultivar as revealed by eSSR primer mRcDOR130 and CES49 (Fig 1) on 6 % polyacrylamide page gel and mRcDOR07, mRcDOR24 & CES09 (Fig 2) on 3.5 % agarose gel.

Table 3: Detail information obtained with EST-SSR marker in North-East castor germplasm

S. No.	Indexes	Abbreviations	EST-SSR
1	Number of assay units	U	35.0
2	Number of polymorphic bands	n_p	70.0
3	Number of monomorphic bands	n_{np}	6.0
4	Average Number of polymorphic bands/assay unit	n_p/U	2.0
5	Number of loci	L	35.0
6	Number of loci/assay unit	n_u	1.0
7	Average confusion probability	C	0.402
8	Average limit of discriminating power	D_L	0.598
9	Effective number of patterns/assay unit	P	2.5
10	Average number of alleles per locus	n_{av}	2.0
11	Fraction of polymorphic loci	B	0.92
12	Total number of effective alleles	N_e	59.5
13	Assay efficiency index	A_i	1.7
14	Effective multiple ratio	E	0.92
15	Marker index	MI	0.55

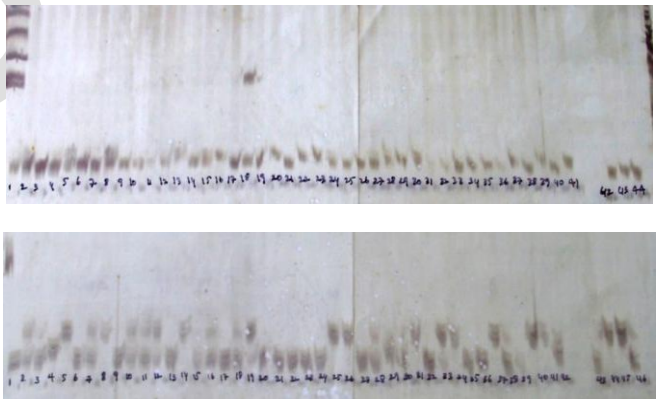
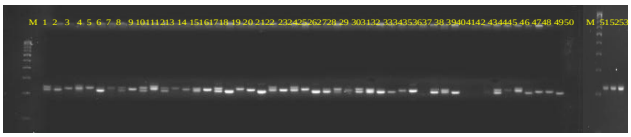


Fig 1: Screening of NE castor germplasm with polymorphic markers mRcDOR 130 and CES 49, respectively on 6%

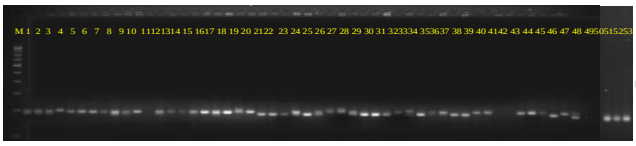
With the aid of 76 allelic information dice similarity coefficient of UPGMA cluster analysis was used to construct a dendrogram from DARWin 5.0 (Fig 3), which illustrated the overall genetic relationship among the 51 accessions studied. Based on the dendrogram, the 51 North-east castor accessions and two castor cultivars were grouped into three major cluster. Although there were three distinct clusters (I, II and III), only one cluster (III) was dominant. Among the three major clusters, Cluster III grouped the largest number of accessions (25). The major cluster-III generated 2 minor clusters IIIa (16 accessions) and IIIb (9 accessions). The other two major clusters (I and II) grouped 14 accessions each. The cultivar, 48-1

was grouped in the cluster-II, similarly major cluster-I grouped the cultivar Aruna along with other accessions of the group indicating genetic relationship with members of cluster II & I respectively. It is important to note that all the 3 clusters generated included accessions from all the region except the 4 accession of Nagaland, which were separated together in cluster I. This indicates that the germplasm was not restricted to a particular region they were highly distributed throughout the North east region.

EST SSR profile of mRcDOR07



EST SSR profile of mRcDOR24



EST SSR profile of CES 09

Fig 2. EST-SSR profile for 51 accessions of NE collection along wi castor cultivars on agarose

EST-SSR marker system revealed low genetic variation among the castor accession collected from North-east hill region of India. It is similar to the results of Foster et al., 2010 and Allan et al., 2008, who found limited genetic diversity and structure for populations in genome wide diversity studies in castor with SNP, AFLP and SSRs. In worldwide genotyping of castor germplasm from different geographical areas [21], no unique alleles in any of the accessions across different regions were found. Based on different parameters, such as the number of alleles, PIC, DP value used in this study to analyze the data, polymorphic information for 51 castor accessions with 35 eSSR markers have been generated.

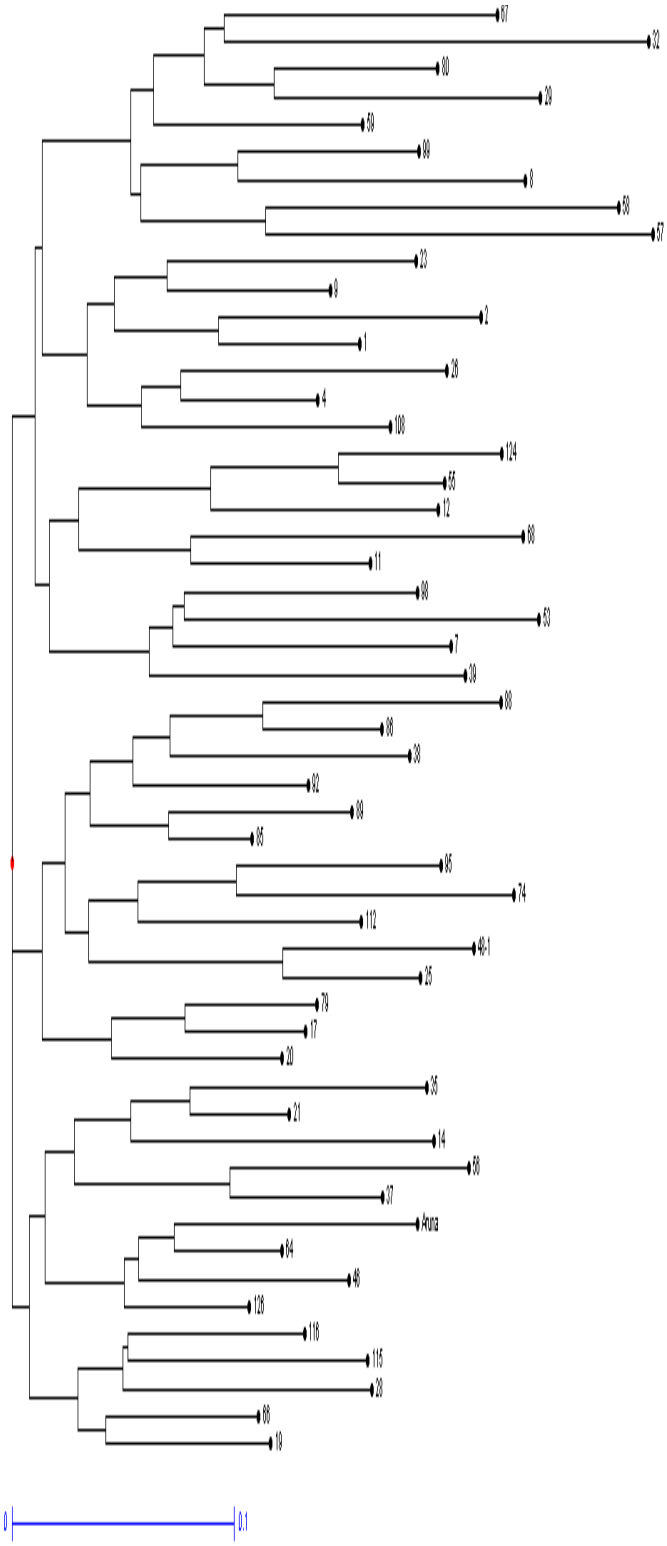


Table 2: Genetic diversity of 35 eSSR loci in Nort-east castor germplasm as indicated by No. of alleles, range of allele size, PIC, confusion probability (*C*), discriminating power (Dp) for EST-SSR primers

S.No	Marker	No. of alleles	Expected size (bp)	Range of size (bp)	Effective allele (Ae)	PIC	*C*	Dp
1.	mRcDOR 04	2	276	270-280	1.03	0.04	0.021	0.979
2.	mRcDOR 07	2	342	320-350	2.08	**0.52***	0.442	0.558
3.	mRcDOR 09	2	161	150-170	1.66	0.40	0.354	0.646
4.	mRcDOR 11	2	118	115-135	1.44	**0.69***	0.605	0.395

5.	mRcDOR 24	3	239	220-250	3.70	**0.73***	0.623	0.377
6.	mRcDOR 26	1	195	180	-	0	-	-
7.	mRcDOR 28	2	244	244-253	1.23	0.19	0.102	0.898
8.	mRcDOR 49	2	266	260-300	1.40	0.29	0.252	0.748
9.	mRcDOR 51	2	289	280-290	1.85	0.46	0.447	0.553
10.	mRcDOR 55	2	196	180-195	2.17	**0.54***	0.446	0.554
11.	mRcDOR 65	3	245	230-270	1.85	0.46	0.436	0.564
12.	mRcDOR 69	1	195	189	-	0	-	
13.	mRcDOR 76	2	162	150-185	1.19	0.16	0.158	0.842
14.	mRcDOR 77	2	247	240-250	1.66	0.40	0.374	0.626
15.	mRcDOR 86	1	306	300	-	0	-	-
16.	mRcDOR 92	3	190	150-195	2.38	**0.58***	0.553	0.447
17.	mRcDOR 103	2	185	182-195	1.40	0.29	0.252	0.748
18.	mRcDOR 119	3	180	180-195	2.27	**0.66***	0.652	0.348
19.	mRcDOR 130	3	130	130-145	2.22	**0.55***	0.457	0.543
20.	mRcDOR 142	3	216	210-216	1.72	0.42	0.401	0.599
21.	mRcDOR 153	1	360	340	-	0	-	-
22.	mRcDOR 166	2	200	180-190	1.96	0.49	0.477	0.523
23.	mRcDOR 203	2	110	110-125	2.04	**0.51***	0.477	0.523
24.	mRcDOR 206	2	180	180-195	1.78	0.44	0.397	0.603
25.	mRcDOR 215	2	280	270-282	1.54	0.35	0.240	0.760
26.	mRcDOR 223	3	180	180-193	2.38	**0.58***	0.470	0.530
27.	mRcDOR 225	1	320	320	-	0	-	-
28.	CES 09	3	188	185-200	1.53	**0.63***	0.577	0.423
29.	CES 10	2	217	217-225	2.05	**0.60***	0.465	0.535
30.	CSE 28	2	239	240-255	1.66	0.40	0.254	0.745
31.	CES 49	3	209	140-230	2.50	**0.63***	0.602	0.398
32.	CES 80	3	189	190-180	1.51	0.34	0.277	0.723
33.	CES 123	3	242	242-255	3.12	**0.68***	0.589	0.411
34.	CES 125	1	172	170	-	0	-	-
35.	CES 137	2	200	200-215	1.44	0.31	0.255	0.745
Total no. of alleles		**76**	**Average**		**1.33**	**0.38**	**0.402**	**0.598**

* denotes eSSR with PIC > 0.5

IV. CONCLUSION

This information generated from the above work allows us to hypothesize the fact that these loci could generate sufficient polymorphism that could allow the use of above markers for establishing genetic relationships among closely related accessions. The primers showing PIC >0.5 can also be useful for a wide range of genetic investigations such as association mapping, linkage map and other crop improvement studies involved in castor.

ACKNOWLEDGEMENT

The authors gratefully acknowledge V. Dinesh Kumar, Principal Scientist, Directorate of Oilseed Research, Rajendranagar, Hyderabad, for kindly providing castor EST-SSR used in the study.

REFERENCES

[1]. Govaerts Frodin, RDG., Radcliffe-Smith, A. (2000). World Cheklist and bibliography of Euphorbiaceae (with Pandaceae). Redwood Books limited, Trowbridge, Wiltshire.

[2]. Meinders, HC., Jones, MD. (1950). Pollen shedding and dispersal in the castor plant *Ricinus communis* L. Journal of Agronomy. 4: 206–209.

[3]. Brigham, RD. (1967). Natural out crossing in dwarf internode castor, *Ricinus communis* L. Crop Science. 7: 353-355.

[4]. Jeong, GT., Park, DH. (2009). Optimization of Biodiesel Production from Castor Oil Using Response Surface Methodology. Applied Biochemistry and Biotechnology. 156: 431–441.

[5]. Ogunniyi, DS. (2006). Castor oil: a vital industrial raw material. Bioresources and Technology. 97(9): 1086-1091.

[6]. Allan, G., Williams, A., Rabinowicz, PD., Chan, AP., Ravel, J., Keim, P. (2008). Worldwide genotyping of castor bean germplasm (*Ricinus communis* L.) using AFLP and SSRs. Genetics Resources and Crop Evolution. 55: 365-378.

[7]. Anjani, K. (2012). Castor genetic resources: A primary gene pool for exploitation. Industrial Crops Products. 35: 1–4.

[8]. Anjani, K.., Chakravarty, SK.., Prasad, MVR. (1994). Collecting castor (*Ricinus communis* L.) germplasm in North-Eastern Hill Province of India. IBPGR News Lett. for Asia, The Pacific and Oceania, 17: 13.

[9]. Doyle, JJ., Doyle, JL. (1990). Isolation of plant DNA from fresh tissue. Focus. 12: 13–15.

[10]. Athma, P., Vaidyanatha, K., Reddy, TP. (1982). Peroxidase isoenzyme studies in certain varieties of *Ricinus communis* L. Industrial Journal of Botany. 5: 178–182.

[11]. Williams, JGK., Kubelik, AR., Livak, KJ., Rafalsky, A., Tingey, SV. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research. 18: 6531-6535.

[12]. Rohlf, FJ. (2000). NTSYS-PC: Numerical taxonomy and multivariate analysis system version 2.02k. State University of New York, Stony Brook N.Y.

[13]. Tessier, C., David, J., This, P., Boursiquot, JM., Charrier, A. (1999). Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* L. Theoretical and Applied Genetics. 98: 171–177.

[14]. Anderson, JA., Churchill, GA., Autroque, JE., Tanksley, SD., Swells, ME. (1993). Optimizing selection for plant linkage map. Genome, 36: 181–186.

[15]. Sathaiah, V., Reddy, TP. (1988). Esterase isozyme patterns of nine parental diallel set of castor (*Ricinus communis* L.). Proceedings of Conference Cytological and Genetics. 1: 240–246.

[16]. Sathaiah, V., Reddy, TP. (1983). Inheritance of isoperoxidases in castor. Paper presented at the XV Intl Congr Genet, December 12-21, New Delhi, India, Abstracts Part 1; Pp. 387.

[17]. Sathaiah, V., Reddy, TP. (1984). Qualitative and quantitative analysis of peroxidase and esterase enzymes in castor (*Ricinus communis* L.). Indian Journal of Botany. 7: 68–73.

[18]. Bajay, MM., Pinheiro, JD., Batista, CEA., Nobrega, MBM., Zucchi, MI. (2009). Development and characterization of microsatellite markers for castor (*Ricinus communis* L.), An important oleaginous species for biodiesel production. Conservation Genetics Resources. 1(1): 237-239.

[19]. Vivodík, M., Balážová, Z., Gálová, Z. (2014). RAPD Analysis of Genetic Diversity of Castor Bean. World Academy of Science, Engineering and Technology. 8(7): 615-618.

[20]. Bhavesh, BG., Kumar, N., Amritpal, SS., Bhupendra, SP., Ravikiran, R., Subhash, N., Jadeja, GC. (2010). Assessment of genetic diversity in castor (*Ricinus communis* L.) using RAPD and ISSR markers. Industrial Crops Products. 32: 491-498.

[21]. Foster, JT., Allan, GJ., Chan, AP., Rabinowicz, PD., Ravel, J., Jackson, PJ., Keim, P. (2010). Single nucleotide polymorphisms for assessing genetic diversity in castor bean (*Ricinus communis*). BMC Plant Biology. 10: 13-22.

[22]. Qiu, L., Yang, C., Tian, B., Yang, JB., Liu, A. (2010). Exploiting EST database for the development and characterization of EST-SSR markers in castor bean (*Ricinus communis* L.). BMC Plant Biology. 10: 278-287.

[23]. Plaschke, J., Ganal, MW., Röder, MS. (1995). Detection of genetic diversity in closely related bread wheat using microsatellite markers. Theatrical Applied Genetics. 91: 1001-1007.

[24]. Fu, YB., Peterson, GW., Richards, KW., Somers, D., DePauw, RW., Clarke, JM. (2005). Allelic reduction and genetic shift in the Canadian hard red spring wheat germplasm released from 1845 to 2004. Theatrical Applied Genetics. 110: 1505–1516.

[25]. Kupper, CPJM.., Chiotatto, A., Zucchi, M., Colombo, C., Carbonell, S., Costa, MJM., Gazaffi, R., Franco, GAA., de Campos, T., de Souza, A. (2011). Genetic diversity in cultivated carioca common beans based on molecular marker analysis. Genetics and Molecular Biology. 34: 88.

[26]. Gogoi, SN., Sarma, RN., Rajan, RK. (2011). Genetic diversity of castor (*Ricinus communis* L.) germplasm in North-eastern region of India. Indian Journal of Sericulture. 50(1): 51-57.