

Biochemical Characterization and Genetic Diversity of *Pongamia pinnata* (L.) Pierre in Eastern India

Kanchan Kumari¹, Amrita Sinha¹, Sanjay Singh^{1,*} and B.N. Divakara²

¹Molecular Biology and Tree Physiology Lab, Botany, Silviculture and NWFP Division, Institute of Forest Productivity, Ranchi-835303, India

²Institute of Wood Science and Technology, Malleśvaram, Bangalore-560 003, India

Abstract

Biochemical characteristics of 24 *Pongamia pinnata* genotypes (candidate plus trees) from three agroclimatic zones were estimated and molecular characterization through RAPD markers was done. Various biochemical characters viz. seed oil, total carbohydrates, protein, acid value and Iodine number recorded significant variation among different genotypes. The highest seed oil content was 41.87% while seeds of 14 genotypes recorded above average (32.11%) for the trait. Seed oil and protein content exhibited a significant positive correlation and moderate heritability. Out of the initially selected twenty-five random primers, twenty-two RAPD primers were found to be highly reproducible and produced a total of 183 loci of which 147 (80.32%) loci were polymorphic. Percentage of polymorphism varied from 44% to 100% with an average of 80.62%. High level of genetic variation was found among different genotypes of *P. pinnata*. Both molecular and oil content (biochemical) markers appeared useful in analyzing the extent of genetic diversity in *Pongamia* and the result of these analyses will help to better understand the genetic diversity and relationship among populations. Overall, the *Pongamia* genotypes included in the study showed a correlation with their geographical origins such that genotypes from the same region tend to have higher genetic similarity as compared to those from different regions. However, in UPGMA based Nei's analysis, some genotypes were found not to be grouped based on geographical origins possibly due to the exchange of germplasm over time between farmers across the regions. The results from oil content analyses showed that several genotypes in 'Central and Western Plateau' agroclimatic zone of Jharkhand displayed a good potential for high oil content. The study provides insight about *P. pinnata* populations in Jharkhand (India) and constitutes a set of useful background information that can be used as a basis for future breeding strategy and improvement of the species.

Key Words: biofuel, genetic diversity, polymorphism, RAPD markers, seed oil

Introduction

With declining reserves of oil for the production of fossil fuels, biofuels can make a substantial contribution to the future energy demands of the domestic and commercial economies. Among all TBOs (Tree Bourne Oilseed species) of India, which can yield oil as a source of energy in the

form of biodiesel, Karanj [*Pongamia pinnata* (L.) Pierre] has been found to be one of the most suitable species due to its various favorable attributes like its hardy nature, high oil recovery and quality of oil. *P. pinnata* is a fast growing leguminous deciduous tree, commonly known as Indian Beech, Pongam, Honge and Karanj, grows about 15-25 m tall and is well-adapted to semi-arid and humid zones and can be

Received: April 12, 2013. Revised: July 26, 2013. Accepted: August 8, 2013.

Corresponding author: Sanjay Singh

Molecular Biology and Tree Physiology Lab, Botany, Silviculture and NWFP Division, Institute of Forest Productivity, NH-23, Lalgutwa, Ranchi-835303, India

Tel: 91-651-2340548, Fax: 91-651-2526006, E-mail: sanjaysingh@lycos.com

grown up to 1,000 m above sea level (Allen and Allen 1981). It is naturally distributed along the coasts and river banks in India and Myanmar (Janardhanan et al. 2003). *P. pinnata* is indigenous to the Indian subcontinent and has been successfully introduced to humid tropical regions of the world as well as parts of Australia, New Zealand, China and the United States (Wani et al. 2006; Scott et al. 2008). It is a nitrogen-fixing hardy tree with tolerance to a wide range of abiotic stresses like drought, frost, heat, limestone, salinity, sand and shade.

The use of vegetable oils from plants such as *P. pinnata* has the potential to provide an environmentally acceptable fuel, the production of which is greenhouse gas neutral, with reductions in current diesel engine emissions (Raheman and Phadatare 2004). *P. pinnata* seed contains high amount of polyunsaturated fatty acids which are an important source of biodiesel (Sarma et al. 2005; Sharmin et al. 2006). The potentiality of *P. pinnata* oil as a source of fuel for the biodiesel industry is well recognized (Azam et al. 2005; Karmee and Chadha 2005). However, to increase the biodiesel production it's important to identify genotypes of *P. pinnata* bearing high oil-yielding seeds.

Genetic improvement of any species depends on the amount of intra-specific genetic variability. Of the several methods developed to measure the genetic diversity, the molecular approach which directly quantifies genetic variability at DNA level is the most efficient. In fact molecular markers serve as a powerful tool for breeders to search for new sources of variation and to investigate genetic factors controlling quantitatively inherited characters. PCR based molecular markers have been widely employed in many forest plant species including *Pongamia* for identification, phylogenetic analysis, population studies and genetic linkage mapping (Kesari et al. 2010; Thudi et al. 2010; Sharma et al. 2010a; Sujaththa et al. 2010; Acharya et al. 2004).

Despite reported to be sensitive to amplification conditions (Ellsworth et al. 1993) and to different genomic backgrounds, showing nonparental bands (Ayliffe et al. 1994) RAPD (Random Amplified Polymorphic DNAs) technique has been indicated as one of the most powerful methods for obtaining a large number of genetic markers in a wide range of species (Williams et al. 1990; Hadrys et al. 1992). The high number of random primers of arbitrary sequence available may allow an extensive, random sampling

of the genome (Michelmore et al. 1991), while the outstanding degree of variability detected at the DNA level may increase the amount of genetic information. Therefore, to exploit the potentiality of available resource base, variability and genetic analysis of twenty-four *Pongamia* genotypes selected from various locations in Jharkhand (Eastern India) were assessed for seed oil traits as a scope for further breeding program.

Materials and Methods

Plant material

Germplasm of twenty four genotypes (CPTs- morphologically superior trees) of *P. pinnata* collected from different agroclimatic zones of Jharkhand (latitude and longitudinal range between 22° N to 24° 50' N and 83° 30' E to 87° E, respectively) and maintained at Institute of Forest Productivity, Ranchi, Jharkhand (India) as clonal (grafted plants) trial were utilized (Table 1).

Biochemical analysis

Seed of different genotypes of *P. pinnata* was subjected to biochemical analysis *viz.* total carbohydrate (Hedge et al. 1962; glucose as standard), protein (Lowry et al. 1951; Bovine serum albumin as standard), acid value (Cox and Pearson 1962), iodine number (Horowitz 1975) and total oil content (Sadasivam and Manickam 1992).

The data obtained were subjected to statistical analysis, employing analysis of variance (ANOVA), 'F'-test for significance at $p \leq 0.05$ and computing LSD values to separate means in different statistical groups using statistical software IBM SPSS version 18. The Statistical Package for Agriculture Research, version 2.0 (SPAR 2.0) was used for calculating various variability parameters (Phenotypic Coefficient of Variability, Genotypic Coefficient of Variability, Heritability, Expected Genetic Advance and Genetic Gain).

DNA extraction

Young leaf tissues visually clean and unaffected by pathogens of each genotypes were collected for genomic DNA isolation and analysis. DNA extractions following CTAB method with slight modification (Murray and Thompson 1980; Doyle and Doyle 1990) was done to remove the inter-

Table 1. Details of *Pongamia pinnata* genotypes used for genetic diversity analysis

Sl. no.	Code	Forest division	Latitude	Longitude	Altitude (m)	Agroclimatic zone
1	PPHBGU	Hazaribag north	24°27'10"N	85°31'42"E	370	JH-1
2	PPGDPS	Giridih	24°17'11"N	86°21'55"E	390	Central and North Eastern Plateau Zone
3	PPHBK	Hazaribag west	23°54'19"N	85°19'04"E	610	
4	PPKDBD	Koderma	24°27'21"N	85°46'12"E	380	
5	PPCTUT	Chatra south	23°54'19"N	85°19'04"E	640	
6	PPRCHT	Ranchi	23°33'02"N	84°54'43"E	790	JH-2
7	PPRCPD	Ranchi east	23°17'07"N	85°10'35"E	810	Central and Western Plateau Zone
8	PPLDBN	Lohardaga	23°33'02"N	84°54'43"E	640	
9	PPRCCP	Ranchi east	23°27'45"N	85°28'39"E	630	
10	PPKTIT	Khunti	23°03'05"N	85°13'40"E	700	
11	PPGWVP	Garhwa south	23°55'30"N	83°46'11"E	410	
12	PPSDPS	Simdega	22°35'47"N	84°40'49"E	370	
13	PPGLIG	Gumla	23°07'02"N	84°33'21"E	520	
14	PPLDCN	Lohardaga	23°26'17"N	84°38'36"E	590	
15	PPRCCH	Ranchi	23°28'40"N	85°10'17"E	790	
16	PPRCPS	Ranchi east	23°09'04"N	85°28'40"E	500	
17	PPRCBH	Ranchi east	23°28'36"N	85°01'06"E	610	
18	PPRCJT	Ranchi east	23°33'55"N	85°05'05"E	650	
19	PPLDKD	Lohardaga	23°21'06"N	84°39'16"E	570	
20	PPGLBJ	Gumla	23°08'22"N	84°46'47"E	800	
21	PPGLBB	Gumla	22°52'39"N	84°53'36"E	500	
22	PPPHMB	Porahat	22°52'35"N	85°18'15"E	690	JH-3
23	PPDBPM	Dhalbum	22°27'10"N	86°37'09"E	350	South Eastern Plateau Zone
24	PPSKHN	Saraikeela	22°51'42"N	85°56'55"E	390	

fering polyphenolic compounds and other polysaccharides. Quantification of DNA was accomplished by UV spectrophotometer (Spectroquant[®] Pharo 300) and also by analyzing the DNA on 0.8% agarose gel alongside λ DNA/HindIII digested as standard. DNA was diluted in TE buffer to a concentration of 50 ng/ μ l for use in PCR analysis.

RAPD analysis

For RAPD fingerprinting PCR amplification of genomic DNA was carried out using 25 arbitrary decamer oligonucleotide primers (RPI series, Bangalore GeNei). The PCR reaction mix included 50 ng/ μ l template DNA, 1X Taq Buffer A (Tris with 15 mM MgCl₂), 0.2 mM each deoxyribonucleotide triphosphate (Bangalore GeNei, India). All PCR reactions were performed in a GeneAmp^{*} PCR 9700 Thermal cycler.

The PCR conditions included: the first step as denaturation at 94°C for 5 minutes, followed by 42 cycles of 1 min at

94°C, annealing temperature of 1 min at 30-50°C (annealing temperature was chosen on the basis of gradient) and extension at 72°C for 2 min. A final extension of 7 min at 72°C was used for primer extension. The amplicons obtained were separated on 1.5% Agarose gel in 1X TAE buffer containing 5 μ g/ml Ethidium bromide. The size of amplified fragments was estimated by comparison to the 100 bp DNA ladder (Bangalore GeNei) and gel was photographed on Gel Documentation System (Syngene). To test the reproducibility of the RAPD markers, the reaction were repeated at least twice.

Data analysis

Each unambiguous band was assumed to represent a dominant allele at a unique genetic locus and was recorded as 1 (present) or 0 (absent) for each sample. Only clear, well defined bands were recorded as potential RAPD markers. The number of polymorphic and monomorphic products was determined for each primer for twenty four genotypes.

To avoid taxonomic weighing the intensity of bands was not taken into consideration and only the presence of band was taken as indicative.

The parameters need to be assessed for determination of genetic diversity including, p value i.e. polymorphism of amplified products, the number of observed alleles (N_a), the mean number of effective alleles (N_e), the mean Nei's gene diversity index (H), the Shannon index (I) and the level of gene flow (N_m) were estimated. Within population diversity, total gene diversity and inter population differentiation (G_{st}) were calculated by subjecting the data to the POPGENE software (Yeh et al. 1999).

Results and Discussion

Genetic improvement for enhanced biomass, oil yield and stress tolerance depends on selection, maintenance and use of elite *P. pinnata* accessions/genotypes. The present

study was conducted to find out the superior germplasm in terms of high oil yield by exploiting the available RAPD markers and correlating it to the various characters of seed oil content and environmental factors as totality.

Biochemical characteristics

Plant seeds are designed to carry the genetic material and all the nutrients that are required to establish the next generation of the species. Typical storage forms of carbon in seeds are starch, protein, and oil, the economically valuable products used for food, feed, and industrial applications. However, the proportion of these constituents varies in different species and environmental conditions. In the present study various biochemical characters recorded significant variation among the different genotypes (Table 2). The highest seed oil content (41.87%) was found in the PPCHT followed by PPCPD (39.33%) and PPPHMB (38.80%). Average total oil content was 32.11% and seeds of 14 geno-

Table 2. Biochemical characteristics of *P. pinnata* genotypes

S. no.	Genotype	Total oil content (%)	Carbohydrate (%)	Protein	Acid value (mg KOH/g)	Iodine number
1	PPHBGU	33.600	20.880	7.227	14.597	64.893
2	PPGDPS	33.333	24.680	8.157	5.087	84.503
3	PPHBNK	32.933	18.693	7.430	14.403	87.813
4	PPKDBD	32.800	14.467	7.553	15.930	75.410
5	PPCTUT	26.113	13.333	8.020	21.967	83.853
6	PPRCHT	41.865	18.627	7.687	11.173	66.180
7	PPRCPD	39.333	12.130	7.667	14.737	68.037
8	PPLDBN	36.613	17.987	8.020	17.103	87.900
9	PPRCCP	36.413	15.297	8.013	20.953	88.170
10	PPKTIT	36.133	20.473	7.420	9.247	66.420
11	PPGWVP	33.273	13.440	8.327	16.357	89.937
12	PPSDPS	33.113	14.787	7.170	6.373	77.293
13	PPGLIG	33.067	17.407	7.117	6.083	62.123
14	PPLDCN	32.667	13.407	7.297	5.937	71.137
15	PPRCCH	32.000	15.547	7.487	8.900	77.457
16	PPRCPS	31.600	18.920	7.027	14.203	91.187
17	PPRCBH	31.067	19.533	8.247	8.610	65.243
18	PPRCJT	28.800	17.813	7.630	6.983	67.467
19	PPLDKD	26.000	26.257	7.170	5.943	77.153
20	PPGLBJ	25.867	26.923	7.240	8.080	78.247
21	PPGLBB	21.867	17.713	5.303	15.430	84.283
22	PPPHMB	38.800	17.643	7.113	8.493	61.677
23	PPDBPM	26.993	22.753	6.807	4.413	61.387
24	PPSKHN	26.400	17.993	6.877	4.757	88.610
LSD _{0.01}		6.04	1.91	0.85	2.01	2.23

types recorded above average for the trait. Highest values for carbohydrates (26.92%), protein (8.33%), acid value (21.97 mg KOH/g) and iodine number (91.19) were exhibited by PPGLBJ, PPGWVP, PPCTUT and PPCPS, respectively. The seed biochemical contents in the selected genotypes indicate towards some very promising candidates (e.g. PPCHT) for clonal development and utilization in large scale plantation drives.

Various genotypes were also evaluated for parameters important in biodiesel production. The iodine value (IV) or iodine number is a measure of average amount of unsaturation of fats/oils which is used in standard for biodiesel and accessing their oxidative stability (Knothe 2002). Iodine value (IV) was introduced in biodiesel quality standards for evaluating their stability to oxidation. Biodiesel with high IV is easily oxidized in contact with air. Biodiesel with high IV tends to polymerize and form deposits on injector nozzles, piston rings and piston ring grooves. The tendency of polymerization increases with the degree of unsaturation of the fatty acids. In the *P. pinnata* genotypes IV values were well below the permissible standards required in biodiesels (Barabás and Todoruț 2011).

Similarly acid value (variously referred as neutralization number, acid number or acidity) is a measure of the amount of carboxylic acid groups in a chemical compound, such as a fatty acid and used to quantify the amount of acid present, for example in a sample of biodiesel. This parameter is a direct measure of the content of free fatty acids, thus the corrosiveness of the fuel, of filter clogging and the presence of water in the biodiesel. This parameter can also be used to measure the freshness of the biodiesel. Fuel that has oxidized after long-term storage will probably have a higher acid value. Acid value in most of the genotypes under study was low in contrast to previously reported very high values

(Sharma et al. 2010b) indicating suitability for biodiesel preparation after appropriate modifications.

In selection and improvement of plant species heritability of a certain desirable trait is important consideration. Heritability is defined as the proportion of total variance that is genetic (Burton 1987). Heritability is important because it aids the breeder in determining the amount of progress or genetic gain that can be made through certain breeding strategies. A high heritability means the offspring have a greater chance of exhibiting an expression of the trait that is similar to that of the parents; whereas a low heritability shows a lower level of resemblance. Variability and genetic parameters of seed oil characteristics have been presented in the Table 3. Highest genotypic and phenotypic coefficient of variability was exhibited by acid value (47.33 and 47.81) and the lowest found for protein content (7.47 and 10.11). Acid value also exhibited highest genetic advance. Maximum heritability (98.28) was for iodine number. Protein (54.68) and seed oil percentage (60.22) had moderate heritability. Little difference between phenotypic coefficient of variation (PCV) and genotypic coefficient of variation (GCV) and high estimates of heritability (broad sense) for seed biochemical characters under study revealed the heritable nature of variability present among the genotypes. High estimates of heritability suggested that these traits may serve as important genetic markers in predicting the selection of best individuals.

Correlation of seed oil and other characteristics was also examined. The negative correlation between seed oil and seed protein concentration in many important oilseed crops is a concern for breeders wishing to increase both of these seed quality traits (Wiggins 2012). Generally observed strong inverse correlation between oil and protein content makes it difficult to improve both traits simultaneously.

Table 3. Coefficient of variation (%) and genetic estimates of biochemical characteristics of *P. pinnata* genotypes

Biochemical characteristics	Coefficient of Variability (%)		Heritability (%)	Genetic advance (%)
	Phenotypic	Genotypic		
Carbohydrate (%)	21.68	22.55	92.42	42.93
Protein	7.47	10.11	54.68	11.39
Acid value (mg KOH/g)	47.33	47.81	98.02	96.53
Iodine Number	13.45	13.57	98.28	27.47
Total oil content (%)	13.71	17.67	60.22	21.92

However, in our study seed oil content notably increased with a concomitant increase in protein content. Linear regression analysis and specific correlation coefficient (R^2) revealed a significant positive correlation between total oil

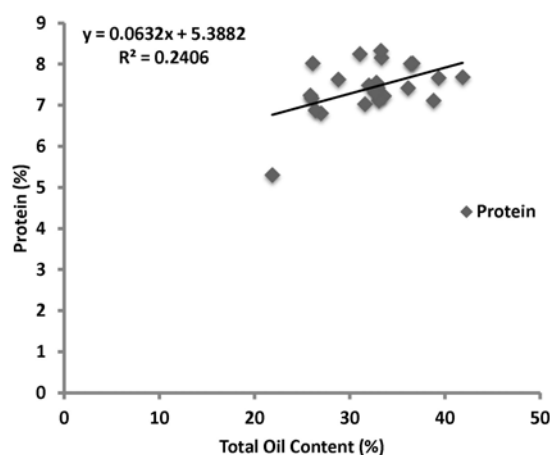


Fig. 1. Linear regression analysis and specific correlation coefficient (R^2) between total oil content and protein.

content and protein ($r=0.491$, $p < 0.05$) (Fig. 1). Overall, moderate heritability and positive correlation among seed oil and protein content in *P. pinnata* suggest possibility of simultaneous selection of these traits for tree improvement and breeding.

Molecular characterization

Out of 25 primers used to assess the existing genetic diversity among 24 genotypes of *P. pinnata*, 22 primers were found to be reproducible across all genotypes, while rest of the primers resulted in either no amplification or stable band patterns. The band patterns produced by some of the RAPD primers have been shown in Fig 2. The selected 22 primers generated a total of 183 distinct bands/fragments out of which 147 (80.32%) bands were found to be polymorphic and 36 bands were monomorphic (Table 4). Number of bands produced by primers ranged from 5 to 12 with an average of 8.3 bands per primer. The primers RPI#13, RPI#14 and RPI#15 produced maximum number (10-11) of polymorphic bands, while RPI#6 and

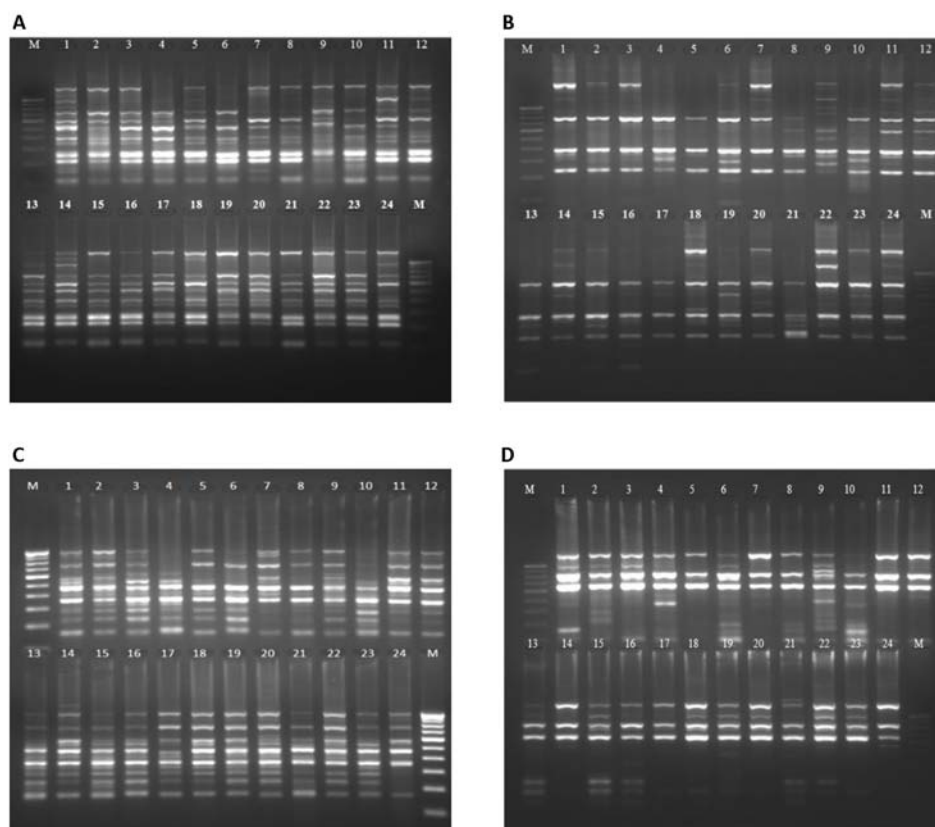


Fig. 2. RAPD pattern of 24 genotypes of *Pongamia pinnata* using RPI#08 (A), RPI#10 (B), RPI#16 (C) and RPI#20 (D), respectively.

Table 4. Scorable DNA bands generated by different random decamer primers through PCR.

Sl. no.	Primer	Sequence	Total number of amplified products	Number of polymorphic products	Number of monomorphic products	Polymorphism (%)	T _m (°C)	Annealing temp (°C)
1	RPI#1	AAAGCTGCGG	7	7	0	100	32	35
2	RPI#2	AACGCGTCGG	6	6	0	100	34	35
3	RPI#3	AAGCGACCTG	7	7	0	100	32	40
4	RPI#4	AATCGCGCTG	6	6	0	100	32	45
5	RPI#5	AATCGGGCTG	9	9	0	100	32	40
6	RPI#6	ACACACGCTG	5	3	2	60	32	35
7	RPI#7	ACATCGCCCA	10	7	3	70	32	45
8	RPI#8	ACCACCCACC	12	7	5	58.33	34	50
9	RPI#9	ACCGCCTATG	5	5	0	100	32	35
10	RPI#10	AGCATGAGCG	8	6	2	75	32	40
11	RPI#11	ACGGAAGTGG	7	7	0	100	32	30
12	RPI#12	ACGGCAACCT	11	7	4	63.63	32	40
13	RPI#13	ACGGCAAGGA	11	10	1	90.90	32	40
14	RPI#14	ACTTCGCCAC	11	10	1	90.90	32	40
15	RPI#15	ACCTGAAGCC	11	11	0	100	32	40
16	RPI#16	AGGCGGCAAG	11	7	4	63.63	34	40
17	RPI#17	AGGCGGGAAC	9	4	5	44.44	34	40
18	RPI#18	AGGCTGTGTC	7	5	2	71.42	32	40
19	RPI#19	AGGTGACCGT	-	-	-	-	32	-
20	RPI#20	AGTCCGCCTC	5	2	3	40	34	40
21	RPI#21	CACGAACCTC	9	8	1	88.88	32	45
22	RPI#22	CATAGAGCGC	-	-	-	-	32	-
23	RPI#23	CCAGCAGCTA	-	-	-	-	32	-
24	RPI#24	CCAGCCGAAC	6	4	2	66.66	34	45
25	RPI#25	GAGCGCCTTC	10	9	1	90	34	40
Total			183	147	36	-	-	-
Average			8.31	6.68	1.636	80.62	32.63	40

RPI#20 produced minimum number of polymorphic bands i.e. 3 and 2, respectively. The percentage of polymorphism ranged from 44% to 100%. Primers RPI#1, RPI#2, RPI#3, RPI#4, RPI#5, RPI#9, RPI#11 and RPI#15 revealed 100% polymorphism where as RPI#17 showed 44% polymorphism. Different primers showed variation in their ability to detect polymorphism but all the twenty-two primers were effective in determining variation among the *P. pinnata* genotypes. A high level of genetic polymorphism (80.32%) is expected for this outcrossing species and comparable with figures reported for some other trees such as *Caesalpinia echinata* (Cardoso et al. 1998) and *Populus tremuloides* (Yeh et al. 1995) using random amplified polymorphic DNA. However, Acharya et al. (2004) detected 99.13% polymorphic RAPD bands in nine

species of the tribe Millettieae.

UPGMA dendrogram based on Nei's (1973) genetic distance between different pairs was correlated with their seed's oil content. The dendrogram (Fig. 3) indicated segregation of 24 genotypes into two distinct groups of population. One group consisted of only one germplasm viz. PPGDPS having its origin as Giridih (belonging to the JH-1, Agroclimatic zone of Jharkhand). The other group consisted of rest of the 23 genotypes. The later larger group was further subdivided into five sub groups. First sub group consisted of 7 members, all of them were collected from Central and Western plateau of Jharkhand (JH-2). In spite of different oil content proportion, they group under a common agroclimatic zone. In this division the germplasms PPGLIG and PPHBNK were found

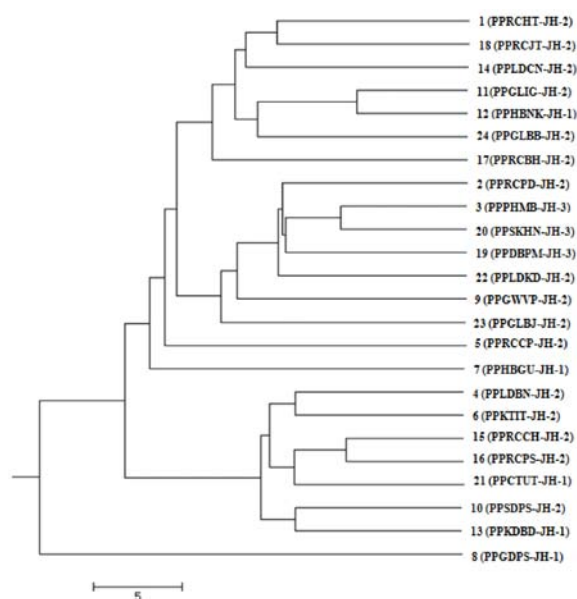


Fig. 3. Dendrogram showing genetic relationship among various *P. pinnata* genotypes.

highly similar and had separated with each other very recently. In the same division the genotypes which yielded high oil content PPRCHT, taken from Ranchi (JH-2; Agroclimatic zone of Jharkhand), was coupled with the plant yielding lowest oil yield PPGLBB, obtained from Gumla (JH-2 Agroclimatic zone of Jharkhand). Despite being from same Agroclimatic zone, the two populations showed greater variation in terms of oil yield.

The next sub-group was further comprised of seven members. Here the populations which belonged to two different agroclimatic zones i.e. Central and Western plateau (JH-2 Agroclimatic zone of Jharkhand) and South-Eastern plateau (JH-3 Agroclimatic zone of Jharkhand) was placed under same subgroup. In this group itself the germplasm PPPHMB and germplasm PPSKHM were very much genetically close to each other and both belonged to the same agroclimatic region (JH-3; South-Eastern plateau of Jharkhand) and have been recently separated from a common ancestor. These two germplasm were in-turn close to PPDBPM which was also obtained from the nearby area of Dhalbhum (belonging to a common agroclimatic zone of Jharkhand i.e. JH-3). The two germplasm PPRCCP and PPHBGU were separated from the above two divisions. The last group of the subdivisions contained seven

members. In this group germplasms PPKDBD and PPCTUT; belonging to Central and North Eastern plateau (JH-1) is placed with other members belonging to Western Plateau (JH-2).

Nei's Analysis of genetic distance and identity (1973) was used to establish the existing genetic variation between the considered genotypes. The wide range (0.55 to 0.88) of pair wise genetic identity between the genotypes indicates the existence of higher genetic diversity among them. Mean heterozygosity (H_t) was 0.2809 and Shanon Information index (Lewontin 1972) was 0.4187 among the genotypes. Average genetic identity of 0.72 suggests that the selected 24 genotypes (CPTs) have contained the vast majority of the natural variation in *P. pinnata*.

According to the Nei's data, the highest genetic similarity exists between the germplasm PPGLIG and germplasm PPHBNK. In the dendrogram both of the germplasm were placed close to each other under one sub-division. The lowest similarity was observed between germplasm PPGDPS and germplasm PPSKHN. They belonged to different divisions in the dendrogram and both of them belonged to different agroclimatic zones of Jharkhand (JH-1 and JH-3, respectively).

Considering that *Pongamia* is indigenous to India, high level of genetic diversity is expected to exist in this species within India. A wide range of variability exists in *Pongamia* with respect to seed and oil traits (Kaushik et al. 2007; Kesari et al. 2009; Mukta et al. 2009). However, there is a large phenotypic diversity in this species, thus providing an opportunity for genetic improvement (Kaushik et al. 2007). Scott et al. (2008) reported wide genetic diversity in plants collected from India and Australia on the basis of DNA fingerprinting. Moderate levels of genetic diversity within and between populations were observed based on RAPD (Kesari et al. 2010). Other woody perennial species such as *Populus tremuloides* and *Isotoma petraea* have displayed lower levels of genetic variation detected with RAPD methodology (Yeh et al. 1995; Bussell 1999).

Jharkhand is divided into three Agroclimatic regions i.e. Central and North Eastern Plateau (JH-1), Western Plateau (JH-2) and South Eastern Plateau (JH-3). In the dendrogram it has been shown that the genotypes which yielded high oil percentage belonged to the second agroclimatic zone of Jharkhand i.e. Central and Western plateau

(JH-2) which account for 33% natural forests of Jharkhand and characterized by late arrival, early cessation and uneven distribution of monsoon rains; and poor water retention capacity of soil. These conditions might have an effect on enhanced seed oil in the genotypes of the region. However, overall evaluation of seed oil content, it appears that individual genotypic factors supersede over environmental factors in determining oil yield. Therefore irrespective of the location or distribution of the *P. pinnata* genotypes, different groups and subgroups have been generated in the dendrogram.

P. pinnata is an important TBO crop in the region valued and exploited for seed oil widely used as fuel and medicinal properties by indigenous communities and farmers. The random distribution of genotypes on the basis of oil yield in the present investigation might be partly due to germplasm exchange between the farmers. Wide range of edaphic and ecological adaptability of the species indicates towards considerable amount of genetic variability to be exploited for potential realization. Comprehensive work on collection of germplasm and evaluating of growth, morphology, seed characteristics and yield traits is ongoing world over but a few reports on genetic variability and divergence studies in *Pongamia* in terms of PCR amplification of the genomic DNA; biodiesel traits; pod and seed traits have appeared (Kaushik et al. 2007; Kesari et al. 2009; Mukta et al. 2009).

Natural distribution of genetic diversity in plant species depends on its isolation, habitat alteration due to climatic change or its evolution, ecological, geographical factors, and more often on human breeding activities (Rao and Hodgkin 2002). Overall, the genetic distance among populations may or may not correlate with the geographical distance between them in some species depending on natural and artificial factors involved in shaping the population genetic structure of the species (Stankiewicz et al. 2001). However, recent studies have reported human selection and genetic drift as prime cause for genetic diversity in sesame (Wei et al. 2008) and safflower (Khan et al. 2009) while ecological and geographical origin playing limited role. Nonetheless variation in genetic diversity within the species is usually related with geographic range, mode of reproduction, mating system, seed dispersal and fecundity (Loveless 1992). The genetic diversity detected in the genotypes of *P. pinnata* in present study may be due to all these

prevalent background factors as the species is widely distributed in different eco-geographical regions. Similar conclusions were reached by Gupta et al. (2008) while assessing genetic variation in 14 accessions of *J. curcas* from different agro-climatic regions of India using RAPD markers. The result of the present study showed that *P. pinnata* germplasm within eastern India constitutes a broad genetic base.

RAPD marker system is thus suitable for conducting genetic diversity analysis for this tree species, where not much molecular genetics research has been done and sequence information is largely unavailable.

Conclusion

A high genetic diversity exists in the *P. pinnata* genotypes collected from different agroclimatic zone of Jharkhand indicate promising potential of the crop for genetic improvement in the region. A high variation was also found among the populations growing in the different agroclimatic regions and localities that would be useful for plant breeding and conservation purposes. The physico-chemical and molecular markers were complementary in assessing the extent of genetic variation in *P. pinnata* and the combined use of these methods helps to better understand the genetic diversity and relationship within and among *P. pinnata* populations.

There was a weak association between genetic variation of *P. pinnata* genotypes and their ecological regions of origin and the most likely important factor affecting the genetic structure of *Pongamia* in this region is human activities. Some *Pongamia* accessions collected from different eco-geographic regions of Jharkhand may have actually originated from the same population. The accessions collected from different agroclimatic zones of Jharkhand, showed significant variation in oil content, and the oil content had a strong positive correlation with protein. The results achieved from oil content and genetic structure analyses are useful as background information for *Pongamia* breeding program to formulate future breeding strategies. Desirable genotypes in terms of oil content have been identified in this study. The use of these genotypes in the breeding program would lead to improved *Pongamia* genotypes with high oil content.

References

- Acharya L., Mukherjee AK, Panda PC. 2004. Genome relationship among nine species of Millettiae (Leguminosae: Papilionoideae) based on random amplified polymorphic DNA (RAPD), *Zeitschrift fuer Naturforschung Section C. J Biosciences* 59: 868-873.
- Allen ON, Allen EK. 1981. *The Leguminosae: A Source Book of Characteristics, Uses and Nodulation*. University of Wisconsin Press, Madison, WI/Macmillan Publishing, London.
- Ayliffe MA, Lawrence GJ, Ellis JG, Pryor AJ. 1994. Heteroduplex molecules formed between allelic sequences cause non-parental RAPD bands. *Nucleic Acids Res* 22: 1632-1636.
- Azam MM, Waris A, Nahar NM. 2005. Prospects and potential of fatty acid methyl esters of some non-traditional seed oils for use as biodiesel in India. *Biomass Bioener* 29: 293-302.
- Barabás I, Todoruț IA. 2011. Biodiesel quality, standards and properties, biodiesel-quality, emissions and by-products. In: *InTech* (Montero G, ed). Available from: <http://www.intechopen.com/books/>.
- Burton JW. 1987. Quantitative genetics: Results relevant to soybean breeding. In: *Soybeans: Improvement, Production and Uses*. *Agronomy* 16: 211-242.
- Bussell JD. 1999. The distribution of random amplified polymorphic DNA (RAPD) diversity among populations of *Isotoma petraea* (Lobeliaceae). *Mol Ecol* 8: 775-789.
- Cardoso MA, Provan J, Powell W, Ferreira CG, Oliveira DE. 1998. High genetic differentiation among remnant populations of the endangered *Caesalpinia echinata* Lam. (Leguminosae-Caesalpinoideae). *Mol Ecol* 7: 601-608.
- Cox HE, Pearson D. 1962. *The Chemical Analysis of Foods*. Chemical Publishing Company Inc., New York.
- Doyle JJ, Doyle JL. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.
- Ellsworth DL, Rittenhouse KD, Honeycutt RL. 1993. Artfactual variation in randomly amplified polymorphic DNA banding patterns. *Biotechniques* 14: 214-217.
- Gupta S, Srivastava M, Mishra GP, Naik PK, Chauhan RS. 2008. Analogy of ISSR and RAPD markers for comparative analysis of genetic diversity among different *Jatropha curcas* genotypes. *Afr J Biotech* 7: 4230-4243.
- Hadrys H, Balick M, Schierwater B. 1992. Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. *Mol Ecol* 1: 55-63.
- Hedge R, Verghese M, Padmini S, Jayaraj RCS. 1962. Carbohydrates estimation in food samples. In: *Methods In Carbohydrates Chemistry* (Whistler RL, Miller JNB, eds). Academic Press, New York, pp 471-513.
- Horowitz W. 1975. *Official Methods of Analysis Of AOAC*. Association of Official Analytical Chemists, Washington.
- Janardhanan K, Vadivel V, Pugalenti M. 2003. Biodiversity in Indian underexploited/tribal pulses. In: *Improvement Strategies for Leguminosae Biotechnology* (Jaiwal PK, Singh RP, eds). The Netherlands: Kluwer Academic Publishers, pp 353-405.
- Karmee SK, Chadha A. 2005. Preparation of biodiesel from crude oil of *Pongamia pinnata*. *Biores Tech* 96: 1425-1429.
- Kaushik N, Kumar S, Kumar K, Beniwal RS, Kaushik N, Roy S. 2007. Genetic variability and association studies in pod and seed traits of *Pongamia pinnata* (L.) Pierre in Haryana, India. *Genet Resour Crop Evol* 54: 1827-1832.
- Kesari V, Madurai Sathyanarayana V, Parida A, Rangan L. 2010. Molecular marker-based characterization in candidate plus trees of *Pongamia pinnata*, a potential biodiesel legume. *AoB Plants* 2010: plq017.
- Kesari V, Sudarshan M, Das A, Rangan L. 2009. PCR amplification of the genomic DNA from the seeds of Ceylon Ironwood, *Jatropha* and *Pongamia*. *Biomass Bioener* 33: 1724-1728.
- Khan M, von Witzke-Ehbrecht S, Maass B, Becker H. 2009. Relationships among different geographical groups, agromorphology, fatty acid composition and RAPD marker diversity in Safflower (*Carthamus tinctorius*). *Genet Resour Crop Evol* 56: 19-30.
- Knothe G. 2002. Structure indices in FA chemistry, how relevant is the Iodine value. *J Am Oil Chem Soc* 79: 847-854.
- Lewontin RC. 1972. The apportionment of human diversity. *Evol Biol* 6: 381-398.
- Loveless MD. 1992. Isozyme variation in tropical trees. *New For* 6: 67-94.
- Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. 1951. Protein measurement with the Folin phenol reagent. *J Biol Chem* 193: 265-275.
- Michelmore RW, Paran I, Kesseli RV. 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc Natl Acad Sci U S A* 88: 9828-9832.
- Mukta N, Murthy IYLN, Sripal P. 2009. Variability assessment in *Pongamia pinnata* (L.) Pierre germplasm for biodiesel traits. *Indus Crops Prod* 29: 536-540.
- Murray MG, Thompson WF. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* 8: 4321-4325.
- Namkoong G. 1986. Genetics and the forests of the future. *Unasylva* 152: 2-18.
- Nei M. 1973. Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci U S A* 70: 3321-3323.
- Raheman H, Phadatare AG. 2004. Diesel engine emission and performance from blends of karanja methyl ester and diesel. *Biomass Bioener* 27: 393-397.
- Rao VR, Hodgkin T. 2002. Genetic diversity and conservation and utilization of plant genetic resources. *Plant Cell Tiss Org* 68: 1-19.
- Sadasivam S, Manickam A. 1992. *Biochemical Methods for Agricultural Sciences*. Wiley Eastern Ltd, Madras.
- Sarma AK, Konwer D, Bordoloi PK. 2005. A comprehensive anal-

- ysis of fuel properties of biodiesel from Koroch seed oil. *Energ Fuels* 19: 656-657.
- Scott PT, Pregelj L, Chen N, Hadler JS, Gresshoff PM. 2008. *Pongamia pinnata*: an untapped resource for the biofuels industry of the future. *Bioener Res* 1: 2-11.
- Sharma SS, Negi MS, Sinha P, Kumar K, Tripathi SB. 2010a. Assessment of genetic diversity of biodiesel species *Pongamia pinnata* accessions using AFLP and three endonuclease-AFLP. *Plant Mol Biol Rep* 29: 12-18.
- Sharma YC, Singh B, Korstad J. 2010b. Application of an efficient nonconventional heterogeneous catalyst for biodiesel synthesis from *Pongamia pinnata* oil. *Energ Fuels* 24: 3223-3231.
- Sharmin E, Ashraf SM, Ahmad S. 2007. Synthesis, characterization, antibacterial and corrosion protective properties of epoxies, epoxy-polyols and epoxy-polyurethane coatings from linseed and *Pongamia glabra* seed oils. *Int J Biol Macromol* 40: 407-422.
- Stankiewicz M, Gadamski G, Gawronski SW. 2001. Genetic variation and phylogenetic relationships of triazine-resistant and triazine-susceptible biotypes of *Solanum nigrum* - analysis using RAPD markers. *Weed Res* 41: 287-300.
- Sujatha K, Rajwade AV, Gupta VS, Hazra L. 2010. Assessment of *Pongamia pinnata* (L.) - a biodiesel producing tree species using ISSR markers. *Curr Sci* 99: 1327-1329.
- Thudi M, Manthena R, Wani SP, Tatikonda L, Hoisington DA, Varshney RK. 2010. Analysis of Genetic Diversity in *Pongamia* [*Pongamia pinnata* (L.)Pierrre] using AFLP Markers. *J Plant Biochem Biotech* 19: 209-216.
- Wani SP, Osman M, Emmanuel D'Silva, Sreedevi TK. 2006. Improved livelihoods and environmental protection through biodiesel plantations in Asia. *Asian Biotechnol Dev Rev* 8: 11-29.
- Wei L, Zhang H, Zheng Y, Guo W, Zhang T. 2008. Development and utilization of EST-derived microsatellites in sesame (*Sesamum indicum* L.). *Acta Agron Sin* 34: 2077- 2084.
- Wiggins BT. 2012. Heritability and genetic gain of seed protein, oil, and yield among ril of soybean. Master's Thesis, University of Tennessee, USA.
- Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* 18: 6531-6535.
- Yeh F, Boyle CT, Ronagca Y, Xian JM. 1999. Popgene version 3.1 (www.ualberta.ca/~fyeh/popgene_download.html)
- Yeh FC, Chong DK, Yang RC. 1995. RAPD variation within and among natural populations of trembling aspen (*Populus tremuloides* Michx.) from Alberta. *J Hered* 86: 454-460.