ISSN 1479-2621
Plant Genetic Resources: Characterization and Utilization (2017) **15**(6); 474–487
doi:10.1017/S1479262116000186

Genetic distinctiveness of safflower cultivars of India and Mexico as revealed by SSR markers

P. Kadirvel[1]*, D. Ravi[1,2], N. Mukta[1], M.C.L. Montoya-Coronado[3], S.B. Ghuge[4], J. Singh[5], V. Singh[6], S.K. Shinde[7], S.N. Deshmukh[8], P. Yadav[1] and K.S. Varaprasad[1]

[1]*ICAR-Indian Institute of Oilseeds Research, Rajendranagar, Hyderabad-500030, Telangana, India,* [2]*Institute of Biotechnology, Professor Jayashankar Telangana State Agricultural University, Rajendranagar, Hyderabad-500030, Telangana, India,* [3]*Instituto Nacional de Investigaciones Forestales, Agricolas y Pecuarias (INIFAP), Campo Experimental Valle del Yaqui, Dr Norman E. Borlaug km 12, C.P. 85000 Cd. Obregón, Sonora, México,* [4]*Vasantrao Naik Marathwada Krishi Vidyapeeth (VNMKV), Parbhani-431402, Maharashtra, India,* [5]*Rajmata Vijayaraje Scindia Krishi Vishwa Vidyalaya (RVSKVV), Indore-452001, Madhya Pradesh, India,* [6]*Nimbkar Agricultural Research Institute (NARI), Phaltan-415537, Maharashtra, India,* [7]*Mahatma Phule Krishi Vidyapeeth (MPKV), Solapur-413001, Maharashtra, India and* [8]*Dr Panjabrao Deshmukh Krishi Vidyapeeth (Dr PDKV), Akola-444001, Maharashtra, India*

Received 7 March 2016; Accepted 29 April 2016 – First published online 29 June 2016

Abstract

Safflower is a traditional oilseed crop in the world. Its seed oil is a healthy edible oil containing high amount of unsaturated fatty acids. Genetically diverse exotic cultivars are valuable germplasm for introducing new diversity in safflower improvement programmes. In this study, we characterized safflower cultivars of India (30) and Mexico (23) comprising varieties, hybrids and advanced lines developed over 50 years for genetic distinctiveness using 38 simple sequence repeat (SSR) loci. Genetic diversity estimates across cultivar groups (total, India and Mexico) were as follows: mean number of alleles (3.2, 3.1, 2.6), expected heterozygosity (0.42, 0.37, 0.37) and polymorphism information content (0.36, 0.33, 0.32) respectively, which suggested narrow SSR allelic diversity within and between cultivar groups. However, distance-based cluster analysis (neighbour-joining tree) and model-based STRUCTURE analysis revealed that safflower cultivars of India and Mexico, with the exception of a few, form two genetically distinct groups. High level of genetic variation explained between the populations (40%) and F_{st} estimate (0.4) suggested that the cultivar groups were highly differentiated with limited gene flow supporting a strong genetic structuring. High oil (~38%) and high oleic (73–79%) contents of a subset of Mexican safflower varieties and advanced lines were confirmed in field trials in India. These exotic sources from Mexico are valuable for safflower breeding programmes in India to develop new cultivars with high oil yielding potential and high oleic acid content, which is the current market demand.

Keywords: *Carthamus tinctorious* L., cultivars, fatty acid composition, genetic structuring, oilseed, oil content

Introduction

Safflower (*Carthamus tinctorius* L.) is a traditional oilseed crop in the world. Its seeds, flowers and foliage are used for diverse human needs. Its seed oil is a healthy edible oil due to high unsaturated fatty acid content (~90%) (Li and Mündel, 1996). It originated in the Fertile Crescent region over 4000 years ago and domesticated in the Far East, India, Pakistan, the Middle East, Egypt, Sudan, Ethiopia and Europe (Chapman and Burke, 2007). Currently, it is grown commercially in 15 countries including India, Mexico, the USA and Argentina as major producers (FAOSTAT, 2012). Though safflower is a valuable crop, the cultivation is rapidly

*Corresponding author. E-mail: kadirvel.palchamy@icar.gov.in

decreasing in India due to low productivity (~800 kg/ha). Furthermore, the popular cultivars are low oil (~30%) and high linoleic (~70%) types, which are less attractive to food industry. Considering the emerging markets, concerted efforts are required in India to breed cultivars with high oil yield potential coupled with high oleic acid content and to achieve greater profitability for safflower cultivation.

Genetic diversity in the germplasm collection is crucial to achieve these goals through targeted breeding programmes. Excellent variability for seed yield related traits, oil content (13–46%) and unsaturated fatty acids (oleic or linoleic) (~20 or ~70%) have been preserved in the global safflower germplasm collection (Fernández-Martinez *et al.*, 1993; Johnson *et al.*, 1999; Usha Kiran *et al.*, 2015). However, among trait specific germplasm sources, the exotic varieties that are genetically diverse and show no genotype × environment (G × E) interaction for the target traits could be highly valuable because they could contribute unique alleles and could be readily deployed without much pre-breeding efforts (Holland, 2004). For instance, the safflower varieties of the USA were used to develop improved varieties in Mexico (Muñoz-Valenzuela *et al.*, 2007) and Morocco (Nabloussi and Boujghagh, 2006). Some of the Mexican safflower varieties are reported to be high oil (~40%) and high oleic (~70%) types (Montoya-Coronado, 2008) which may have the potential to contribute for safflower improvement in India. Understanding the genetic relatedness of the Mexican cultivars with the Indian counterparts would be helpful for their utilization in Indian breeding programmes. Simple sequence repeat (SSR) markers are considered the best among genetic markers to reveal genetic diversity and relatedness in crop germplasm and cultivar collections due to desirable properties (Powell *et al.*, 1996). Extensive information on SSR markers is available in safflower (Chapman *et al.*, 2009; Ambreen *et al.*, 2015). However, only two skeleton SSR linkage maps have so far been reported (Mayerhofer *et al.*, 2010), which are useful for selection of SSR loci for diversity studies (Usha Kiran *et al.*, 2015). To date, no high-density genetic map and reference genome sequence have been developed. In this study, our objectives were to investigate genetic distinctiveness between Indian and Mexican safflower cultivars (a set of 53 genotypes consisting of varieties, hybrids and advanced lines developed over 50 years), using SSR markers and to assess the potential of Mexican varieties and advanced lines for improvement of oil and oleic acid contents in the Indian safflower cultivars.

Materials and methods

Plant material

A panel of 53 safflower genotypes consisting of 30 Indian cultivars/advanced lines and 23 varieties/advanced lines from Mexico was used for molecular characterization. The Indian group included varieties, hybrids and advanced lines developed at breeding centres: Annigeri, University of Agriculture Sciences-Dharwad (UAS-D) (Karnataka); Phaltan, Nimbkar Agricultural Research Institute (NARI) (Maharashtra); Solapur, Mahatma Phule Krishi Vidyapeeth (MPKV) (Maharashtra); Akola, Dr Panjabrao Deshmukh Krishi Vidyapeeth (Dr PDKV) (Maharashtra); Parbhani, Vasantrao Naik Marathwada Krishi Vidyapeeth (VNMKV), (Maharashtra) and Indore, Rajmata Vijayaraje Scindia Krishi Vishwa Vidyalaya (RVSKVV) (Madhya Pradesh). The Mexican safflower varieties and advanced lines were obtained from Instituto Nacional de Investigaciones Forestales, Agricolas y Pecuarias (INIFAP), Campo Experimental Valle del Yaqui, Sonora. More details of the genotypes are provided in Table 1.

Genotyping using SSR markers

DNA was extracted using pooled leaf samples (ten plants per genotype) following the procedure described by Doyle and Doyle (1987). A set of 38 SSR primer-pairs published by Mayerhofer *et al.* (2010) were synthesized for genotyping work. These primers were derived from expressed sequence tag (EST) sequences of safflower. The repeat motif types were perfect dinucleotide (at 28 loci), trinucleotide (at one locus), hexanucleotide (at one locus) and compound (at eight loci). The primers were chosen based on genetic linkage map positions (Mayerhofer *et al.*, 2010) and random coverage on linkage groups (LG) (10 out of 12) was ensured. Distribution of SSR loci on different LG was as follows: LG1-5, LG2-2, LG3-3, LG4-6, LG5-7, LG6-5, LG7-2, LG8-3, LG9-2 and LG10-3. Details of SSR primer sequences are provided in online Supplementary Table S1. PCR conditions were as follows: 94°C for 5 min for initial denaturation, 35 cycles of 94°C for 30 s for denaturation, 55°C for 30 s for annealing and 72°C for 30 s for extension followed by 72°C for 5 min for final extension in a thermocycler (BioRad Laboratories Inc). Different annealing temperatures were maintained depending upon the requirement for a specific primer pair. The PCR products were resolved in 6% polyacrylamide gel electrophoresis (PAGE) (BioRad Laboratories Inc.) and the polymorphisms were visualized after silver staining. The polymorphic SSR alleles were scored as codominant markers using different characters such as 1 (allele number 1), 2 (allele number 2), 3 (allele number 3), etc.

Genetic diversity analysis

The genetic diversity measures namely number of alleles (N_A), gene diversity (expected heterozygosity; H_e) and polymorphism information content (PIC) were obtained using

https://doi.org/10.1017/S1479262116000186 Published online by Cambridge University Press

Table 1. List of safflower cultivars/advanced lines used in the study

Name	Origin	Year of release	Name	Origin	Year of release
A-1	Annigeri	1969	Humaya-65	Sonora	1965
A-2	Annigeri	1998	Kino-76	Sonora	1976
A-300	Annigeri	1959	Aceitera	Sonora	1976
Nira (NRS-209)	Phaltan	1986	Mante-81	Sonora	1981
NARI-6	Phaltan	2001	Sahuaripa-88	Sonora	1988
NARI-38	Phaltan	2007	Quiriego-88	Sonora	1988
NARI-52 (Advanced line)	Phaltan	–	Sanjose-89	Sonora	1989
NARI-57	Phaltan	2015	Sinaloa-90	Sonora	1990
NARI-NH-1 (Hybrid)	Phaltan	2002	San Ignacio-92	Sonora	1992
NARI-H-15 (Hybrid)	Phaltan	2006	Sonora-92	Sonora	1992
NARI-H-23 (Hybrid)	Phaltan	2014	Bacum-92	Sonora	1992
AKS-207	Akola	2007	Quilantan-97	Sonora	1997
PKV Pink (AKS-311)	Akola	2012	RC-1002-L	Sonora	2010
Bhima (S-4)	Jalgaon	1982	RC-1005-L	Sonora	2010
Phule Kusuma (JLSF-414)	Jalgaon	2003	RC-1033-L	Sonora	2010
Girna (JLSF-88)	Jalgaon	1990	Ciano-Lin	Sonora	2010
Tara	Solapur	1976	Ciano-OL	Sonora	2010
SSF-658	Solapur	2009	CC-1469 (Advanced line)	Sonora	–
SSF-708	Solapur	2010	CC-1577 Advanced line)	Sonora	–
SSF-733	Solapur	2013	CCC-B1 (Advanced line)	Sonora	–
SSF-748 (Phule Chandrabhaga)	Solapur	2008	CCC-B2 (Advanced line)	Sonora	–
Sharda (BSF-168-4)	Parbhani	1990	CCC-B3 (Advanced line)	Sonora	–
PBNS-12 (Parbhani Kusum)	Parbhani	2001	CCC-B4 (Advanced line)	Sonora	–
PBNS-40	Parbhani	2006			
JSF-1	Indore	1984			
JSF-7	Indore	1990			
JSI-73	Indore	1999			
JSI-97	Indore	2004			
JSI-99	Indore	2004			
RVS-113	Indore	2010			

the software program PowerMarker version 3.25 (Liu and Muse, 2005). The N_A refers to total number of alleles detected at a single SSR locus across individuals in the population. The H_e refers to the probability that two randomly chosen alleles from the population are different. The PIC refers to the value of a marker for detecting polymorphism within a population, which is derived from the number of detectable alleles and the distribution across individuals.

Distance-based cluster analysis

The genetic distinctness among genotypes was determined by weighted neighbour-joining (NJ) tree-based cluster analysis using dissimilarity coefficient (simple matching) as implemented in DARwin (Dissimilarity Analysis and Representation for windows) V.5.0.158 (Perrier and Jacquemoud-Collet, 2006). The statistical support for the branching points in the dendrogram was obtained by bootstrap analysis with 1000 replications.

Kinship analysis

Pairwise kinship coefficients were obtained by regression analysis (Ritland, 1996) as implemented in the statistical programme SPAGeDi 1.4c – a program for spatial pattern analysis of genetic diversity (Hardy and Vekemans, 2002). Negative values of the coefficients were considered zero.

Model based structure analysis

The genotypic data of 38 SSR loci were analysed in the STRUCTURE 2.3.3 program (Pritchard *et al.*, 2000) to detect

https://doi.org/10.1017/S1479262116000186 Published online by Cambridge University Press

the possible number of sub-populations (*K*) within the collection of 53 genotypes. Number of possible clusters was assumed to be 1–10. The mean posterior probability [*LnP* (*D*)] values per cluster (*K*) were obtained based on ten replicates. The delta *K* measure (Evanno *et al.*, 2005) was used to determine the number of populations as implemented in the online version of STRUCTURE HARVESTER (http://tayloro.biologyucla.edu/Struct_harvest) (Earl and VonHoldt, 2012). The membership coefficient of each genotype in the putative number of sub-populations was obtained using the admixture model with a burn-in period of 200,000 and replications of 500,000. The genotypes with membership coefficient more than 0.75 were assigned to the respective sub-population and less than 0.75 were assigned to the admixture group.

Analysis of molecular variance

Hierarchical analysis of molecular variance (AMOVA) partitions genetic variation among or within groups in order to understand the amount of population differentiation. The AMOVA and pairwise F_{st} estimates were obtained using the software Arlequin 3.11 with the permutations of 1000 (Excoffier *et al.*, 2005). The Arlequin creates a distance matrix and partitions the sum of squared deviations from distance matrix into hierarchical variance components, which are tested for significance using permutation tests.

Agronomic evaluation

A subset of ten Mexican varieties/advanced breeding lines, which showed more than 35% oil content in a preliminary field trial at Indian Council of Agricultural Research (ICAR)-Indian Institute of Oilseeds Research (IIOR), Hyderabad during 2012–2013 (data not shown), was evaluated for agronomic performance at three major safflower breeding centres (vertisols) in India: IIOR-Hyderabad (17.37°N, 78.48°E), RVSKVV-Indore (20.43°N, 76.54°E) and VNMKV-Parbhani (19.16°N, 76.54°E) during 2013–2014. A recently developed high oil variety NARI-57 was included as a local check. Randomized Complete Block Design (RCBD) was followed with three replications. The trial plot consisted of five rows of 5 m length with the spacing of 45 cm (between rows) × 20 cm (between plants). The field was irrigated once either by flooding or by sprinkler after seed sowing. Weather conditions during post-sowing period (October 2013–April 2014) at the locations were as follows: Hyderabad (rainfall 326 mm, maximum temperature range 29.4–37.1°C, minimum temperature range 11.4–23.2°C), Indore (rainfall 198 mm, maximum temperature range 30.3–38.1°C, minimum temperature range 12.2–20.0 °C) and Parbhani (rainfall 268 mm, maximum temperature range 31.5–39.1°C, minimum temperature range 20.7–20.8 °C). Data on days to 50% flowering, days to maturity, seed yield per plant (g) and oil content (%) were collected using the standard evaluation procedures (IBPGR, 1983). The oil content (%) was measured in nuclear magnetic resonance (NMR) spectroscopy using 20 g of pooled seed samples from each plot (Yadav and Murthy, 2016).

Data on fatty acid composition of a subset of eight Mexican varieties and advanced lines were obtained from a separate trial (RCBD with three replications) at Hyderabad using Indian varieties: A-1, Bhima and NARI-57 as checks. Oil from seed was extracted in hexane on soxhlet apparatus (Extraction unit, E-816, Buchi). Methyl esters were obtained by a two-step catalytic process according to slightly modified method of Ghadge and Raheman (2005). In the first step, the oil (100–150 mg) was treated with 2% sulphuric acid in methanol (5 ml) for 2 h at 60°C. After the reaction, the mixture was allowed to settle for an hour and methanol–water mixture that separated at the top was removed. In the second step, product at the bottom was transesterified using 2 ml of 13% methanolic KOH for 30 min at 55°C. The organic phase was extracted with hexane and washed with water till it reached neutral pH. The hexane was dried over anhydrous sodium sulphate and concentrated with nitrogen to get methyl esters.

Fatty acid composition was determined using an Agilent 7860A gas chromatograph equipped with a flame ionization detector (FID), a split injection port and an autosampler. Peak separation was performed on a DB-225 fused silica capillary column (diameter 250 µm, length 30 meter and film thickness 0.25 µm) from Agilent Technologies. The samples (0.2 µl) were injected in split mode (split ratio 1:20). The initial oven temperature was set at 160°C for 2 min, raised to 220°C (at a rate of 6°C/min) and held at 220°C for 10 min. Both inlet and detector were set to 230°C. The carrier gas was nitrogen set to a constant flow rate of 1.2 ml/min. Peak identification was performed by comparing the relative retention times with those of a commercial standard mixture of FAME (Supelco 37 Component FAME Mix). Fatty acid composition [palmitic acid (C16:0), stearic acid (C18:0), oleic acid (C18:1) and linoleic acid (C18:2)] was determined by calculating relative peak areas per cent by GC post run analysis using EZChrom elite compact software.

Data were subjected to analysis of variance (ANOVA) and mean comparison (LSD5%) using the statistical analysis package Plant Breeding Tools (PBTools) v 1.3 (IRRI, 2013).

Results

Genetic diversity estimates

Range and mean of genetic diversity measures: N_A, H_e and PIC of the total safflower cultivar group of India and

https://doi.org/10.1017/S1479262116000186 Published online by Cambridge University Press

Mexico, Indian group alone and Mexican group alone based on 38 SSR loci are presented in Table 2. All 38 loci were polymorphic in the total group, whereas 30 loci were polymorphic in the Indian group and 35 were polymorphic in the Mexican group. Overall, a total of 120 alleles were detected across 38 loci. The allele number per SSR

Table 2. Genetic diversity estimates within and between Indian and Mexican safflower cultivar groups based on 38 SSR loci

SSR locus	N_A			H_e			PIC		
	Total	India	Mexico	Total	India	Mexico	Total	India	Mexico
ct-006	2	–	2	0.08	–	0.17	0.07	–	0.15
ct-010	2	2	2	0.47	0.49	0.42	0.36	0.37	0.33
cm-022	2	2	2	0.07	0.06	0.08	0.07	0.06	0.08
ct-032	2	2	2	0.50	0.39	0.34	0.37	0.31	0.28
ct-044	2	2	2	0.37	0.14	0.50	0.30	0.13	0.33
ct-047	5	4	3	0.66	0.59	0.66	0.60	0.51	0.58
ct-050	4	3	4	0.59	0.30	0.42	0.51	0.28	0.39
ct-125	4	4	3	0.63	0.64	0.52	0.56	0.57	0.41
ct-127	2	–	2	0.04	–	0.08	0.04	–	0.08
ct-137	2	2	2	0.40	0.10	0.47	0.32	0.09	0.36
ct-138	5	3	4	0.66	0.22	0.62	0.61	0.21	0.55
ct-227	3	–	3	0.45	–	0.39	0.35	–	0.33
ct-246	4	4	2	0.67	0.49	0.45	0.60	0.45	0.35
ct-266	5	5	2	0.62	0.55	0.09	0.55	0.51	0.08
ct-297	2	–	2	0.11	–	0.23	0.10	–	0.20
ct-309	2	2	2	0.22	0.25	0.17	0.19	0.22	0.16
ct-316	2	2	2	0.16	0.15	0.16	0.14	0.14	0.15
ct-331	2	2	2	0.49	0.19	0.31	0.37	0.17	0.26
ct-337	2	2	2	0.45	0.29	0.49	0.35	0.24	0.37
ct-381	10	8	8	0.77	0.62	0.85	0.75	0.60	0.83
ct-384	4	4	4	0.60	0.49	0.65	0.55	0.42	0.58
ct-405	2	–	2	0.04	–	0.08	0.04	–	0.08
ct-419	3	3	–	0.17	0.29	–	0.16	0.26	–
ct-423	3	2	3	0.34	0.06	0.56	0.31	0.06	0.50
ct-440	3	2	2	0.51	0.36	0.08	0.40	0.29	0.08
ct-518	2	2	2	0.47	0.21	0.42	0.36	0.18	0.33
ct-520	2	2	2	0.50	0.28	0.23	0.37	0.24	0.20
ct-558	2	2	–	0.20	0.33	–	0.18	0.27	–
ct-590	3	–	3	0.11	–	0.24	0.11	–	0.23
ct-598	3	3	3	0.58	0.62	0.51	0.51	0.54	0.45
ct-599	5	5	–	0.68	0.66	–	0.63	0.61	–
ct-605	2	–	2	0.33	–	0.50	0.27	–	0.38
ct-623	4	4	3	0.56	0.34	0.64	0.50	0.32	0.56
ct-657	4	4	3	0.50	0.47	0.52	0.44	0.43	0.42
ct-788	3	3	2	0.61	0.51	0.19	0.53	0.44	0.18
ct-820	2	–	2	0.04	–	0.08	0.04	–	0.08
ct-831	5	5	3	0.66	0.56	0.63	0.60	0.52	0.56
ct-863	4	4	3	0.55	0.57	0.20	0.47	0.49	0.18
Range	2–10	2–8	2–8	0.04–0.77	0.06–0.66	0.08–0.85	0.04–0.75	0.06–0.61	0.08–0.83
Mean	3.2	3.1	2.6	0.42	0.37	0.37	0.36	0.33	0.32

N_A, number of alleles at a SSR locus; H_e, expected heterozygosity (gene diversity); PIC, polymorphism information content. '–' indicates monomorphic locus.

https://doi.org/10.1017/S1479262116000186 Published online by Cambridge University Press

locus ranged from 2 to 10 for the total group and from 2 to 8 for the Indian as well as the Mexican group with the mean of 3.2, 3.1 and 2.6, respectively. Out of 120 alleles, 23 (19%) specific to Indian group and 19 (16%) specific to Mexican group were detected at 23 loci. The expected heterozygosity (H_e) ranged from 0.04 to 0.77 for the total group, from 0.06 to 0.66 for the Indian group and from 0.08 to 0.85 for the Mexican group with the mean of 0.42, 0.37 and 0.37, respectively. The PIC values for SSR primer pair ranged from 0.04 to 0.75 for the total group, from 0.06 to 0.61 for the Indian group and from 0.08 to 0.83 for the Mexican group with the mean of 0.36, 0.33 and 0.32, respectively.

Genetic distinctness

NJ tree based on pairwise dissimilarity coefficients revealed three major clusters of cultivars. Indian cultivars grouped in two clusters: cluster-I consisted of 26 cultivars and cluster-II consisted of four cultivars (NARI-57, RVS-113, JSI-99 and NARI-6). The Mexican cultivars grouped in a single cluster, which was clearly distinct from the cluster-I of the Indian cultivars with high bootstrap support (64%). Very low bootstrap value (8%) did not support distinctness of the Indian cluster-II from both the Indian cluster-I and the Mexican cluster. Dendrogram showing the clusters of cultivars is presented in Fig. 1.

In the total group, pairwise dissimilarity coefficients ranged from 0.03 (RC-1002-L and RC-1005-L) to 0.69 (JSI-97 and Kino-76) with an average of 0.40. In Indian group, the pairwise dissimilarity coefficients ranged from 0.05 (SSF-748 and Sharda) to 0.50 (Phule Kusuma and NARI-57, PKV-Pink and NARI-57, SSF-708 and NARI-57, PKV-Pink and JSI-73) with an average of 0.31. In Mexican group, the pairwise dissimilarity coefficients ranged from 0.03 (RC-1002-L and RC-1005-L) to 0.51 (Humaya-65 and San Ignacio-92, Kino-76 and San Ignacio-92, CCC-B2 and Ciano-Lin, Quilantan-97 and Ciano-Lin) with an average of 0.32.

Kinship

Twenty-two of the Indian safflower cultivars and advanced lines had either zero or negative kinship coefficients with Mexican safflower varieties. Eight cultivars had positive kinship values: NARI-6 (range 0.004–0.15), NARI-57 (range 0.01–0.33), NARI-NH-1 (range 0.01–0.16), PKV Pink (0.12), SSF-658 (range 0.04–0.14), SSF-708 (range 0.02–0.06), JSI-99 (0.15) and RVS-113 (0.16). NARI-57 shared the maximum kinship with Ciano-Lin (0.33) and JSF-1 (0.32). Within Indian group, JSF-1 and NARI-57 had the highest kinship coefficient of 0.43. The other Indian cultivar pairs that shared substantial kinship were NARI-6 and PKV Pink (0.31), JSI-7 and JSI-73 (0.35), A-1 and A-300 (0.28), NARI-52 and NARI-57 (0.28), NARI-6 and JSI-99 (0.27), NARI-38 and SSF-733 (0.27), JSI-97 and JSI-73 (0.27), and RVS-113 and NARI-57 (0.26). Within Mexican group, RC-1002-L and RC-1005-L (0.89) and RC-1033-L and Ciano-Lin (0.81) had high kinship coefficients. The other Mexican varietal pairs that shared substantial kinship were Sahuaripa-88 and Sanjose-89 (0.35), Quiriego-88 and San Ignacio-92 (0.34), San Ignacio-92 and Bacum-92 (0.31), Quiriego-88 and Bacum-92 (0.27), Sanjose-89 and San Ignacio-92 (0.27), and Humaya-65 and Aceitera (0.27).

Population structure

The delta-K (DK) analysis of LnP(D) values showed a sharp peak at $K = 2$, suggesting two sub-populations (Indian group and Mexican group) within a collection of 53 genotypes (Fig. 2). The genotypes were assigned to specific population group based on the threshold value (≥0.75) of membership coefficients. In the Indian group, 26 cultivars had the membership coefficients more than the threshold value of 0.75, while NARI-6, NARI-57, SSF-658 and RVS-113 showed admixture. All Mexican cultivars and advanced lines had the membership coefficients more than the threshold value of 0.75, which ranged from 0.835 to 0.997. The Indian variety NARI-57 had the maximum admixture (towards Mexican group-0.619 and towards Indian group-0.381) followed by NARI-6 (towards Mexican group-0.532 and towards Indian group-0.468). The STRUCTURE plot is presented in Fig 2.

Analysis of molecular variance

Partitioning of variation through hierarchical AMOVA showed that about 40% of variation is explained between populations and 60% between individuals within the population. The pairwise F_{st} estimate between Mexican and Indian groups was 0.40.

Agronomic performance of Mexican genotypes

To check for variance homogeneity in ANOVA test, the residuals were plotted against the predicted values. This produced a band of nearly constant width centred near zero over the whole range of fitted values with no patterns for all traits in all trials. The histogram of residuals represented an approximate normal distribution. In the QQ-plot the points fell nearly on a straight diagonal line from bottom left to top right.

In ANOVA, genotypic effects were significant (at 5% probability level) for days to 50% flowering and maturity at all three locations, for seed yield per plant at Indore and Parbhani, and for oil content only at Hyderabad (Table 3). Days to 50% flowering of the Mexican genotypes ranged from 87 to 97 at Hyderabad, from 118 to 121 at Indore and from 98 to 101 at Parbhani, while the check

https://doi.org/10.1017/S1479262116000186 Published online by Cambridge University Press

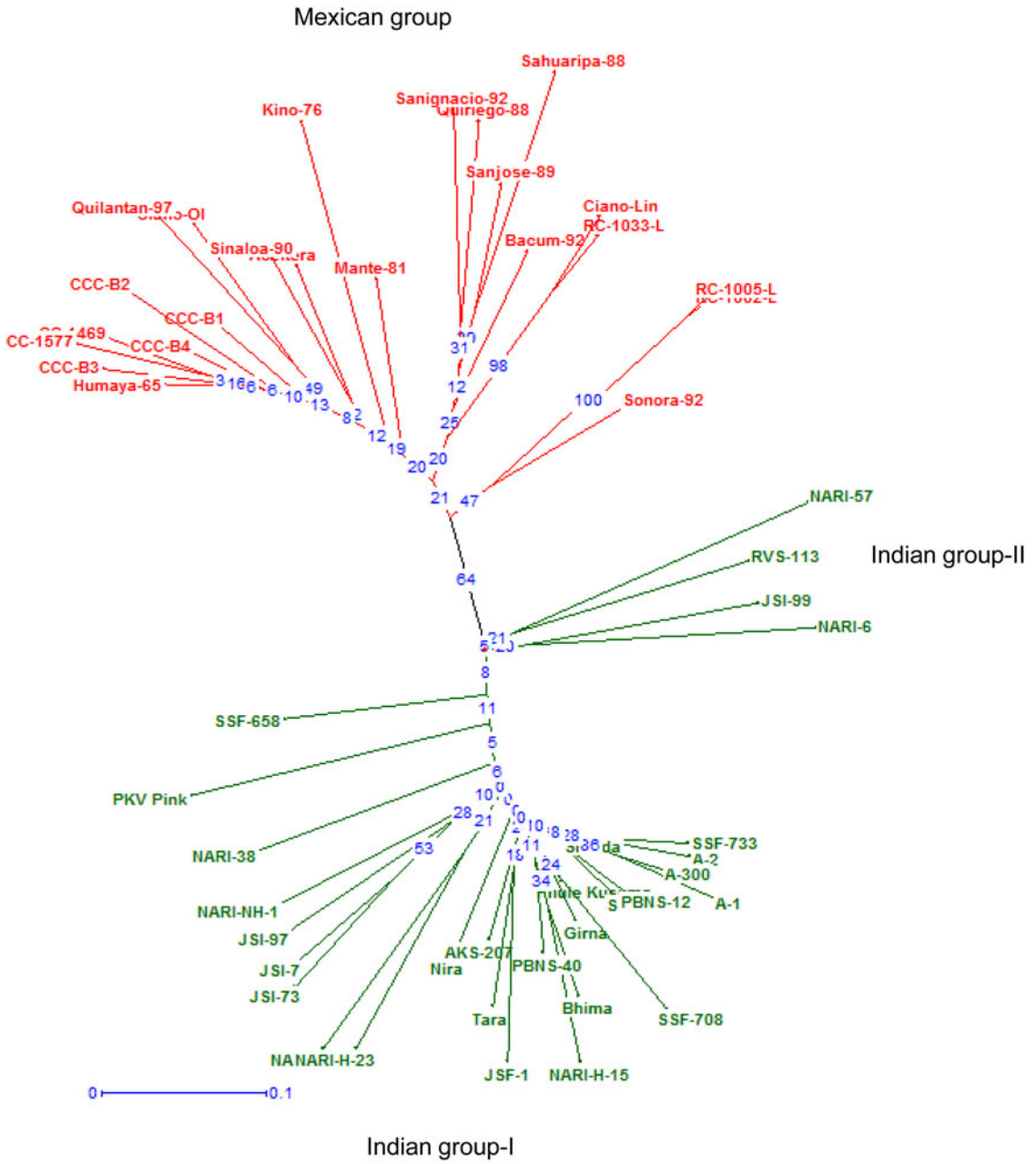

Fig. 1. Neighbour-joining tree showing genetic distinctiveness of Indian and Mexican safflower cultivars/advanced lines based on 38 SSR loci. Bootstrap percentage values are shown on the branches.

NARI-57 recorded days to 50% flowering of 88 at Hyderabad, 119 at Indore and 101 at Parbhani. Days to maturity of the Mexican genotypes ranged from 129 to 137 at Hyderabad, from 149 to 152 at Indore and from 136 to 140 at Parbhani, while the check NARI-57 recorded days to maturity of 133 at Hyderabad, 149 at Indore and 139 at Parbhani. The seed yield per plant of the Mexican genotypes ranged from 4.4 to 15.5 g at Hyderabad, from 9.7 to 25 g at Indore and from 12.5 to 16 g at Parbhani. The oil content of the Mexican genotypes ranged from 35.9 to 40% at Hyderabad, from 37 to 38.9% at Indore and from 36.3 to 38.5% at Parbhani, while the check NARI-57 recorded the oil content of 39% at Hyderabad, 38% at Indore and 37% at Parbhani.

Genotypic effects on the fatty acid composition were very large in ANOVA (Table 4). The Mexican genotypes RC-1002-L, RC-1033-L and Ciano-Lin recorded high linoleic acid content (range 75–79%) and Quilantan-97, Ciano-OL, CC-1469, CCC-B2 and CCC-B4 recorded high oleic acid content (range 73–79%). The Indian popular varieties

https://doi.org/10.1017/S1479262116000186 Published online by Cambridge University Press

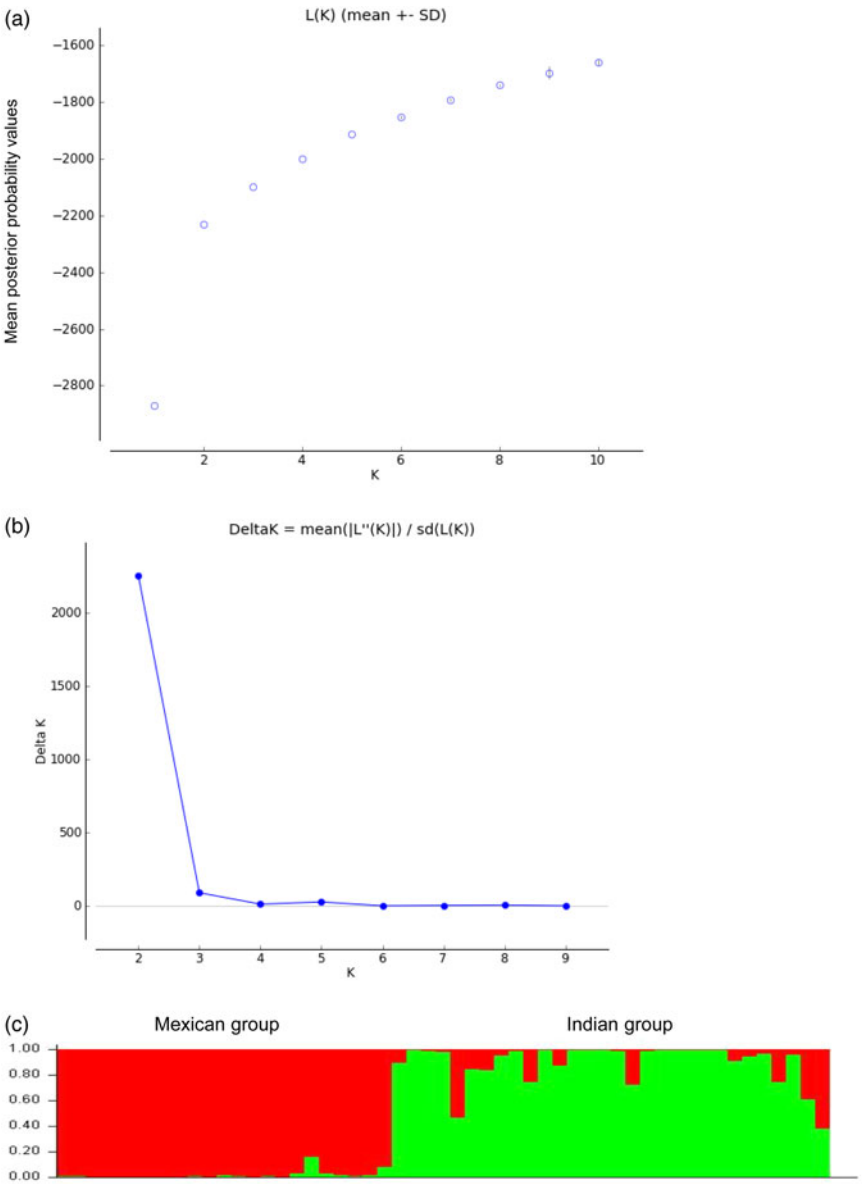

Fig. 2. Number of sub-populations (*K*) within a collection of 53 safflower cultivars/advanced lines detected in STRUCTURE as per the procedures described by (a) Pritchard *et al.* (2000) and (b) Evanno *et al.* (2005). (c) STRUCTURE Plot showing model-based clustering of safflower cultivars into two distinct groups with a few admixtures.

A-1, Bhima and NARI-57 recorded high linoleic acid content (76%). Palmitic acid content ranged from 4.4 to 6.7% in the Mexican genotypes and from 5.6 to 7.6% in the Indian varieties. Stearic acid content was the lowest, which ranged from 1.8 to 2.6% in the Mexican genotypes and from 2.4 to 3.1% in the Indian varieties. Data on fatty acid composition are presented in Table 4.

Discussion

In this study, genetic distinctiveness between safflower cultivar groups of India and Mexico was established using SSR markers. The cultivar groups represented a set of 53 genotypes comprising released varieties, hybrids and advanced lines developed over 50 years in India and Mexico, which are the major safflower breeding countries in the world. Genetic diversity in worldwide collection of cultivated safflower germplasm has been extensively investigated using molecular markers (Johnson *et al.*, 2007; Yang *et al.*, 2007; Khan *et al.*, 2009; Sehgal *et al.*, 2009; Derakhshan *et al.*, 2014; Lee *et al.*, 2014; Pearl and Burke, 2014; Kumar *et al.*, 2015). However, studies on genetic characterization of safflower cultivars from different countries are very limited. Sehgal and Raina (2005)

https://doi.org/10.1017/S1479262116000186 Published online by Cambridge University Press

Table 3. Agronomic performance of Mexican safflower genotypes at major safflower breeding centres in India

Variety/line	Days to 50% flowering			Days to maturity			Seed yield/plant (g)			Oil content (%)		
	Hyderabad	Indore	Parbhani	Hyderabad	Indore	Parbhani	Hyderabad	Indore	Parbhani	Hyderabad	Indore	Parbhani
Ciano-OL	95	120	98	131	150	136	4.40	22.33	15.17	36.41	36.95	37.20
CC-1469	90	120	100	135	149	140	15.50	18.53	13.07	37.92	38.87	36.53
CCC-B2	89	118	101	132	150	140	9.53	25.00	13.30	36.85	37.18	36.51
CCC-B4	92	119	101	129	150	140	4.96	19.60	14.00	35.85	38.59	38.52
Humaya-65	97	121	101	137	150	138	7.40	17.90	13.43	39.97	38.54	37.71
Aceitera	97	121	99	137	149	136	7.53	16.90	16.00	38.75	38.81	36.86
RC-1002-L	87	120	101	134	150	139	7.93	20.93	13.33	36.98	37.99	36.27
RC-1005-L	93	120	98	137	152	136	7.53	9.70	14.23	37.29	37.13	37.25
RC-1033-L	89	120	99	136	150	137	7.90	20.20	12.50	38.55	38.17	38.19
Ciano-Lin	93	120	98	137	151	136	8.53	18.93	14.63	37.89	38.00	38.47
NARI-57 (Check)	88	119	101	133	149	139	6.37	21.27	14.20	38.68	37.66	36.63
Mean	92	120	100	134	150	138	7.96	19.21	13.99	37.74	37.99	37.29
F value	4.386**	2.483*	8.267**	3.085*	2.811*	16.581**	1.823 (NS)	3.316*	7.162**	3.916**	1.862 (NS)	1.096 (NS)
$LSD_{0.05}$	4.6	1.4	1.4	4.6	1.4	1.4	–	6.0	1.0	1.8	–	–

NS indicates non-significant.
*indicates significance at 5% probability level.
**indicates significance at 1% probability level.

Table 4. Fatty acid composition of Mexican safflower genotypes in an evaluation trial at Hyderabad, India

Cultivar/ advanced line	Fatty acid composition (%)			
	Palmitic acid	Stearic acid	Oleic acid	Linoleic acid
Quilantan-97	4.6	2.1	78.9	14.4
Ciano-OL	5.0	1.9	74.6	18.4
CC-1469	4.6	2.2	79.2	14.0
CCC-B2	4.4	1.8	78.8	15.0
CCC-B4	5.2	1.8	72.6	20.3
RC-1002-L	6.6	2.6	15.9	74.9
RC-1033-L	6.7	2.0	12.3	78.9
Ciano-Lin	6.6	2.2	15.1	76.1
A-1 (Check)	5.6	2.4	15.6	76.4
Bhima (Check)	5.7	2.6	16.2	75.5
NARI-57 (Check)	7.6	3.1	13.4	75.9
F value	44.247**	15.055**	335.189**	332.896**
$LSD_{0.05}$	0.44	0.29	5.02	4.86

**indicates significance at 1% probability level.

analysed the efficiency of multi-locus randomly amplified polymorphic DNA (RAPD), inter-SSR and amplified fragment length polymorphic (AFLP) markers for their ability to discriminate safflower varieties and found that AFLP markers had more discriminative power. However, SSRs are considered highly informative and discriminative markers for the characterization of plant genetic diversity due to desirable properties: abundance, locus specific, multi-allelic and codominant nature. For instance, Powell *et al.* (1996) reported that SSR markers revealed highest *He* in soybean germplasm when compared with restriction fragment length polymorphic (RFLP), RAPD and AFLP markers. Use of SSR markers to map genetic diversity in safflower has become feasible with the availability of extensive information on SSR markers (Chapman *et al.*, 2009; Mayerhofer *et al.*, 2010; Hamdan *et al.*, 2011; Yamini *et al.*, 2013; Lee *et al.*, 2014; Ambreen *et al.*, 2015). For the first time, Pearl and Burke (2014) have used single nucleotide polymorphism (SNPs) markers to characterize genetic diversity in a worldwide collection of safflower germplasm, which included 25 commercial varieties from North America.

The estimates of N_A (mean range 2.6–3.2), H_e (0.37–0.42) and PIC (0.32–0.36) of SSR loci across safflower cultivar groups of India and Mexico suggested low SSR allelic diversity within each cultivar collection. However, the total cultivar group showed a slight increase in N_A, H_e and PIC than the respective country groups suggesting that both the cultivar groups have preserved specific alleles. The SSR markers have long been used to analyse the extent of diversity within the crop varietal collections. Roussel *et al.* (2005) found that the SSR allelic diversity in the European wheat varieties released over a period of 160 years (1840–2000) reflected the time and place of their development. Choudhary *et al.* (2013) compared the diversity of major Indian rice varieties released at different time periods using SSR markers and found that recently released varieties were more diverse than the older ones. In contrast, Wu *et al.* (2014) reported that genetic diversity of improved sesame varieties was lower than landraces.

The current study provided an opportunity to compare genetic diversity levels between a cultivar collection and a regional core germplasm collection in safflower. In our previous study (Usha Kiran *et al.*, 2015), we examined the genetic diversity in a sub-core collection (148 accessions) of Indian safflower germplasm (N_A 3.6, H_e 0.31, PIC 0.28) based on a set of 44 SSR loci. In this study, using 38 loci of the same set, slightly higher level of genetic diversity (H_e 0.42 and PIC 0.36) was detected in the cultivar collection of India and Mexico but with little decrease in N_A (3.2). More alleles could be detected in the sub-core collection probably due to larger sample size. Overall, it appears that genetic diversity in the safflower cultivar collection does not strikingly differ from the germplasm collections. Narrow range of N_A, H_e and PIC estimates from different germplasm collections using SSRs: 2.8–3.81, 0.31–0.39, 0.30–0.33, respectively (Hamdan *et al.*, 2011; Barati and Arzani, 2012; Derakhshan *et al.*, 2014; Lee *et al.*, 2014) and H_e (0.256) using SNPs (Pearl and Burke, 2014) indicate a similar trend.

Low genetic diversity estimates in safflower cultivar groups could perhaps be due to choice of SSRs, which were derived from conserved EST sequences. Mayerhofer *et al.* (2010) reported that these EST-SSRs were less polymorphic than genomic SSRs. However, it seems difficult to associate EST-SSRs with low genetic diversity in safflower because Hamdan *et al.* (2011) and Lee *et al.* (2014) reported low diversity estimates using genomic SSRs. Furthermore, it is important to note that type, length and position of SSRs are some of the factors that would affect SSR diversity in populations. Perfect and compound SSRs tend to reveal different pattern of genetic diversity due to differences in their evolutionary process. A longer repeat is expected to be highly polymorphic due to more chances of mutability. The SSRs located in the centromeric region would be less polymorphic due to recombination suppression (Li *et al.*, 2000). The SSRs analysed in this study were predominantly of perfect dinucleotide repeat type with the repeat length ranging from 7 to 25 (online Supplementary Table S1). Comparisons of the repeat length and number of alleles at different SSR loci indicated no relationship between them. For instance, the SSR locus ct-381 with the repeat length of (AG)14 had maximum of ten alleles, while ct-047 with the repeat length of (GA)25 had five alleles and ct-520 with the repeat length of (AG)

https://doi.org/10.1017/S1479262116000186 Published online by Cambridge University Press

24 had only two alleles. Overall, average number of alleles was 3 and 3.4 for perfect and compound SSRs, respectively. Similar results suggesting inconsistency of the relationship between repeat length and the amount of expected diversity have been reported in other studies as well (Plaschke *et al.,* 1995). The SSRs used in this study have been genetically mapped (Mayerhofer *et al.*, 2010) but their position and distance from the centromere is not available; therefore, it was not possible to verify if centromeric suppression could have caused low SSR diversity. Though choice of SSRs could influence genetic diversity estimates, narrow genetic base of the breeding programmes cannot be ruled out. However, Fu (2015) cautions that such conclusion would be risky unless a thorough study is made with careful sampling of the cultivars and high-throughput genome-wide markers.

Though SSR allelic diversity was low within Indian and Mexican safflower cultivar groups, clear distinction between them emerged from the genetic distance-based cluster analysis (NJ tree) with the exception of four Indian varieties: NARI-6, JSI-99, RVS-113 and NARI-57 (Indian group-II) (Fig. 1). Genetic relationships revealed by SSR markers showed congruence with the pedigree information of the cultivars suggesting the usefulness of the clusters. Within Indian group, NJ tree revealed strong genetic similarity of A-1 with A-300 (86% bootstrap support) and JSI-73 with JSI-7 (52% bootstrap support), which clearly reflected their pedigree. The highly popular spiny variety A-1 was developed by pedigree method of selection from the cross between A-482-1 x A-300 (Sehgal and Raina, 2005). The spineless variety JSI-73 was developed by pedigree method of selection from the cross involving JSI-7 (Saxena *et al.*, 2008). Similarly, within Mexican group, strong similarity of RC-1002-L with RC-1005-L (100% bootstrap support) and Ciano-Lin with RC-1033-L (98% bootstrap support) was observed, which is also supported by pedigree information. Ciano-Lin has been derived from the crosses involving RC Original (Borbón-Gracia *et al.*, 2011). However, no clear pattern of classification of Indian safflower cultivars was found according to the location of their origin, which suggested that the breeding centres shared a common source of germplasm and parental lines. This is expected because the Indian safflower breeding is highly networked through its coordinated research programmes on safflower since 1972 (DOR, 2006). The Indian varieties from Phaltan centre: NARI-6 and NARI-57 and Indore centre: JSI-99 and RVS-113 clustered separately but were not distinct enough from the Indian group-I and the Mexican group of cultivars as indicated by low bootstrap support, which suggested that breeding of these varieties involved exotic germplasm sources. For instance, NARI-57, a high oil variety has been derived from the cross involving Carmex, which is an American cultivar giving high yield and oil content (41%) (Gonzalez *et al.*, 1994). Similarly, JSI-99, a dwarf and extra early duration variety has been derived from the cross involving a Mexican dwarf germplasm (Saxena *et al.*, 2008). Relative kinship values among cultivar pairs also supported the genetic relationships. None of the Indian cultivars except NARI-57 had higher kinship coefficient (>0.25) with the Mexican varieties suggesting that they are highly unrelated and possibly have unique pedigree. The relative kinship estimates provide useful guidance to group genotypes according to their genetic similarity levels especially when the pedigree information is not available or incomplete.

Model-based STRUCTURE analysis also supported strong genetic structuring between Indian and Mexican safflower cultivar groups as two distinct populations with the exception of a few admixed Indian cultivars NARI-6, NARI-57, NARI-H-23, JSI-99 and RVS-113. High proportion of genetic variance (40%) explained between the populations and the high pairwise F_{st} estimate (0.40) in AMOVA suggested that Indian and Mexican safflower cultivar groups are highly differentiated and the gene flow between them would have been very limited. Therefore, the Mexican safflower varieties could be valuable sources to generate more useful diversity in Indian safflower breeding programmes. The admixed cultivars may carry mixed ancestry involving exotic and indigenous germplasm sources. This is expected because the global safflower germplasm collection of the United States Department of Agriculture (USDA) has been the common source for safflower breeders across the world (Li and Mündel, 1996). Contrary to the cultivar groups, the safflower germplasm collections did not show strong genetic structuring as per the geographical origin (Johnson *et al.*, 2007; Khan *et al.*, 2009; Chapman *et al.*, 2010; Majidi and Zadhoush, 2014; Usha Kiran *et al.*, 2015), which suggests that the regional breeding programmes would accumulate unique alleles. This observation is supported by the results of this study that 19% of the SSR alleles were found to be specifically present in the Indian group and 16% were present in the Mexican group.

The results of field evaluation indicated a good agronomic potential of the Mexican safflower varieties for the Indian situations. The flowering duration of the Mexican safflower varieties was similar with the Indian check variety NARI-57 at all three locations. The cultivars reached 50% flowering at about 92 days after sowing at Hyderabad, 100 days at Parbhani and 120 days at Indore. Similarly, the cultivars reached maturity at about 134 days after sowing at Hyderabad, 138 days at Parbhani and 150 days at Indore. Overall, late flowering and maturity of about 2 to 3 weeks was observed at Indore, which corresponded with Sonora. Days to 50% flowering of Ciano-Lin, RC-1005-L, RC-1002-L and RC-1033-L at Sonora were 121, 117, 125 and 121 (Montoya-Coronado, 2008).

Mean seed yield per plant of the Mexican safflower genotypes was highest at Indore followed by Parbhani and the

https://doi.org/10.1017/S1479262116000186 Published online by Cambridge University Press

least at Hyderabad. Large variation in seed yield performance of genotypes among locations indicate severe influence of G × E interactions. Use of RCBD when number of entries exceeded 10 would have been a limitation in this study because maintaining homogeneity conditions within the block in such situations is very difficult. However, RCBD is still considered effective when the number of treatments is up to 20 (Casler, 2015). Therefore, the results presented in this study are only indicative and a thorough field evaluation at different locations is warranted to obtain more dependable data on the seed yield performance of the Mexican safflower varieties in India. Interestingly, mean seed oil content (∼38%) of the Mexican safflower genotypes did not differ across three Indian locations and was also comparable with the Indian high oil check variety NARI-57 (38%). However, those genotypes were reported to be showing about 3% higher oil content at Sonora. The oil content of Ciano-Lin, RC-1002-L, RC-1005-L and RC-1033-L at Sonora were 41.5, 40.5, 41.9 and 40.8%, respectively (Montoya-Coronado, 2008). Discrepancies in seed oil content of safflower genotypes across countries have been observed by other researchers as well. For instance, safflower varieties of the USA namely Centennial and Montola-2000 recorded 42.3 and 38.3%, respectively at Sidney, Montana (Armah-Agyeman *et al.,* 2002) while the same varieties recorded 29 and 35.2%, respectively in Turkey (Arslan, 2007). This is expected because the seed oil content in safflower is reported to be a quantitative trait, low heritable (Golkar *et al.*, 2011) and influenced by environments (Coşge *et al.,* 2007; Ashrafi and Razmjoo 2010; Yeilaghi *et al.*, 2012). Nevertheless, it is encouraging to note that the oil content of the Mexican genotypes did not vary in three important locations within India but only Indore had the potential for obtaining high seed yield. Contrary to oil content, distinct genetic variation exists for fatty acid composition particularly linoleic acid and oleic acid content in safflower (Fernández-Martinez *et al.,* 1993) and is highly heritable (Golkar *et al.,* 2011). The present study also found that fatty acid composition of the Mexican safflower varieties is stable under Indian conditions. High linoleic acid content in Ciano-Lin, RC-1002-L, RC-1005-L and RC-1033-L (75–79%) and high oleic acid content in Ciano-OL (75%) at Hyderabad was highly comparable with the data from Sonora (Montoya-Coronado, 2008).

Genetic enhancement of safflower for oil yield and quality requires simultaneous improvements in seed yield, oil content and fatty acids. Though the fatty acid content can be easily manipulated in breeding programmes, it might be a great challenge to combine seed yield and oil content to achieve desirable oil yield due to genetic complexities (Rao *et al.,* 1977; Golkar *et al.,* 2011). Therefore, it would be essential to know the combining ability of the high oil genotypes with the high seed yielding cultivars to exploit them in breeding programmes to improve oil yield potential (Golkar *et al.,* 2011).

Breaking the low productivity and profitability scenario of the safflower cultivation in India would warrant diverse strategies including broadening the genetic base of the breeding programmes. Concerted efforts are required to develop safflower cultivars with resistance to wilt, *Alternaria* blight and aphid, tolerance to moisture stress, high oil content (5–8% more than the current levels) and high nutritional quality with increased oleic acid and antioxidants (tocopherol) (Nimbkar, 2008; Anjani, 2012). Currently, most of the released safflower cultivars in India possess low oil content of about 30% except HUS-305 (∼36%) and NARI-57 (∼38%) (DOR 2006, 2014). Furthermore, all the Indian safflower cultivars are high linoleic types (∼75%). High linoleic (polyunsaturated fatty acid) oil is considered healthy as it reduces blood cholesterol level but its poor shelf life makes it unsuitable for frying purpose. High oleic (monounsaturated fatty acid) oil (∼80%) is naturally more stable and is highly preferred by the food industry. Till date, no high oleic safflower cultivar has been released for cultivation in India. The high oil and high oleic safflower varieties and advanced lines of Mexico that are genetically distinct and stable across different locations in India, as reported in this study, could contribute for achieving the breeding goals for development of safflower cultivars with these specific traits in India.

Supplementary material

The supplementary material for this article can be found at http://dx.doi.org/10.1017/S1479262116000186.

Acknowledgements

The authors thank Dr Senthilvel Senapathy, IIOR and Anitha Raman, IRRI for statistical analysis support and critical discussions; Dr Rajesh S. Patil, Plant Breeder, UAS-D for sharing seeds of the safflower varieties developed at Annigeri; Mr. P. Kishore for technical assistance; ICAR for financial support.

References

Ambreen H, Kumar S, Variath MT, Joshi G, Bali S, Agarwal M, Kumar A, Jagannath A and Goel S (2015) Development of genomic microsatellite markers in *Carthamus tinctorious* L. (Safflower) using next generation sequencing and assessment of their cross species transferability and utility for diversity analysis. *PLoS ONE* 10: e0135443. doi: 10.1371/journal.pone.135443.

Anjani K (2012) Genetic improvement in safflower: possible avenues. In: Murthy IYLN, Basappa H, Varaprasad KS, Padmavathi P (eds) *Safflower Research and Development in the World: Status and Strategies*. Hyderabad, India: Indian Society of Oilseeds Research, pp. 1–26.

Armah-Agyeman G, Loiland J, Karow R and Hang AN (2002) *Safflower*. Oregon State University Extension Service, EM872. Corvallis, Oregon, USA. 7 p.

Arslan B (2007) The determination of oil content and fatty acid compositions of domestic and exotic safflower (*Carthamus tinctorius* L.) genotypes and their interactions. *Journal of Agronomy* 6: 415–420.

Ashrafi E and Razmjoo K (2010) Effect of irrigation regimes on oil content and composition of safflower (*Carthamus tinctorius L.*) cultivars. *Journal of the American Oil Chemists Society* 87: 499–506.

Barati M and Arzani A (2012) Genetic diversity revealed by EST-SSR markers in cultivated and wild safflower. *Biochemical Systematics and Ecology* 44: 117–123.

Borbón-Gracia A, Ochoa-Espinoza XM, Montoya-Coronado L, Márquez JP and Camarena MGG (2011) CIANO-LIN: a new cultivar of linoleic safflower. *Revista Mexicana de Ciencias Agrícolas* 2: 791–794.

Casler MD (2015) Fundamentals of experimental design: guidelines for designing successful experiments. *Agronomy Journal* 107: 692–705.

Chapman MA and Burke JM (2007) DNA sequence diversity and the origin of cultivated safflower (*Carthamus tinctorius* L.; Asteraceae). *BMC Plant Biology* 7: 60. doi: 10.1186/1471-2229-7-60.

Chapman MA, Hvala J, Strever J, Matvienko M, Kozik A, Michelmore RW, Tang S, Knapp SJ and Burke JM (2009) Development, polymorphism, and cross-taxon utility of EST–SSR markers from safflower (*Carthamus tinctorius* L.). *Theoretical and Applied Genetics* 120: 85–91.

Chapman MA, Hvala J, Strever J and Burke JM (2010) Population genetic analysis of safflower (*Carthamus tinctorius*; Asteraceae) reveals a near eastern origin and five centres of diversity. *American Journal of Botany* 97: 831–840.

Choudhary G, Ranjitkumar N, Surapaneni M, Deborah DA, Vipparla A, Anuradha G, Siddiq EA and Vemireddy LR (2013) Molecular genetic diversity of major Indian rice cultivars over decadal periods. *PLoS ONE* 8: e66197. doi: 10.1371/journal.pone.0066197.

Coşge B, Gürbüz B and Kiralan M (2007) Oil content and fatty acid composition of some safflower (*Carthamus tinctorius* L.) varieties sown in spring and winter. *International Journal of Natural and Engineering Sciences* 1: 11–15.

Derakhshan E, Majidi MM, Sharafi Y and Mirlohi A (2014) Discrimination and genetic diversity of cultivated and wild safflowers (*Carthamus* spp.) using EST-microsatellites markers. *Biochemical Systematics and Ecology* 54: 130–136.

DOR (2006) *Research Achievements in Safflower*. Hyderabad, India: All India Coordinated Research Project on Safflower, Directorate of Oilseeds Research, 111 p.

DOR (2014) *Annual Report, Safflower, 2013–14*. Hyderabad, India: All India Coordinated Research Project on Safflower, Directorate of Oilseeds Research, 146 p.

Doyle JJ and Doyle JN (1987) A rapid DNA isolation protocol for small quantity of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

Earl DA and vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361.

Evanno G, Regnaut S and Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611–2620.

Excoffier L, Laval G and Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* 1: 47–50.

FAOSTAT (2012) http://faostat.fao.org

Fernández-Martinez J, del Rio M and de Haro A (1993) Survey of safflower (*Carthamus tinctorius* L.) germplasm for variants in fatty acid composition and other seed characters. *Euphytica* 69: 115–122.

Fu Y-B (2015) Understanding crop genetic diversity under modern plant breeding. *Theoretical and Applied Genetics* 128: 2131–2142.

Ghadge SV and Raheman H (2005) Biodiesel production from mahua (*Madhuca indica*) oil having high free fatty acids. *Biomass and Bioenergy* 28: 601–605.

Golkar P, Arzani A and Rezaei AM (2011) Genetic analysis of oil content and fatty acid composition in safflower (*Carthamus tinctorius* L.). *Journal of the American Oil Chemists Society* 88: 975–982.

Gonzalez JL, Schneiter AA, Riveland NR and Johnson BL (1994) Response of hybrid and open-pollinated safflower to plant population. *Agronomy Journal* 86: 1070–1073.

Hamdan YAS, Garcia-Moreno MJ, Redondo-Nevado J, Velasco L and Perez-Vich B (2011) Development and characterization of genomic microsatellite markers in safflower (*Carthamus tinctorius* L.). *Plant Breeding* 130: 237–241.

Hardy OJ and Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* 2: 618–620.

Holland JB (2004) *Breeding: Incorporation of Exotic Germplasm*. New York: Encyclopaedia of Plant and Crop Science Marcel Dekker, Inc., pp. 222–224. doi: 10.1081/e-epcs 120010536.

IBPGR (1983) *Safflower Descriptors*. Rome: International Board for Plant Genetic Resources, (IBPGR), p. 22.

IRRI (2013) *Plant Breeding Tools (PBTools) Version 1.3*. Manila: International Rice Research Institute.

Johnson RC, Bergman JW and Flynn CR (1999) Oil and meal characteristics of core and non-core safflower accessions from the USDA collection. *Genetic Resources and Crop Evolution* 46: 611–618.

Johnson RC, Kisha TJ and Evans MA (2007) Characterizing safflower germplasm with AFLP molecular markers. *Crop Science* 47: 1728–1736.

Khan MA, Witzke-Ehbrecht SV, Maass BL and Becker HC (2009) Relationships among different geographical groups, agromorphology, fatty acid composition and RAPD marker diversity in safflower (*Carthamus tinctorius*). *Genetic Resources and Crop Evolution* 56: 19–30.

Kumar S, Ambreen H, Murali TV, Bali S, Agarwal M, Kumar A, Goel S and Jagannath A (2015) Assessment of genetic diversity and population structure in a global reference collection of 531 accessions of *Carthamus tinctorius* L. (safflower) using AFLP markers. *Plant Molecular Biology Reporter* 33: 1299-1313.

Lee G-A, Sung J-S, Lee SY, Chung JW, Yi JY, Kim YY and Lee M-C (2014) Genetic assessment of safflower (*Carthamus tinctorius* L.) collection with microsatellite markers acquired *via* pyrosequencing method. *Molecular Ecology Resources* 14: 69–78.

Li D and Mündel H-H (1996) *Safflower. Carthamus tinctorius* L. *Promoting the conservation and use of underutilized and neglected crops*. 7. Rome: Institute of Plant Genetics and Crop Plant Research, Gatersleben/International Plant Genetic Resources Institute.

Li Y, Fahima T, Korol AB, Peng J, Röder MS, Kirzhner V, Beiles A and Nevo E (2000) Microsatellite diversity correlated with ecological-edaphic and genetic factors in three microsites of wild emmer wheat in North Israel. *Molecular Biology and Evolution* 17: 851–862.

Liu K and Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21: 2128–2129.

Majidi MM and Zadhoush S (2014) Molecular and morphological variation in a world-wide collection of safflower. *Crop Science* 54: 2109–2119.

Mayerhofer R, Archibald C, Bowles V and Good AG (2010) Development of molecular markers and linkage maps for the *Carthamus* species *C. tinctorius* and *C. oxyacanthus*. *Genome* 53: 266–276.

Montoya-Coronado MCL (2008) Mexican safflower varieties with high tolerance to *Ramularia carthami*. In: Knights SE and Potter TD (eds) *Safflower: Unexploited Potential and World Adaptability*. Proceedings of the 7th International Safflower Conference, New South Wales: Australia, Wagga Wagga.

Muñoz-Valenzuela S, Musa GC, Montoya-Coronado L and Rivera-Rojas VM (2007) Evaluation of safflower genotypes in Northwest Mexico. In: Janick J and Whipkey A (eds) *Issues in new crops and new uses*. Alexandria, VA: ASHS Press, pp. 193–195.

Nabloussi A and Boujghagh M (2006) Safflower breeding in Morocco: history and current situation. *Sesame and Safflower Newsletter* 21: 84–87.

Nimbkar N (2008) Issues in safflower production in India. In: Knights SE and Potter TD (eds) *Safflower: Unexploited Potential and World Adaptability*. Proceedings of the 7th International Safflower Conference, New South Wales, Australia: Wagga Wagga.

Perrier X and Jacquemoud-Collet JP (2006) DARwin software. Available at http://darwin.cirad.fr/darwin

Pearl SA and Burke JM (2014) Genetic diversity in *Carthamus tinctorius* (Asteraceae; safflower), an underutilized oilseed crop. *American Journal of Botany* 101: 1640–1650.

Plaschke J, Ganal MW and Röder MS (1995) Detection of genetic diversity in closely related bread wheat using microsatellite markers. *Theoretical and Applied Genetics* 91: 1001–1007.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S and Rafalski A (1996) The comparison of RFLP, RAPD, AFLP, and SSR (microsatellite) marker for germplasm analysis. *Molecular Breeding* 2: 225–238.

Pritchard JK, Stephens M and Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

Rao VR, Ramachandram M and Arunachalam V (1977) An analysis of association of components of yield and oil in safflower (*Carthamus tinctorius* L.). *Theoretical and Applied Genetics* 50: 185–191.

Ritland K (1996) Estimators for pairwise relatedness and individual inbreeding coefficients. *Genetics Research*, Cambridge, 67: 175–185.

Roussel V, Leisova L, Exbrayat F, Stehno Z and Balfourier F (2005) SSR allelic diversity changes in 480 European bread wheat varieties released from 1840 to 2000. *Theoretical and Applied Genetics* 111: 162–170.

Saxena MK, Singh J, Deshpande SL and Choudhari R (2008) Two decades of safflower in Madhya Pradesh from 1984–2004. In: Knights SE and Potter TD (eds) *Safflower: Unexploited Potential and World Adaptability*. Proceedings of the 7th International Safflower Conference, New South Wales, Australia: Wagga Wagga.

Sehgal D and Raina SN (2005) Genotyping safflower (*Carthamus tinctorius*) cultivars by DNA fingerprints. *Euphytica* 146: 67–76.

Sehgal D, Rajpal VR, Raina SN, Sasanuma T and Sasakuma T (2009) Assaying polymorphism at DNA level for genetic diversity diagnostics of the safflower (*Carthamus tinctorius* L.) world germplasm sources. *Genetica* 135: 457–470.

Usha Kiran B, Mukta N, Kadirvel P, Alivelu K, Senthilvel S, Kishore P and Varaprasad KS (2015) Genetic diversity of safflower (*Carthamus tinctorius* L.) germplasm as revealed by SSR markers. *Plant Genetic Resources* available on CJO2015. doi:10.1017/S1479262115000295.

Wu K, Yang M, Liu H, Tao Y, Mei J and Zhao Y (2014) Genetic analysis and molecular characterization of Chinese sesame (*Sesamum indicum* L.) cultivars using Insertion-Deletion (InDel) and Simple Sequence Repeat (SSR) markers. *BMC Genetics* 15: 35.

Yadav P, Murthy IYLN (2016) Calibration of NMR Spectroscopy for accurate estimation of oil content in sunflower, safflower and castor seeds. *Current Science* 110: 73–76.

Yamini KN, Ramesh K, Naresh V, Rajendrakumar P, Anjani K and Dinesh Kumar V (2013) Development of EST-SSR markers and their utility in revealing cryptic diversity in safflower (*Carthamus tinctorius* L.). *Journal of Plant Biochemistry and Biotechnology* 22: 90–102.

Yang Y-X, Wu W, Zheng Y-L, Chen L, Liu R-J and Huang C-Y (2007) Genetic diversity and relationships among safflower (*Carthamus tinctorius* L.) analysed by inter-simple sequence repeats (ISSRs). *Genetic Resources and Crop Evolution* 54: 1043–1051.

Yeilaghi H, Arzani A, Ghaderian M, Fotovat R and Feizi M, Pourdad SS (2012) Effect of salinity on seed oil content and fatty acid composition of safflower (*Carthamus tinctorius* L.) genotypes. *Food Chemistry* 130: 618–625.