Trait selection by path and principal component analysis in *Jatropha curcas* for enhanced oil yield

Sarnam Singh^b, Aruna Prakash^b, N R Chakraborty^c, Candace Wheeler^d, P K Agarwal^{ab}, Arup Ghosh^{ab*}

^aAcademy of Scientific and Innovative Research (AcSIR) and CSIR-Central Salt & Marine Chemicals Research Institute Premises, G. B. Marg, Bhavnagar-364002, Gujarat, India.

^bDepartment of Wasteland Research, CSIR-Central Salt & Marine Chemicals Research Institute, G. B. Marg, Bhavnagar-364002, Gujarat, India. Fax: +91-278- 2567562, Tel: +91-278-2567760

^cDepartment of CIHAB, Palli-Siksha Bhavana (Institute of Agriculture), Visva-Bharati, Sriniketan-731236, West Bengal, India.

^dGlobal Energy System Intelligence Center, GM Research & Development, MC #480-106-RE3, 30500 Mound Road, Warren, MI 48090-9055, United States

*Author for correspondence:

Arup Ghosh, Tel: +912782567760 Ext 6440, Email: arupghosh@csmcni.org

Abstract

Path analysis was done on different morphological, physiological and oil quality traits of 15, six-years old candidate plus trees of *Jatropha curcas* to assess their direct and indirect effect on oil yield. Principal component analysis was subsequently done to assess the pattern of the variations in characters and segregate the traits to distinguish parental accessions for plant improvement. Total chlorophyll had a positive direct effect on oil yield and also had indirect effect through single seed weight, single fruit weight and oil content. Male to female flower ratio had negligible direct effect on oil yield but had considerable indirect effect through seed yield per plant. Seed weight, fruit weight and seed yield plant⁻¹ had a high direct effect on oil yield and also had indirect effect through oil content and thus found suitable for direct selection. The first five principal components contributed 75.5% of the total variability. The accessions JCN-14, JCN-08 and JCN-09 had very high values for the first component, while JCN-04 and JCN-05 had lower values for all these variables.

Conflict of interest statement: The authors declare that they have no conflict of interest.

Key words: *Jatropha curcas*, Selection, Path analysis, Principal component analysis, oil yield

1. Introduction

Jatropha curcas can serve as an environment friendly renewable source of liquid (*Jatropha* methyl ester or biodiesel) as well as solid biofuel (briquetted fruit shells) and can substitute to some extent the fossil fuel requirement (Ghosh et al. 2007, 2014; Fairless 2007) besides contributing towards economic upliftment of farmers, reducing soil erosion (Datta and Pandey 1993), enhancing the soil quality (Ogunwole et al. 2008, Anand et al. 2015) and providing diverse ecosystem services and carbon sequestration (Mantri et al. 2014).

Seed and oil yield are known to be quantitative traits that are governed by multiple genes and therefore direct selection may not be much effective in improving these characters. Thus, indirect selection by giving due weightage to other heritable characters having strong association with them may be more helpful in genetic improvement for seed and oil yield. Achieving high oil productivity per plant is the critical objective of *Jatropha* breeding programme. Therefore, the traits having the highest influence on oil yield needs to be identified at the very outset. Accordingly, diverse germplasm possessing other desirable traits like higher biotic (disease and pest) and abiotic (drought) stress tolerance should be bred with high oil yielding accessions to develop robust genotypes.

Path coefficient analysis was carried out to determine the nature of relationship between yield and other characters and also to analyse the direct and indirect influence of different characters on oil yield. Principal components analysis (PCA) is a mathematical technique used to categorize large number of variables into major components and assess their contribution to the total variation. This technique has been applied in *Jatropha curcas* (Zapico et al. 2011; Shabanimofrad et al. 2013; Sharma et al. 2013; Tripathi et al. 2015; Nietsche et al. 2015) in earlier studies. Accordingly, PCA was carried out to determine the relationship among traits and find the traits which varied together, thus segregating the traits to distinguish parental accessions for plant improvement.

2. Materials and Methods

Fifteen candidate plus trees were selected in their native stand based on their phenotypic assessment as also reported in our earlier study (Singh et al. 2016). Out of these, fourteen belonged to diverse locations in India (Gujarat, Odisha and Maharashtra) and one belonged to Cape Verde (Pamidimarri and Reddy 2014). Selection was based on different desirable characters related to growth, yield, oil content, tolerance to pests and diseases or adaptability to dry climate exhibited by them at their native or cultivated stands (**Table 1**). Cuttings were raised from all of these mother plants except JCN01 which was raised through seeds.

The experimental site was located at Nesvad (21° 30'29.71" N; 72° 02'11.54" E; 92m) in Bhavnagar district of Gujarat state in India. The soil texture was sandy loam with calcareous layer beneath. The soil had an initial pH of 7.3, electrical conductivity of 0.12 dS m⁻¹. It had 0.3% organic carbon, 80 kg ha⁻¹ available N, 2.9 kg ha⁻¹ available P, and 176 kg ha⁻¹ available K.

Five months old saplings of all the fifteen selected germplasm were transplanted in field during 2007 in a randomized complete block design (RCBD) layout with 3 replications having 9 plants per selected germplasm. The plants were spaced at a distance of 3m x 3m. The fertilizer was supplied at the rate of 45:30:20 kg ha⁻¹ N: P₂O₅: K₂O, respectively for the initial three years following which N dose was increased to 60 kg N ha⁻¹.

Data were recorded during 2013-14 for growth characters, yield characters, physiological characters of plants and quality characters of seed oil. The growth characters included in the study were basal stem diameter, number of

primary and secondary branches, diameter of primary and secondary branches, plant height and volume of plant canopy. The plant canopy volume was derived according to Ghezehei et al. (2009).

The yield characters that were studied included male-to-female flower ratio, length and width of fruits and seeds, single fruit weight, number of seed fruit⁻¹, single seed weight, fruit coat weight fruit⁻¹, seed weight fruit⁻¹ and seed yield plant⁻¹. 50 inflorescences were considered to determine the male-to-female ratio. Hundred mature sundried fruits of every germplasm were randomly drawn from each of the 3 replications for measurements on fruits. One hundred seeds of each germplasm were randomly selected from each of the three replications for the measurement of seed characters. Seeds were harvested individually from each plant and were expressed as average seed yield plant⁻¹.

Leaf gas exchange parameters of all the germplasm were recorded in triplicate from three randomly selected plants from each replication. The gas exchange parameters, namely, photosynthetic rate (A), transpiration rate (E), stomatal conductance (Gs), intercellular leaf CO₂ concentration (Ci) and water use efficiency (WUE = A/E) were measured using a portable Photosynthesis System (LICOR 6400-XT, USA) equipped with infrared gas analyser. The total chlorophyll, chlorophyll-a, chlorophyll-b and carotenoids were quantitatively determined by colorimetric method (Arnon 1949) on Epoch BioTek spectrophotometer (USA) using 480, 510, 645, 652, and 663 nm wavelength.

Oil from the seeds was extracted in hexane solvent by using a soxhlet apparatus (Akbar et al. 2009). Fresh acidic methanol (5%) and hexane was used for methyl esterification of *Jatropha* oil (Bojan and Durairaj 2012). The methyl esters were analysed for their fatty acids using GC–MS (Shimadzu, QP-2010).

2.1. Statistical analysis

Path co-efficient analysis was carried out according to Dewey and Lu (1959) using the Windostat version 8.6 from Indostat service Hyderabad, India. The PCA was performed using Minitab 16 software.

3. Results

The genetic variability and divergence studies were reported recently for the same set of 15 germplasm while also identifying the best gain heritable traits (Singh et al. 2016). The present study sought to determine the direct and indirect effect of various traits on oil yield and also assess the pattern of the variations in characters, segregate the traits and distinguish the different parental accessions.

3.1. Path analysis

The seed length, oil content and seed yield plant⁻¹ had a high significant ($p \leq 0.001$) positive correlation of oil yield per plant, while number of secondary branches per plant was inversely correlated with it. The diameter of primary and secondary branches, fruit length, seed width, single seed weight, seed weight fruit⁻¹, rate of photosynthesis and transpiration also had positive correlation with oil yield, while male to female flower ratio had negative

association with oil yield plant⁻¹ at p<0.01 significance level. At p<0.05, the basal stem diameter, fruit width, single fruit weight, chlorophyll b correlated positively with oil yield, whereas, WUE correlated negatively (Table 2).

More consistent interpretation of the associations between the traits and whether the effect was through the influence of other traits was determined by path coefficient analysis. Path coefficient analysis was conducted taking oil yield plant⁻¹ as the dependent variable. Genotypic correlation of traits with oil yield plant⁻¹ was partitioned into direct and indirect effects (**Table 2**). High R² (1.0003) and very low residual effect (0.0173) indicated that most of the important traits that contributed to oil yield were considered in the study.

The significant positively correlated characters which exhibited high and positive direct effects of 0.89, 0.21, 0.73 and 0.58 on oil yield plant⁻¹ were seed yield plant⁻¹, oil content, single seed weight and single fruit weight, respectively, indicating a true relationship among these characters. Therefore, a direct selection through these traits might also be more effective. The length and width of seed as well as fruit, diameter of secondary and primary branches, basal stem diameter, rate of net photosynthesis and transpiration, and chlorophyll b had significant positive correlation, but their direct effects were negative or negligible indicating that the indirect effects were the cause of the association. Seed weight, fruit weight, seed yield plant⁻¹ and oil content were positively and indirectly influenced by seed length, seed width, fruit length, fruit width, diameter of primary as well as secondary branches, chlorophyll b, photosynthetic rate and transpiration.

All the photosynthetic pigments had indirect positive effects on oil yield through seed and fruit weight, seed yield plant⁻¹ and oil content even though their correlation with oil yield was either not significant or low (p<0.05). The male to female flower ratio was significantly but negatively associated with oil yield plant⁻¹, while path analysis revealed a positive direct effect for this trait which may be due to a negative correlation getting hindered by negative indirect effects through seed yield plant⁻¹. Seed weight fruit⁻¹ was significantly and positively associated with oil yield plant⁻¹, but its direct effect was negative. This may be due to high positive indirect effects via seed and fruit weight, seeds fruit⁻¹, seed yield plant⁻¹, total chlorophyll and oil content.

3.2 Principal component analysis

The principal component analysis (PCA) was performed to assess the variation considering 33 traits. Based on factor loadings of these traits, the five components of PCA explained 75.5% of the total variation as shown in **Table 3**. The projection of characters on two principal components is shown in **Fig. 1**. The characters with highest loading on principal component and their value above or below to ± 0.18 were deemed to be important. The first principal component (PC1) was contributing 28.6% of the total variation. Fruit length and width, single seed and fruit weight, seed weight fruit⁻¹, seed length, oil yield plant⁻¹, photosynthesis rate, all photosynthesis pigments and oil content were highest loading on PC1.

The principal component second (PC2) accounted for 18.1% of total variation and the variables, number of primary and secondary branches, chlorophyll-a, palmitic acid and stearic acid had positive support whereas diameter of primary and secondary branches, plant height, plant canopy volume, seed yield, oil yield and linoleic acid had substantial negative involvement. The third principal component (PC3) explained 12.7% of total variation with positive association by water use efficiency and palmitic acid while the characters like number of primary and secondary branches, volume of plant canopy, stomatal conductance, internal leaf CO₂, transpiration rate and linoleic acid had negative loading on this component. Remaining two principal components had low participation in total variation i.e. 10.5 and 7.6%.

4. Discussions

Oil yield plant⁻¹ in *Jatropha* or other crops is a complex trait influenced directly or indirectly by the interaction of various growth, physiological, fruit and seed characters, seed yield and seed oil content. An understanding of interrelationships of traits that significantly contribute towards oil yield in plant may substantially enhance the efficiency of breeding programme. Furthermore, comprehending the relationship between *Jatropha* oil yield and its component traits is of paramount importance in breeding for selecting desirable genotypes for yield improvement programs. Correlation coefficient determines simple relations among the characters, so it doesn't always determine decisive results about determination of plant selection criteria (Cakmakci et al. 1998).

Path coefficient analysis on the other hand reveals more comprehensive information on the interrelations of complex characters by partitioning the correlation coefficient into direct and indirect effects and therefore has been used to determine the aspect of selection criteria for genetic improvement. The nature of association between oil yield and its related characters would determine the appropriate trait to be used in indirect selection for improvement in the oil yield of *Jatropha*.

The number of secondary branches plant⁻¹ was strongly and negatively correlated with oil yield but had negligible direct effect. However, it had an indirect negative effect via single fruit weight and single seed weight, seed yield plant⁻¹ and seed oil content. There was negligible direct effect of diameter of primary and diameter of secondary branches on oil yield plant⁻¹. The influence of these characters on oil yield was more indirect through single seed weight, single fruit weight and seed oil content and hence they may be considered together for indirect selection. Fruit length and fruit width, seed length and seed width had high association with oil yield as well as seed yield but had no direct effect on oil yield. However, they had indirect effect through traits like single seed and fruit weight, seeds fruit⁻¹, seed yield and oil content and hence can be considered together for indirect selection.

Fruit weight and seed weight had a high correlation and direct effect towards oil yield and in addition to this they had an indirect effect via each other as well as via oil content. Therefore, these characters may be important for

direct selection purposes. Similarly, seed yield plant⁻¹ was an important trait found suitable for direct selection. Further, we found high direct effect and indirect effects of seed yield on oil yield through oil content as well. The male to female flower ratio was inversely correlated with oil yield but had negligible direct effect on oil yield. However, oil yield was indirectly affected by male to female flower ratio via seed yield plant⁻¹. Similar high direct effect on seed yield was also previously reported by Rao et al. (2008). The total chlorophyll had a direct effect on oil yield and an indirect effect through seed weight, fruit weight and oil content and, therefore, these characters together may be pertinent for indirect selection purposes. The gas exchange parameters viz., net photosynthesis and transpiration rate were highly correlated with oil yield. The effects were primarily due to indirect effects via seed weight and fruit weight, total chlorophyll and oil content and, therefore, might be considered for indirect selection.

Principal component analysis revealed that the first principal component strongly correlated with chlorophyll-b, fruit length, total chlorophyll, photosynthesis rate, single seed weight, oil content, chlorophyll-a, seed weight fruit¹, fruit width, single fruit weight, carotenoids, seed length and oil yield plant⁻¹ in the sequential order indicating that these traits vary together in the same direction. Therefore, the first component can be varied as a measure of the quality of fruit (length, width and weight), seed (width, weight and oil content), photosynthesis and pigments (total chlorophyll, chl-a, b and carotenoids). It can, thus, be inferred that high oil yield plant⁻¹ trait is linked with yield attribute traits as well as physiological traits. Similarly, the second principal component varied as a measure of plant architecture and fatty acid profile. The third principal component showed the relationship of physiological parameters with plant morphology in respect to the number of branches and plant volume. In our study, the first five components contributed 75.5% of the total variation and our results are in conformity with other studies on *Jatropha* where first three or four components explained more than 70% of total variation (Shabanimofrad et al. 2013; Nietsche et al. 2015).

A scatter plot of the component scores has been produced in **Fig. 2** indicating the second component plotted against the first component. Looking at the germplasm out by itself to the right, we may conclude the germplasm JCN14 (IC 565735), followed by JCN08 and JCN09 (IC-565733) had very high values for the first component and it is expected that these germplasm would have high values for the characters for which are strongly correlated (**Fig. 1 and 2**).

Incidentally, as also reported earlier (Singh et al., 2016), JCN14 and JCN09 exhibited significantly higher oil content (37.49, and 37.34%, respectively) and oil yield plant⁻¹ (0.45 and 0.47 kg plant⁻¹, respectively) than almost all other germplasm. JCN14 also exhibited maximum values for some other biological characters as well, viz., diameter of primary and secondary branches, fruit and seed length, chlorophyll-a, carotenoid, total chlorophyll, and photosynthesis rate, whereas JCN09 had significantly highest basal stem diameter, plant height, fruit weight

and seed yield per plant under semi-arid condition. JCN14 (IC 565735) was earlier reported to be the highest yielder even under sub-humid climatic conditions (Ghosh et al. 2007). In the present study, JCN04 (IC-565732) followed by JCN05 (IC-565739) were located extremely left of the spectrum and, thus, had lower values for each of these variables. Interestingly, both of these germplasm had the lowest oil yields (0.16 and 0.19 kg plant⁻¹, respectively), however, they possessed desirable characters like open canopy or ability to adapt in dry climate which formed the basis of their initial selection (**Table 1**). Located at the top of the plot, the exotic JCN01 had a high value for the second component, while JCN11 and JCN02 (IC-565734) at bottom of the plot had low values for the same component. Thus, plants from such distinct group of germplasm associated with certain particular characters may be utilized as parents for breeding purpose to develop elite *Jatropha* varieties.

5. Conclusion

The study thus helped in identification of some important characters which directly or indirectly have bearing on the oil productivity of *Jatropha*. The traits selected may serve as the selection criteria for *Jatropha* oil yield improvement programme by hybridization and further recurrent selection. The study also identified diverse germplasm with distinct set of traits which may be employed as parents for breeding elite varieties.

Acknowledgement

The authors gratefully acknowledge CSIR, New Delhi, General Motors, U.S. and U.S. Department of Energy vide grant number DE-FG36-08GO18125 for financial support of this project. The authors acknowledge Springer for permitting the reproduction of Table 1 from our earlier authored article in the journal, *Trees- Structure and Function* entitled, “Genetic variability, character association and divergence studies in *Jatropha curcas* for improvement in oil yield”, DOI 10.1007/s00468-016-1354-0 (2016) authored by Sarnam Singh, Aruna Prakash, N. R. Chakraborty, Candace Wheeler, P. K. Agarwal & Arup Ghosh whose copyright is held exclusively by Springer-Verlag Berlin Heidelberg.

References

- Akbar, E., Yaakob, Z., Kamarudin, S.K., Ismail, M., Salimon, J., 2009. Characteristic and composition of *Jatropha curcas* oil seed from Malaysia and its potential as biodiesel feedstock feedstock. Eur. J. Sci. Res. 29, 396-403.
- Anand, K.V., Kubavat, D., Trivedi, K., Agarwal, P.K., Wheeler, C., Ghosh, A., 2015. Long-term application of *Jatropha* press cake promotes seed yield by enhanced soil organic carbon accumulation, microbial biomass and enzymatic activities in soils of semi-arid tropical wastelands. Eur. J. Soil. Biol. 69, 57-65.
- Arnon, D.I., 1949. Copper enzymes in isolated chloroplasts. Polyphenoloxidases in *Beta vulgaris*. Plant Physiol. 24, 1-15.

222 Bojan, S.G., Durairaj, S.K., 2012. Producing biodiesel from high free fatty acid *Jatropha curcas* oil by a two-step
 223 method—an Indian case study. J. Sustain. Energ. Environ. 3, 63-66.

224 Cakmakci, S., Unay, A., Acikgoz, E., 1998. An investigation on determination of characters regarding to seed and
 225 straw yield using different methods in common vetch. Turk. J. Agric. For. 22, 161-166.

226 Datta, S.K., Pandey, R.K., 1993. *Jatropha curcas*-a promising crop for new source of fuel. Appl. Bot. Abstr. 13,
 227 108-118.

228 Dewey, D.R., Lu, K., 1959. A correlation and path-coefficient analysis of components of crested wheatgrass seed
 229 production. Agron. J. 51, 515-518.

230 Fairless, D., 2007. Biofuel: the little shrub that could-maybe. Nature. 449, 652-655.

231 Ghezehei, S.B., Annandale, J.G., Everson, C.S., 2009. Shoot allometry of *Jatropha curcas*. Southern Forests. 71,
 232 279-286.

233 Ghosh, A., 2014. Plane truth: aviation's real impact on people and the environment. Int. J. Environ. Stud. 71, 888-
 234 898.

235 Ghosh, A., Chaudhary, D.R., Reddy, M.P., Rao, S.N., Chikara, J., Pandya, J.B. et al., 2007. Prospects for *Jatropha*
 236 methyl ester (biodiesel) in India. Int. J. Environ. Stud. 64, 659-674.

237 Mantri, V.A., Parmar, D.R., Rao, P.N., Ghosh, A., 2014. Observations on ecosystem services in *Jatropha curcas*
 238 plantations established in degraded lands in India. Int. J. Environ. Stud. 71, 209-214.

239 Nietzsche, S., Vendrame, W.A., Crane, J.H., Pereira, M.C., Costa, A., Reis, S.T., 2015. Variability in reproductive
 240 traits in *Jatropha curcas* L. accessions during early developmental stages under warm subtropical conditions.
 241 Glob. Change Biol. Bioenerg. 7, 122-134.

242 Ogunwale, J.O., Chaudhary, D.R., Ghosh, A., Daudu, K., Chikara, J., Patolia, J.S., 2008. Contribution of *Jatropha*
 243 *curcas* to soil quality improvement in a degraded Indian entisol. Acta. Agr. Scand B-S P. 58, 245-251.

244 Pamidimarri, D.S., Reddy, M.P., 2014. Phylogeography and molecular diversity analysis of *Jatropha curcas* L.
 245 and the dispersal route revealed by RAPD, AFLP and nrDNA-ITS analysis. Mol. Biol. Rep. 41, 3225-3234.

246 Rao, G.R., Korwar, G.R., Shanker, A.K., Ramakrishna, Y.S. 2008. Genetic associations, variability and diversity
 247 in seed characters, growth, reproductive phenology and yield in *Jatropha curcas* (L.) accessions. Trees. 22, 697-
 248 709.

249 Singh, S., Prakash, A., Chakraborty, N. R., Wheeler, C., Agarwal, P. K., Ghosh, A. 2016. Genetic variability,
 250 character association and divergence studies in *Jatropha curcas* for improvement in oil yield. Trees. doi:
 251 10.1007/s00468-016-1354-0.

252 Shabanimofrad, M., Rafii, M.Y., Wahab, P.M., Biabani, A.R., Latif, M.A., 2013. Phenotypic, genotypic and
 253 genetic divergence found in 48 newly collected Malaysian accessions of *Jatropha curcas* L. Ind. Crop. Prod. 42,
 254 543-551.

255 Sharma, H.K., Shukla, A., Kumar, A., Shukla, A., Choudhary, S.B., Jatothu, J.L., 2013. Variability and genetic
 256 diversity assessment in physic nut (*Jatropha curcas* L.). J. Med. Plants Res. 7, 2380-2391.

257 Tripathi, A., Mishra, D.K., Shukla, J.K., 2015. Genetic diversity and trait association between growth, yield and
 258 seed component of *Jatropha curcas* (L.) source collection from Indian sub-continent. J. Plant Breeding Crop. Sci.
 259 7, 143-156.

260 Zapico, F.L., Nival, S.K., Aguilar, C.H., Eroy, M.N., 2011. Phenotypic diversity of *Jatropha curcas* L. from
 261 diverse origins. J. Agric. Sci. Technol. 5, 215-219.

262 **Figure caption**

263 **Fig. 1 Principal component analysis loading plot of characters of *J. curcas***

264 **Fig. 2 Principal component analysis score plot of candidate plus trees of *J. curcas***

Table-1: Selection criteria for fifteen candidate plus trees (CPTs) of *J. curcas*, their source location and geographical description

Genotype code	Criteria of selection	Accession number, if any	Location/source	Latitude	Longitude
JCN01	Size and shape of seed (long) and fruit (Oblong) and difference in leaf morphology (Exotic germplasm)	EU700455	Cape Verde	-	-
JCN02	High yield ¹	IC-565734	Banaskantha, Gujarat	24°17'53.23"N	72°44'44.65"E
JCN03	High Yield ¹	IC-565730	Ranpur, Banaskantha, Gujarat	24° 17'56.74"N	72° 55'40.90"E
JCN04	Open canopy ³	IC-565732	Chikhla, Banaskantha, Gujarat	24° 19'33.71"N	72° 53'31.69"E
JCN05	Visually more branching, profuse flowering in native stand and good adaptation to dry climate in cultivated condition	IC-565739	Zanjmer, Gopnath, Bhavnagar, Gujarat	21°11'36.36"N	72° 3'46.37"E
JCN06	Visually high number of primary, secondary and tertiary branches	IC-565738	Akola, Maharashtra	20°42'31.92"N	77° 2'57.92"E
JCN07	High yield ¹ and open canopy ³	IC-565731	Panchmahal, Gujarat	22°50'54.74"N	73°55'2.28"E
JCN08	High yield ¹ and oil content ²	-	Sardar Krushinagar Agriculture University, Dantiwada, Gujarat	24°19'28.61"N	72°19'37.60"E
JCN09	High yield ¹ , oil content ² and open canopy ³ , less affected by virus and fungus	IC-565733	Shyamlaji, Gujarat	23°41'38.65"N	73°22'42.28"E
JCN10	High yield ¹	-	Udaipur, Rajasthan	24°46'21.43"N	73°43'56.39"E
JCN11	High yield ¹	-	Udaipur, Rajasthan	24°45'17.06"N	73°43'12.19"E
JCN12	Oil content ²	-	Ambaji –Khedbrahma area Banaskantha, Gujarat	24°18'41.11"N	72°57'58.64"E
JCN13	High yield ¹ in native stand and visual insect resistance to leaf miner in cultivated stand	-	Bhiloda, Sabarkantha, Gujarat	23°48'29.42"N	73°16'50.24"E
JCN14	High yield in native stand and relatively high and consistent yield and oil content in cultivated conditions (CP-9) at Odisha (sub-humid climate) and Gujarat (semi-arid climate) ⁴	IC-565735	Ganjam, Odisha	19°27'18.43"N	85° 2'3.09"E
JCN15	Open canopy ³	-	Ambalimal, Banaskantha, Gujarat	24°21'57.88"N	72°49'30.65"E

NB: Selection of plus trees from native based on: high yield¹ (>2.5 kg seed yield plant⁻¹), seed oil content² (>33%), open canopy³ (angle between main stem to primary and angle between primary to secondary branch >65°; angle of secondary to tertiary > 50°. ⁴Selection of IC-565735 was based on consistent high yield obtained among other cultivated germplasm at Odisha.

Table 2: Genotypic path analysis results showing direct (diagonal and bold lettered) and indirect effects of all the traits over oil yield plant⁻¹

	BSD	NPB	DPB	NSB	DSB	PH	VPC	FL	FW	SSW	SFW	SW/F	SPF	SFCW
BSD	-0.0088	-0.0009	-0.0033	-0.0009	-0.0007	-0.0041	-0.0039	-0.0014	-0.0033	-0.0009	0.0004	-0.0021	-0.0048	0.0032
NPB	0.0007	0.0071	-0.0048	0.0073	-0.0047	-0.0016	-0.0005	0.0023	0.0006	-0.0014	-0.0019	-0.0020	-0.0024	-0.0012
DPB	-0.0172	0.0309	-0.0459	0.0200	-0.0408	-0.0376	-0.0202	0.0004	0.0019	-0.0071	-0.0111	-0.0104	-0.0155	-0.0086
NSB	-0.0002	-0.0022	0.0009	-0.0022	0.0011	-0.0003	-0.0003	-0.0001	0.0006	0.0009	0.0007	0.0010	0.0003	0.0001
DSB	-0.0002	0.0018	-0.0024	0.0014	-0.0027	-0.0018	-0.0003	-0.0006	-0.0001	-0.0008	-0.0011	-0.0008	-0.0003	-0.0011
PH	0.0279	-0.0140	0.0494	0.0077	0.0397	0.0603	0.0381	0.0094	-0.0020	0.0023	0.0087	0.0033	0.0065	0.0125
VPC	-0.0121	0.0018	-0.0121	-0.0035	-0.0034	-0.0174	-0.0275	0.0125	0.0148	0.0079	0.0068	0.0068	-0.0050	0.0047
FL	0.0025	0.0049	-0.0001	0.0004	0.0033	0.0024	-0.0070	0.0154	0.0143	0.0101	0.0100	0.0094	-0.0026	0.0078
FW	0.0030	0.0007	-0.0003	-0.0023	0.0002	-0.0003	-0.0042	0.0073	0.0078	0.0058	0.0053	0.0062	0.0019	0.0026
SSW	0.0778	-0.1433	0.1130	-0.3131	0.2119	0.0278	-0.2121	0.4842	0.5446	0.7348	0.6600	0.7094	-0.0306	0.4031
SFW	-0.0241	-0.1529	0.1409	-0.1772	0.2357	0.0841	-0.1437	0.3803	0.3938	0.5230	0.5822	0.5156	0.0244	0.4863
SW/F	-0.2879	0.3378	-0.2731	0.5485	-0.3675	-0.0663	0.2997	-0.7361	-0.9576	-1.1656	-1.0692	-1.2073	-0.2634	-0.5845
SPF	0.1326	-0.0836	0.0821	-0.0310	0.0247	0.0263	0.0440	-0.0407	0.0593	-0.0101	0.0102	0.0530	0.2430	-0.0436
SFCW	0.1142	0.0522	-0.0593	0.0115	-0.1272	-0.0655	0.0542	-0.1600	-0.1047	-0.1731	-0.2635	-0.1528	0.0566	-0.3155
SL	0.0073	0.0014	0.0030	-0.0029	0.0023	0.0059	-0.0030	0.0154	0.0181	0.0116	0.0111	0.0113	0.0001	0.0075
SW	0.0002	-0.0001	-0.0005	0.0007	-0.0003	-0.0007	0.0006	-0.0018	-0.0029	-0.0009	-0.0012	-0.0007	0.0007	-0.0015
M:F	-0.0024	0.0019	-0.0013	0.0037	-0.0014	-0.0001	-0.0014	-0.0007	0.0006	-0.0013	-0.0008	-0.0017	-0.0018	0.0005
SYPP	0.2324	-0.2524	0.3527	-0.6104	0.3696	0.2663	0.2066	0.2752	0.2588	0.3478	0.2241	0.2965	-0.1539	0.0679
A	-0.0013	0.0104	0.0005	0.0002	0.0205	-0.0034	-0.0248	0.0435	0.0254	0.0320	0.0331	0.0292	-0.0101	0.0278
Gs	0.0025	0.0086	-0.0032	0.0100	-0.0020	0.0016	-0.0005	0.0106	0.0054	0.0045	0.0047	0.0036	-0.0029	0.0047
CI	0.0022	0.0046	-0.0024	0.0071	-0.0043	0.0021	0.0042	0.0019	0.0001	-0.0011	-0.0013	-0.0013	-0.0007	-0.0008
E	0.0114	-0.0539	0.0360	-0.0385	0.0153	0.0108	0.0019	-0.0684	-0.0260	-0.0369	-0.0361	-0.0232	0.0508	-0.0405
WUE	-0.0051	0.0342	-0.0279	0.0293	-0.0312	-0.0047	0.0246	0.0144	-0.0038	0.0005	-0.0023	-0.0082	-0.0318	0.0053
TChl	0.0275	0.0602	-0.0285	0.0031	0.0250	-0.0558	-0.1476	0.1816	0.1489	0.1018	0.0891	0.1028	0.0056	0.0459
Chl-a	-0.0039	-0.0396	0.0254	-0.0091	-0.0041	0.0347	0.0729	-0.0791	-0.0578	-0.0387	-0.0342	-0.0375	0.0050	-0.0199
Chl-b	-0.0158	-0.0025	-0.0063	0.0066	-0.0157	0.0058	0.0364	-0.0558	-0.0551	-0.0380	-0.0341	-0.0402	-0.0109	-0.0164
CRT	-0.0056	-0.0077	0.0001	-0.0016	-0.0067	0.0072	0.0298	-0.0318	-0.0260	-0.0119	-0.0128	-0.0141	-0.0078	-0.0073
OC	0.0408	-0.0776	0.0551	-0.1611	0.0959	-0.0038	-0.0573	0.1079	0.1101	0.1613	0.1253	0.1610	0.0081	0.0452
OYPP	0.2985*	-0.2722	0.3880**	-0.6964***	0.4325**	0.2720	0.1587	0.3855**	0.3660**	0.4566**	0.3022*	0.4070**	-0.1417	0.0840
Partial R²	-0.0026	-0.0019	-0.0178	0.0015	-0.0012	0.0164	-0.0044	0.0059	0.0029	0.3355	0.1760	-0.4913	-0.0344	-0.0265

	SL	SW	M:F	SYPP	A	Gs	Ci	E	WUE	TChl	Chl-a	Chl-b	CRT	OC
BSD	-0.0034	0.0005	0.0031	-0.0023	0.0002	-0.0017	-0.0018	0.0009	-0.0008	-0.0011	-0.0003	-0.0025	-0.001	-0.0017
NPB	0.0005	0.0001	0.0019	-0.0020	0.0012	0.0045	0.0031	0.0036	-0.0040	0.0019	0.0026	0.0003	0.0011	-0.0026
DPB	-0.0074	-0.0065	0.0087	-0.0182	-0.0004	0.0109	0.0106	0.0156	-0.0214	0.0058	0.0108	-0.0052	0.0001	-0.0120
NSB	0.0003	0.0005	-0.0012	0.0015	0.0000	-0.0016	-0.0015	-0.0008	0.0011	0.0000	-0.0002	0.0003	-0.0001	0.0017
DSB	-0.0003	-0.0003	0.0006	-0.0011	-0.0009	0.0004	0.0011	0.0004	-0.0014	-0.0003	-0.0001	-0.0008	-0.0004	-0.0012
PH	0.0191	0.0116	-0.0008	0.0181	-0.0034	0.0074	0.0123	-0.0062	0.0047	-0.0149	-0.0194	-0.0063	-0.0090	-0.0011
VPC	0.0044	0.0051	0.0057	-0.0064	0.0113	0.0011	-0.0109	0.0005	0.0113	0.0179	0.0186	0.0179	0.0169	0.0075
FL	0.0126	0.0080	-0.0017	0.0048	0.0112	0.0121	0.0028	0.0100	-0.0037	0.0123	0.0113	0.0154	0.0101	0.0079
FW	0.0076	0.0066	0.0007	0.0023	0.0033	0.0032	0.0000	0.0019	0.0005	0.0052	0.0042	0.0078	0.0042	0.0041
SSW	0.4550	0.1916	-0.1364	0.2878	0.3930	0.2445	-0.0798	0.2564	-0.0057	0.3300	0.2640	0.5010	0.1811	0.5636
SFW	0.3448	0.2100	-0.0688	0.1469	0.3221	0.2067	-0.0720	0.1991	0.0222	0.2290	0.1852	0.3561	0.1539	0.3470
SW/F	-0.7267	-0.2392	0.3046	-0.4031	-0.5881	-0.3252	0.1535	-0.2650	-0.1646	-0.5475	-0.4211	-0.8723	-0.3515	-0.9245
SPF	0.0007	-0.0467	-0.0655	-0.0421	-0.0410	-0.0534	-0.0172	-0.1169	0.1292	0.0060	-0.0113	0.0477	0.0391	0.0093
SFCW	-0.1261	-0.1388	-0.0241	-0.0241	-0.1464	-0.1104	0.0255	-0.1208	0.0278	-0.0639	-0.0584	-0.0931	-0.0478	-0.0678
SL	0.0188	0.0144	0.0012	0.0096	0.0064	0.0101	0.0042	0.0078	-0.0060	0.0066	0.0046	0.0117	0.0072	0.0075
SW	-0.0026	-0.0034	-0.0013	-0.0017	-0.0002	-0.0001	0.0001	-0.0006	0.0008	-0.0003	-0.0001	-0.0008	-0.0009	-0.0003
M:F	0.0004	0.0026	0.0068	-0.0025	-0.0023	-0.0016	-0.0005	-0.0022	0.0015	-0.0025	-0.0021	-0.0033	-0.0017	-0.0030
SYPP	0.4529	0.4396	-0.3268	0.8880	0.2956	0.0011	-0.2045	0.3544	-0.3113	0.0836	0.0188	0.2231	0.0006	0.5703
A	0.0204	0.0039	-0.0206	0.0199	0.0599	0.0349	-0.0192	0.0447	-0.0107	0.0497	0.0507	0.0514	0.0434	0.0424
Gs	0.0072	0.0005	-0.0031	0.0000	0.0078	0.0134	0.0081	0.0115	-0.0096	0.0074	0.0083	0.0057	0.0053	0.0023
CI	0.0023	-0.0003	-0.0007	-0.0024	-0.0033	0.0063	0.0104	0.0028	-0.0070	-0.0018	-0.0013	-0.0033	-0.0026	-0.0043
E	-0.0440	-0.0184	0.0345	-0.0422	-0.0789	-0.0905	-0.0286	-0.1057	0.0832	-0.0495	-0.0578	-0.0352	-0.0323	-0.0381
WUE	0.0191	0.0136	-0.0130	0.0210	0.0107	0.0430	0.0400	0.0471	-0.0598	-0.0050	0.0006	-0.0179	-0.0122	-0.0069
TChl	0.0798	0.0198	-0.0823	0.0213	0.1881	0.1247	-0.0399	0.1062	0.0190	0.2266	0.2282	0.2365	0.2105	0.1212
Chl-a	-0.0266	-0.0039	0.0332	-0.0023	-0.0911	-0.0670	0.0133	-0.0588	0.0012	-0.1084	-0.1076	-0.1170	-0.0964	-0.0505
Chl-b	-0.0346	-0.0131	0.0266	-0.0140	-0.0478	-0.0239	0.0176	-0.0186	-0.0167	-0.0581	-0.0606	-0.0557	-0.0601	-0.0395
CRT	-0.0186	-0.0132	0.0122	0.0000	-0.0350	-0.0193	0.0123	-0.0147	-0.0099	-0.0448	-0.0432	-0.0521	-0.0483	-0.0175
OC	0.0844	0.0191	-0.0937	0.1350	0.1488	0.0367	-0.0865	0.0759	0.0244	0.1124	0.0987	0.1491	0.0761	0.2102
OYPP	0.5398***	0.4639**	-0.4001**	0.9917***	0.4206**	0.0663	-0.2473	0.4285**	-0.3059*	0.1962	0.1233	0.3587*	0.0855	0.7238***
Partial R ²	0.0101	-0.0016	-0.0027	0.8807	0.0252	0.0009	-0.0026	-0.0453	0.0183	0.0445	-0.0133	-0.0200	-0.0041	0.1522

NB: $R^2 = 1.0003$, Residual effect = 0.0173; *, **, *** significance at 0.05, 0.01, and 0.001 probability level.

BSD-Basal stem diameter, **NPB**-Number of primary branches, **DPB**-Diameter of primary branches, **NSB**-Number of secondary branches, **DSB**-Diameter of secondary branches, **PH**-Plant height, **VPC**-Volume of plant canopy, **FL**-Fruit length, **FW**-Fruit width, **SSW**-Single seed weight, **SFW**-Single fruit weight, **SW/F**-Seed weight fruit⁻¹, **SPF**-Seeds fruit⁻¹, **SFCW**-Single fruit-coat weight, **SL**-Seed length, **SW**-Seed width, **M:F**-Male to Female flower ratio, **SYPP**-Seed yield plant⁻¹, **A**-Photosynthetic rate (Net assimilation rate), **Gs**-Stomatal conductance, **CI**-Intercellular leaf CO₂ concentration, **E**-Transpiration rate, **WUE**-Water use efficiency, **TChl**-Total chlorophyll, **Chl-a**-Chlorophyll-a, **Chl-b**-Chlorophyll-b, **CRT**-Carotenoids, **OC**-Oil content, **OYPP**-Oil yield plant⁻¹.

Table 3: Eigen vectors and eigen values of the first five principal components for different traits of *J. curcas* germplasm

Traits	PC1	PC2	PC3	PC4	PC5
Basal stem diameter (BSD)	0.050	-0.070	-0.019	0.486	0.108
Number of primary branches (NPB)	0.000	0.314	-0.223	0.065	0.046
Diameter of primary branches (DPB)	0.042	-0.323	0.081	0.117	0.098
Number of secondary branches (NSB)	-0.072	0.239	-0.229	0.036	0.262
Diameter of secondary branches (DSB)	0.111	-0.279	0.119	-0.046	0.003
Plant height (PH)	0.007	-0.251	-0.142	0.157	0.174
Volume of plant canopy (VPC)	-0.152	-0.209	-0.225	0.227	-0.073
Male to Female flower ratio(M:F)	-0.100	0.080	-0.016	-0.251	0.316
Fruit length (FL)	0.284	0.071	-0.124	0.015	0.119
Fruit width (FW)	0.241	0.033	0.013	0.089	0.259
Seed length (SL)	0.220	-0.053	-0.160	0.104	0.193
Seed width (SW)	0.128	-0.057	-0.094	-0.089	0.159
Single fruit weight (SFW)	0.241	-0.159	-0.083	-0.163	0.172
Single seed weight (SSW)	0.253	-0.128	-0.051	-0.068	0.050
Seed weight fruit ⁻¹ (SW/F)	0.247	-0.142	0.004	0.010	0.112
Seeds fruit ⁻¹ (SPF)	-0.008	-0.085	0.191	0.298	0.258
Single fruit-coat weight (SFCW)	0.160	-0.131	-0.160	-0.319	0.189
Seed yield plant ⁻¹ (SYPP)	0.152	-0.194	-0.062	0.106	-0.342
Total chlorophyll (TChl)	0.269	0.175	0.105	0.046	-0.008
Chlorophyll-a (Chl-a)	0.248	0.220	0.082	0.013	-0.031
Chlorophyll-b (Chl-b)	0.285	0.069	0.137	0.093	0.058
Carotenoids (CRT)	0.232	0.165	0.153	0.059	0.094
Photosynthetic rate (A)	0.266	0.067	0.000	-0.078	-0.152
Stomatal conductance (Gs)	0.156	0.148	-0.313	0.071	0.014
Intercellular leaf CO ₂ conc. (Ci)	-0.053	0.087	-0.354	0.161	0.102
Transpiration rate (E)	0.174	0.118	-0.311	-0.046	-0.261
Water use efficiency (WUE)	-0.012	-0.091	0.415	-0.031	0.252
Oil content (OC)	0.253	-0.122	0.109	0.026	-0.228
Oil yield plant ⁻¹ (OYPP)	0.180	-0.179	-0.042	0.118	-0.334
Palmitic acid (PA)	-0.036	0.215	0.181	0.224	-0.094
Stearic acid (SA)	0.044	0.207	0.103	0.352	0.035
Linoleic acid (LA)	-0.072	-0.286	-0.223	-0.041	0.079
Oleic acid (OA)	0.087	0.153	0.145	-0.313	-0.079
Eigen value	9.430	5.988	4.194	3.466	2.506
Proportion	0.286	0.181	0.127	0.105	0.076
Cumulative %	28.600	46.700	59.400	69.900	75.500

Fig.1

Fig. 2