Genetic Resources and Crop Evolution (2005) 52: 821–830
DOI 10.1007/s10722-003-6111-8

Development of core collection using geographic information and morphological descriptors in safflower (*Carthamus tinctorius* L.) germplasm

Sangam Lal Dwivedi[1,*], Hari Deo Upadhyaya[2] and Dattatray Mahabaleshwar Hegde[1]
[1]*Directorate of Oilseeds Research, Rajendra nagar, Hyderabad 500030, India;* [2]*International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, India; *Author for correspondence*

Received 28 May 2003; accepted in revised form 7 November 2003

Key words: Carthamus tinctorius, Core collection, Genetic resources, Safflower, Shannon-Weaver diversity index

Abstract

Safflower (*Carthamus tinctorius* L.) ranks eighth among the major oilseeds crop grown worldwide. The leaves, flower, and seeds have medicinal and industrial significance. Its seed has the best quality of edible oil. The development of a core collection could facilitate easier access to safflower genetic resources for their use in crop-improvement programs and simplify the genebank management. The present study was initiated to develop a core subset of safflower based on 12 morphological descriptors and geographic information on 5522 safflower accessions. The accessions were stratified by country of origin, and data on 12 descriptors were used for clustering following Ward's method. About 10% of the accessions were randomly selected from each of the 25 clusters to constitute a core subset of 570 accessions. Mean comparisons using *t*-test, frequency distribution using χ^2-test, and Shannon-Weaver diversity index of 12 descriptors indicated that the genetic variation available for these traits in the entire collection has been preserved in the core subset. There was a fair degree of similarity in phenotypic correlation coefficients among traits in the entire collection and core subset, suggesting that this core subset has preserved most of the co-adapted gene complexes controlling these associations. This core subset, provides an opportunity to evaluate agronomic and seed quality traits and resistance to abiotic and biotic stresses to identify diverse germplasm with beneficial traits for enhancing the genetic potential of safflower.

Introduction

Safflower (*Carthamus tinctorius* L.) ranks eighth after soybean, groundnut, rapeseed, sunflower, sesame, linseed, and castor crops grown world-wide. India, Mexico, USA, Ethiopia, Argentina, and Australia together account for 99% and 87% of the world safflower area and production, respectively (Damodaram and Hegde 2002). The world average yield of safflower is much lower (0.72 t ha^{-1}) than those reported for soybean (2.34 t ha^{-1}), rapeseed (1.51 t ha^{-1}), groundnut (1.37 t ha^{-1}), and sunflower (1.14 t ha^{-1}). Nutritionally, the safflower oil is similar to olive oil. The seeds contain 30% oil, 20% protein, and 35% crude fibre. The seeds are also a rich source of minerals (Zn, Cu, Mn, and Fe), vitamins (Thiamine and B-carotene), and the tocopherols (alpha, beta, and gamma) (Nagaraj 2001). Safflower leaves, petals, and seeds have tremendous medicinal and therapeutic significance, and petals are also used for extracting dye for coloring cloths and foodstuffs (Danisova and Sarooka 1994; Varma et al. 1997; Rudometova et al. 2001; Zhaomu and Lijie 2001). The green safflower fodder is highly palatable and comparable to other forage legumes, grasses, and cereals in crude protein and total digestible nutrients (Lachover and Kostrinski

1965). The relative feed value (RFV) of safflower forage from the bud, bloom, and seed fill growth stages are generally above the standard value (RFV = 100) for full bloom alfalfa forage (Wichman et al. 2001).

Frankel (1984) proposed "core collection" which would represent, with a minimum of repetitiveness, the genetic diversity of a crop species and its relatives. Frankel and Brown (1984) and Brown (1989a) outlined the procedure for the development of core collection by using information on the origin and characteristics of the accessions. The issues that should be taken into consideration while developing a core are the size, the sampling strategy, the grouping within the collection, and the number of accessions to be included in the core from each group. The core collection should be about 10% of the total collection that will retain over 70% of the alleles in the whole collection (Brown 1989a). Using the stratified sampling, the collection is first divided into non-overlapping groups or strata, and then a simple random sample is drawn from within each group. Passport and characterization data may be used to determine the groups within the germplasm collection. The hierarchy of groupings begins with the groupings suggested by taxonomy followed by assigning accessions to major geographic groups or agro-ecological regions. Clustering within the broad geographic group could be done based on information from available genetic diversity, cytological variation, marker loci or quantitative traits, and data on stress tolerances. Collection with abundant discriminating data of this type will require a multivariate clustering to discern groups of similar accessions (Spagnoletti Zeuli and Qualset 1987). The number of accessions selected from each cluster will depend on the strategy used. A good core collection should have maximum genetic diversity and no genotypically redundant entries, should represent the whole collection, and should be small enough to manage easily (Brown 1989b).

Core collection has been developed in many crops (Erskine and Muehlbauer 1991; Tohme et al. 1995; Knüpffer and van Hintum 1995; Cordeiro et al. 1995; Bisht et al. 1998; Huamán et al. 1999; Upadhyaya et al. 2001, 2002, 2003). Johnson et al. (1993) were the first to develop a core subset of 210 safflower accessions based on branching pattern, flower color, flowering time, growth habit, head diameter, plant height, iodine number, lysine content, oil content, and spineless on 2000 accessions from 50 countries. Further studies revealed that this core subset capture a large fraction of the diversity in oil and meal characteristics (Johnson et al. 1999). Suresh and Balakrishnan (2001) and Balakrishnan and Suresh (2000, 2001a, b) proposed several strategies and sampling methods to develop core collection using geographic origins and 28 descriptors on 3250 safflower accessions from 32 countries. There are 25,179 accessions of safflower germplasm conserved in 22 genebanks of 15 countries in the world (http://www.ipgri.cgiar.org/regions/apo/safflower.html). It is not known how many of these accessions are common across the genebanks. The Safflower Project Coordinating Unit at Solapur, Maharashtra, India characterized over 7000 safflower accessions from 40 countries for few morphological and agronomic descriptors. The present study was therefore aimed to develop a core subset of safflower based on geographic distribution and 12 morphological descriptors on 5522 safflower accessions.

Materials and methods

We used 5515 accessions from 38 countries to develop the core subset for safflower including seven accessions where information on country of origin was not available but data on 12 morphological descriptors was available. There were 2272 new accessions that were not included by Suresh and Balakrishnan (2001) and Balakrishnan and Suresh (2000, 2001a, b) in their study. They were grown in augmented randomized block design at Solapur (longitude 17° 14' N; altitude 75° 56' E; elevation 483.63 m MSL), India. A control cultivar, Bhima, was planted after every 10 rows of test entries. The accessions were grown on a single row of 5 m long, spaced at 45 cm between rows and 20 cm within the row. Locally recommended cultural practices (40:40:20::N:P:K Kg ha^{-1}, irrigation, and protection against pests and diseases) were adopted to raise a good crop. Data on 12 morphological and 10 agronomic descriptors, chosen from the safflower descriptors list developed by the National Bureau of Plant Genetic Resources, New Delhi, for distinctness, uniformity, and stability testing, were recorded for 10 competitive plants in each accession (Table 1). However, only morphological descriptors were used for the development of a core subset.

The accessions were first stratified by country of origin and then grouped into nine geographic regions following Brown (1989a). Seven accessions of unknown origin were assigned into a separate group. The data on 12 morphological descriptors in each

Table 1. List of 12 morphological and 10 agronomic descriptors recorded on 5522 safflower accessions.

Descriptor	Abbreviation	Classes and stage of evaluation
Morphological descriptor		
Growth habit	GH	1 = Appressed, 2 = Bushy, 3 = Cone shaped, 4 = Cylindrical, and 5 = Errect, recorded a week before harvest.
Shape of upper stem leaves	SUSL	1 = Dentate, 2 = Lanceolate broad, 3 = Lanceolate narrow, 4 = Linear, 5 = Oblong, and 6 = Ovate, recorded at bloom stage.
Margin of upper stem leaves	MUSL	1 = Deeply lobbed, 2 = Deeply serrate, 3 = Entire, 4 = Few serrate, 5 = Oblong, 6 = Serrate, 7 = Slightly serrate, recorded at bloom stage.
Number of spines on outer involucral bracts	NSOIB	1 = Few spines, 2 = Intermediate, 3 = Many spines, and 4 = Non spiny, recorded one week before harvest.
Corolla color at bloom stage	CCBS	1 = Light yellow, 2 = Orange, 3 = Pale yellow, 4 = White, and 5 = Yellow, recorded at bloom stage.
Corolla color at dry stage	CCDS	1 = Deep red, 2 = dirty white, 3 = Golden yellow, 4 = Grey white, 5 = Orange, 6 = Orange red, 7 = Pale yellow, 8 = Pinkish white, 9 = Red, 10 = White, and 11 = Yellow, recorded after petals were dried.
Hull color	HC	1 = Cream, 2 = Partial hull, 3 = Smooth striped, 4 = Striped, and 5 = White, recorded on dry seeds.
Hull thickness	HT	1 = Thin, 2 = Intermediate, and 3 = Thick, recorded on dried seeds
Pappus of the achene	PA	1 = Absent, 2 = Negligible, and 3 = Present, recorded after threshing of the capitula
Seed shape	SS	1 = Conical, 2 = Cresent, and 3 = Oval, recorded on dried seeds
Seed length	SL	1 = large, 2 = Intermediate, and 3 = Small, recorded on dried seeds
Seed width	SW	1 = Broad, 2 = Intermediate, and 3 = Narrow, recorded on dried seeds
Agronomic descriptor		
Days to 50% flowering	DF	Days from seedling emergence to 50% plants produced flowers.
Days to first capitula formation	DFCF	Days from seedling emergence to initiation of first capitula formation.
Plant height (cm)	PH	Recorded one week before harvest
Plant canopy (cm)	PC	Recorded one week before harvest
Length of first primary branch (cm)	LFPB	Recorded one week before harvest
Angle of first primary branch	AFPB	Recorded one week before harvest using protracter?
Capitula per plant	CPP	Recorded at maturity
Diameter of main capitula (cm)	DMC	Recorded on dry capitula at maturity
Seed yield per plant (g)	SYP	Recorded on dry samples with seed moisture $<10\%$.
100-seed weight (g)	HSW	Recorded on 100 randomly selected sound mature seeds from the bulk seed samples.

group was standardized using the range of each variable to eliminate scale differences (Milligan and Cooper 1985). The standardized data was subjected to the hierarchical cluster algorithm of Ward (1963) at an r^2 (squared multiple correlation) value of 0.70, using SAS (SAS Institute 1989). This method optimizes an objective function because it minimizes the sum of the square within groups and maximizes the sum of squares among groups. The agglomerative procedure starts with n groups (i.e., one observation in one group; maximum among group sum of squares), and proceeds by merging observations in groups so that the between-groups sum of squares decreases and within-groups sum of squares increases. In certain cases the within-groups sum of square will remain the same, but it will never decrease. From each cluster, ≈10% of the accessions were randomly selected for inclusion into the core subset. At least one accession was included from those clusters that had less than 10 accessions.

Means of the entire collection and core subset for the 12 morphological descriptors were compared using Newman–Keuls procedure (Newman 1939; Keuls 1952). The homogeneity of variances of the entire collection and core subset was tested with the Levene's test (Levene 1960). The distribution homogeneity for each descriptor among the entire collection and the core subset was analyzed by the χ^2-test. The phenotypic correlation among different traits in the entire collection and core subset was estimated to know whether these associations, which may be under genetic control, were conserved in the core subset.

Results and Discussion

A core subset of 570 accessions was established from 5522 safflower accessions. These 570 accessions were arrayed into 25 distinct clusters. South Asia and Southeast Asia together accounted for 79.8% (4406 accessions) of the accessions in the entire collection, and this predominance was also reflected in the core subset that contain, 77.7% (443 accessions) of the accessions from these regions (Table 2). About 7.2%

Table 2. Region- and-country-wise representation of accessions in entire collection (EC) and core subset (CS) in safflower.

Region	Country	No. of accessions in EC	No. of accessions in CS	Region	Country	No. of accessions in EC	No. of accessions in CS
South Asia	Bangladesh	6 (0.2)*	1 (0.3)	USSR	USSR	21 (0.4)	8 (1.4)
	India	3364 (98.1)	336 (98.0)	Africa	Ethiopia	12 (30.8)	4 (40)
	Pakistan	59 (1.7)	6 (1.7)		Kenya	8 (20.5)	1 (10)
	Total**	3429 (62.1)	343 (60.2)		South Africa	2 (5.1)	1 (10)
South East Asia	China	971 (99.4)	98 (98.0)		Sudan	17 (43.6)	4 (40.0)
	Japan	4 (0.4)	1 (1.0)		Total	39 (0.7)	10 (1.8)
	Thailand	2 (0.2)	1 (1.0)	Americas	Argentina	2 (0.5)	1 (2.4)
	Total	977 (17.7)	100 (17.5)		Canada	14 (3.6)	2 (4.9)
West Asia	Afghanistan	12 (13.2)	2 (11.8)		USA	378 (95.9)	38 (92.7)
	Iran	71 (78)	12 (70.6)		Total	394 (7.1)	41 (7.2)
	Iraq	1 (1.1)	1 (5.9)	Australia	Australia	18 (0.3)	7 (1.2)
	Jordan	7 (7.7)	2 (11.8)	Europe	Austria	1 (1.4)	1 (5.9)
	Total	91 (1.6)	17 (3.0)		Bulgaria	2 (2.9)	1 (5.9)
Mediterranean	Algeria	1 (0.6)	1 (4.3)		Germany	13 (18.6)	2 (11.8)
	Egypt	46 (26.1)	5 (21.7)		Hungary	27 (38.6)	5 (29.4)
	France	3 (1.7)	1 (4.3)		Poland	5 (7.1)	1 (5.9)
	Israel	21 (11.9)	2 (8.7)		Portugal	20 (28.6)	5 (29.4)
	Italy	9 (5.1)	2 (8.7)		Switzerland	1 (1.4)	1 (5.9)
	Lebanon	2 (1.1)	1 (4.3)		United Kingdom	1 (1.4)	1 (5.9)
	Libya	1 (0.6)	1 (4.3)		Total	70 (1.3)	17 (3.0)
	Morocco	5 (2.8)	1 (4.3)	Unknown		7 (0.1)	4 (0.7)
	Spain	8 (4.5)	1 (4.3)				
	Syria	7 (4.0)	1 (4.3)				
	Turkey	73 (41.5)	7 (30.4)				
	Total	176 (3.2)	23 (4.0)				

*Figure in the parenthesis represents the percentage of germplasm representing into the EC and CS within region.
**Figure in the parenthesis represents the percentage of germplasm representing from the region into the EC and into the CS.

(41) of the accessions in the core subset were from the Americas, 4% (23) from the Mediterranean, and 3% each from Europe (17) and West Asia (17) regions. Australia, former USSR, and Africa represented 1.2%, 1.4%, and 1.8% accessions, respectively, in the core subset.

The ranges, means, and variances of the 12 morphological and 10 agronomic descriptors are given in Table 3. Differences among means of the entire collection and core subset for the 12 morphological descriptors used in developing the core subset were not significant, and the variances of the entire collection and core subset were homogeneous for all the traits except for growth habit ($P = 0.024$). The core subset captured 100% range variation for 11 morphological descriptors and 80% for shape of upper stem leaves. The differences between means and variances, except for seed yield per plant ($P = 0.034$), of the entire collection and core subset for the 10 agronomic descriptors were also not significantly different. However, this core subset captured over 70% range variation for six agronomic descriptors (days to first capitula formation, plant canopy, length of the first primary branch, angle of the first primary branch to the main stem, diameter of the main capitula, and seed yield per plant). For the remaining four agronomic descriptors (days to 50% flowering, plant height, number of capitula per plant, and 100-seed weight), this core subset could capture between 41% and 62% range variation of the entire collection. The analysis of frequency distribution of 12 morphological descriptors, except for growth habit ($P = 0.036$), indicated the homogeneity of distribution among the entire collection and core subset (Table 4). Suresh and Balakrishnan (2001) compared the diversity of the core sample with that of the whole collection using six different sampling strategies. The pool diversity index based on 28 descriptors was close to the diversity of the whole collection. However, when accessions from different diversity groups were allocated with equal frequency or in proportion to the logarithm of the number of accessions in each group or in the proportion to the square root-proportion of the number of accessions in each group, the resultant core samples

Table 3. Range, mean, and variance of 12 morphological and 10 agronomic descriptors recorded in the EC and CS of safflower.

	Range		Mean*		Significance	Variance**			
Trait	EC	CS	EC	CS		EC	CS	*F*-value	*P*-value
Morphological descriptor									
Growth habit	1.0–5.0	1.0–5.0	2.90	2.90	NS	1.41	1.59	5.11	0.024
Shape of upper stem leaves	1.0–6.0	2.0–6.0	5.60	5.60	NS	0.81	0.84	0.10	0.754
Margin of upper stem leaves	1.0–7.0	1.0–7.0	6.50	6.50	NS	1.53	1.58	0.06	0.806
Number of spines on outer involucral bracts	1.0–4.0	1.0–4.0	2.10	2.10	NS	0.41	0.41	0.001	0.987
Corolla color at bloom stage	1.0–5.0	1.0–5.0	4.20	4.20	NS	1.70	1.76	0.32	0.573
Corolla color at dry stage	2.0–11.0	2.0–11.0	6.80	6.80	NS	3.85	3.87	0.001	0.977
Hull color	2.0–7.0	2.0–7.0	5.70	5.70	NS	4.57	4.54	0.02	0.879
Hull thickness	1.0–3.0	1.0–3.0	2.60	2.60	NS	0.49	0.49	0.01	0.937
Pappus of the achene	1.0–3.0	1.0–3.0	1.10	1.10	NS	0.08	0.10	0.79	0.374
Seed shape	1.0–3.0	1.0–3.0	1.10	1.10	NS	0.17	0.18	0.04	0.849
Seed length (mm)	1.0–3.0	1.0–3.0	1.90	1.90	NS	0.61	0.62	0.22	0.635
Seed width (mm)	1.0–3.0	1.0–3.0	1.60	1.60	NS	0.25	0.26	0.60	0.437
Agronomic descriptor									
Days to 50% flowering	71.8–86.2	76.8–85.7	81.50	81.60	NS	1.84	1.64	1.85	0.174
Days to first capitula formation	42.9–76.6	49.2–73.2	60.50	60.50	NS	11.94	10.96	0.86	0.354
Plant height (cm)	49.1–106.0	62.2–94.3	74.40	74.60	NS	21.97	21.01	0.26	0.612
Plant canopy (cm)	40.5–46.3	41.0–46.3	41.90	42.00	NS	0.09	0.12	2.58	0.108
Length of first primary branch (cm)	36.9–39.5	37.2–39.5	37.80	37.80	NS	0.07	0.06	2.03	0.155
Angle of first primary branch to main stem	35.9–48.6	38.2–48.6	42.50	42.40	NS	1.57	1.62	0.17	0.677
Capitula per plant	13.5–64.4	14.8–45.1	25.90	25.80	NS	18.08	14.45	2.13	0.145
Diameter of main capitula	1.8–2.8	1.8–2.6	2.20	2.20	NS	0.01	0.01	0.01	0.914
Seed yield per plant (g)	11.8–19.6	11.9–19.6	13.00	13.00	NS	0.33	0.47	4.48	0.034
100–seed weight (g)	3.3–8.3	3.6–5.6	4.50	4.50	NS	0.10	0.10	0.19	0.661

*NS- Non significant at 0.05.

**Differences between mean of EC and CS were tested by Newman-Keuls test, and variance homogeniety was tested by Levene's test.

Table 4. Chi-square test and probability for comparison of frequency distribution of 12 morphological descriptors between the CS and entire safflower collection.

Descriptor	Number of classes	χ^2	Probability
Shape of upper stem leaves	6	7.80	0.167
Margin of the upper stem leaves	7	4.49	0.618
Number of spines on outer involucral bracts	4	0.83	0.841
Growth habit	4	8.57	0.036
Corolla color at bloom stage	4	0.50	0.919
Corolla color at dry stage	7	2.33	0.887
Hull color	5	1.26	0.869
Seed shape	3	1.19	0.203
Pappus of the achene	3	1.77	0.413
Hull thickness	3	0.02	0.992
Seed length	3	0.49	0.783
Seed width	3	1.49	0.474

had higher levels of diversity than the whole collection. Johnson et al. (1993) also reported that a core of 210 *C. tinctorius* L. accessions, roughly 10% of the 2000 accessions, represented the whole collection.

Ortiz et al. (1998) emphasized the importance of proper and adequate sampling for the conservation of phenotypic associations arising from co-adapted gene complexes in core collection. There is a fair degree of similarity in phenotypic correlation coefficients among morphological and quantitative descriptors, suggesting that this core subset has preserved most of the co-adapted gene complexes controlling these associations. The correlation (r) values, except for days to flower and days to first capitula formation ($r = 0.830$ in entire collection and 0.800 in core subset) were low but significant (47 and 25 correlation coefficients in entire collection and core subset, respectively in Table 5 and 40 and 31 correlation coefficients in entire collection and core subset, respectively in Table 6), indicating that these correlation's did not explain a large fraction of the total variation. The Shannon-Weaver diversity index (H′) was used to measure allelic richness and evenness in the entire collection and core subset. A low H′ indicates an extremely unbalanced frequency of classes for an individual trait and a lack of genetic diversity. In the present study, H′ values for all the 12 morphological and 10 agronomic descriptors were similar in the entire collection and core subset (Table 7), indicating that the diversity of the entire collection was represented in the core subset. This core subset was developed using data on morphological descriptors that have high heritability and are least influenced by genotype × environment interaction. It was further validated using 10 agronomic descriptors. Unlike earlier core collections reported in safflower, it represents variability from the large collections of accessions used in the development of this core subset.

Ashri and Knowles (1960) reported the Near East to be the center of origin of cultivated safflower (*C. tinctorius* L.) where the closely related wild species, *Carthamus persicus* Willd. (syn. *Carthamus flavescens*), a self-incompatible species in Turkey, Syria, and Lebanon and *Carthamus palaestinus* Eig, a

Table 5. Correlation coefficients between 12 morphological descriptors in the EC (above diagonal) and CS (below diagonal) of safflower.

	SUSL	MUSL	NSOIB	GH	CCBS	CCDS	HC	SS	PA	HT	SL	SW
SUSL	–	0.479**	−0.178**	0.058**	0.405**	−0.016	0.004	−0.146**	−0.037**	−0.081**	−0.092**	0.207**
MUSL	0.487**	–	−0.285**	0.091**	0.436**	−0.02	−0.046**	−0.220**	−0.001	−0.085**	−0.070**	0.174**
NSOIB	−0.075	−0.204**	–	−0.058**	−0.273**	−0.005	0.001	0.143**	0.011	0.037**	−0.012	−0.097**
GH	0.033	0.08	−0.032	–	0.151**	0.019	−0.034**	−0.089**	0.023	0.123**	0.039**	0.097**
CCBS	0.390**	0.378**	−0.223**	0.173**	–	−0.278**	−0.033**	−0.255**	−0.007	−0.098**	−0.044**	0.260**
CCDS	0.004	0.012	0.001	0.008	−0.273**	–	0.045**	−0.081**	0.001	0.073**	0.011	0.069**
HC	0.024	−0.046	−0.052	−0.04	−0.05	0.027	–	−0.011	0.034**	0.044**	0.045**	0.095**
SS	−0.148**	−0.140**	0.045	−0.079	−0.243**	−0.083*	0.019	–	−0.022	−0.038**	−0.054**	−0.155**
PA	−0.005	0.036	−0.029	−0.027	0.01	0.005	0.058	−0.01	–	0.017	0.006	0.014
HT	−0.069	−0.100*	0.003	0.141**	−0.139**	0.098*	0.032	−0.019	0.051	–	0.012	0.064**
SL	−0.125**	−0.085*	0.004	0.028	−0.098	0.061	0.041	0.013	−0.008	−0.055	–	−0.065**
SW	0.192**	0.110**	−0.069	0.058	0.251**	0.091**	0.146**	−0.149**	0.003	0.068	−0.089*	–

SUSL - Shape of upper stem leaves, MUSL - Margin of upper stem leaves, NSOIB - Number of spines on outer involucral bracts, GH - Growth habit, CCBS- Corolla color at bloom stage, CCDS- Corolla color at dry stage, HC - Hull color, SS - Seed shape, PA - Pappus of the achene, HT - Hull thickness, SL - Seed length, and SW - Seed width.

* and ** - Significant at 0.05 and 0.01 probability level, respectively.

Table 6. Correlation coefficients between 10 agronomic descriptors in the EC (above diagonal) and CS (below diagonal) of safflower.

	DFCF	DF	PH	PC	LFPB	AFPB	CPP	DMC	SYP	HSW
DFCF	–	0.830**	0.389**	0.060**	0.256**	−0.163*	−0.027*	0.029*	−0.107**	−0.257**
DF	0.800**	–	0.423**	0.076**	0.276**	−0.155**	−0.010	0.017	−0.103**	−0.266**
PH	0.405**	0.434**	–	0.202**	0.391**	−0.118**	0.018	0.171**	−0.056**	−0.140**
PC	0.070	0.132**	0.180**	–	0.576**	0.214**	0.534**	0.151**	0.335**	0.043**
LFPB	0.229**	0.268**	0.382**	0.544**	–	−0.011	0.374**	0.148**	0.178**	−0.066**
AFPB	−0.102*	−0.094*	−0.093*	0.305**	0.051	–	0.199**	0.029*	0.183**	0.104**
CPP	−0.011	0.034	−0.007	0.462**	0.310**	0.285**	–	0.017	0.370**	0.077**
DMC	0.041	0.039	0.169**	0.108**	0.196**	0.002	−0.002	–	0.138**	0.041**
SYP	−0.028	0.005	0.001	0.449**	0.291**	0.265**	0.368**	0.125**	–	0.213**
HSW	−0.242**	−0.259**	−0.098*	0.093	0.001	0.146**	0.137**	0.042	0.236**	–

DFCF - Days to first capitula formation, DF - Days to 50% flowering, PH - Plant height, PC - Plant canopy, LFPB - Length of first primary branch, AFPB - Angle of first primary branch to main stem, CPP - Capitula per plant, DMC - Diameter of main capitula, SYP - Seed yield per plant, HSW - 100-seed weight.
* and ** - Significant at 0.05 and 0.01 probability level, respectively.

self-compatible wild species restricted to the deserts of southern Israel and western Iraq (Zeven and Zhukovsky 1975), are found. It is expected that there would be introgression of wild characteristics into domestic types, thus creating new genetic variability for beneficial traits. Accessions from South Asia and South East Asia regions in this core subset are over-represented (77.7%), whereas accessions from West Asia and Mediterranean regions are under- represented (7%). There is, therefore, need to plan for exploration in West Asia and Mediterranean regions to enrich world safflower collection. The safflower cultivars and hybrids have a very narrow genetic base, mainly because of lack of information on the genetic variability in germplasm possessing beneficial traits. For example, in the last four decades (1960–2002) in India, 21 safflower varieties and 3 hybrids have been released for cultivation. Except for A1, Nira, Girna, JS1–73, NARI 6, DSH 129, MKH 11, and PH 6 which have been developed as a result of cross-breeding using single crosses, the remaining plants were either released as direct introduction or re-elected from the original land race populations. This core subset contains 570 accessions arrayed into 25 distinct clusters. However, accessions within the cluster will probably be similar and therefore assist users to identify additional germplasm with similar characteristics. Pests including aphids (*Myzus persicae* Sulz; *Macrosiphum* spp.; *Uroleucon carthami* Hille Ris Lambers), fruit fly (*Acanthiophilus helianthi* Rossi), leafworm (*Spodoptera* spp.), bollworms (*Helicoverpa* spp.), and stem borers (*Melangromyza* spp.) and diseases including leaf blight (*Alternaria carthami* Choudhary), root rot (*Phytophthora* spp.; *Rhizoctonia bataticola* (Taub) Butler), wilt (*Sclerotinia sclerotiorum* (Lib.) de Bary); *Verticillium alboatrum* Reinke and Berthier; *Fusarium oxysporum* spp.), leaf

Table 7. Shannon-Weaver diversity index for 12 morphological and 10 agronomic descriptors in the ES and CS of safflower.

Descriptor	EC	CS
Morphological descriptor		
Shape of upper stem leaves	0.31	0.33
Margin of the upper stem leaves	0.31	0.32
Number of spines on outer involucral bracts	0.40	0.40
Growth habit	0.54	0.56
Corolla color at bloom stage	0.34	0.34
Corolla color at dry stage	0.48	0.48
Hull color	0.30	0.30
Seed shape	0.09	0.10
Pappus of the achene	0.11	0.10
Hull thickness	0.35	0.35
Seed length	0.47	0.47
Seed width	0.31	0.31
Mean ± SE	0.33 ± 0.039	0.34 ± 0.038
Over all mean ± SE	0.46 ± 0.036	0.45 ± 0.035
Agronomic descriptor		
Days to first capitula formation	0.60	0.61
Days to 50% flowering	0.62	0.62
Plant height (cm)	0.62	0.61
Plant canopy (cm)	0.62	0.56
Length of first primary branch (cm)	0.63	0.61
Angle of first primary branch to main stem	0.62	0.63
Capitula per plant	0.57	0.59
Diameter of the main capitula	0.62	0.61
Seed yield per plant (gm)	0.55	0.45
100-seed weight (gm)	0.62	0.62
Mean ± SE	0.61 ± 0.009	0.59 ± 0.016

spots (*Cercospora carthami* (H. and P. Sydow) Subram and Ramkr.); *Ramularia carthami* Zaprometor), rust (*Puccinia carthami* (Hutz) Corda), and bud blight (*Phytophthora drechsleri* Trucker; *Bortrytis cinerea* Pers.: Fr) are the major biotic constraints to world safflower production. There are only few documented cases wherein accessions with beneficial traits have been identified and channeled into breeding programs for enhancing the genetic potential of the safflower crop. EC 210582 (GMU 2047) from cluster 1, EC 181204 (GMU 990) from cluster 8, and EC 181465 (GMU 1136) from cluster 17, for example, are the only accessions in the core subset that were reported immune to *Fusarium* wilt under artificial wilt sick plot conditions from India (R.D. Prasad, pers comm.). For identifying additional germplasm with similar characteristics, the researchers may concentrate on evaluation of accessions falling in clusters 1, 8, and 17, rather than screening the whole core subset. Thus the evaluation of this core subset vis-a-vis the entire collection will save time and resources and at the same time provide useful information on genetic variability for beneficial traits. It will also assist the germplasm curator to acquire new variability for traits showing limited variation in the core subset. With the development of this core subset in safflower, it should now be feasible to conduct extensive multi-location evaluation of these 570 accessions into diverse agro-ecological regions for resistance to abiotic and biotic stresses and for agronomic and seed quality traits. The information obtained from this multi-location evaluation can be used to develop a mini core subset, as done in case of chickpea and peanut (Upadhyaya and Ortiz 2001; Upadhyaya et al. 2002), to identify germplasm with beneficial traits for utilization in safflower breeding programs. Further studies may be planned to characterize a mini core set for genetic diversity using DNA marker technology (Bernatsky and Tanksley 1989; Virk et al. 1995; Gilbert et al. 1999). Isozyme and RAPD techniques have been found useful in characterizing safflower genetic resources (Zhang 2001; Yazdi-Samadi et al. 2001). Thus the genetically diverse germplasm, based on phenotypic and molecular diversity, may become available to breeders for enhancing the genetic potential of safflower crop. The list of accessions included in this core subset is available from the Directorate of Oilseeds Research (e-mail: director@dor-icar.org; Fax 91-040-4017969) and also from the senior author (e-mail: sldwivedi@yahoo.com) on request.

Acknowledgements

The authors gratefully acknowledge the contributions made by scientists of Safflower Project Coordinating Unit, Solapur, India for characterizing germplasm for various descriptors. The senior author also acknowledges the Indian Council of Agricultural Research, New Delhi, India for providing "visiting scientist fellowship" that enabled him to complete this study.

References

Ashri A. and Knowles P.F. 1960. Cytogenetics of safflower (*Carthamus* L.) species and their hybrids. Agron. J. 52: 11–17.

Balakrishnan R. and Suresh K.K. 2000. Some strategies for obtaining core samples from germplasm collections using strata of geographical origins—A case study in safflower (*Carthamus tinctorius* L.). Stat. Appl. 2: 49–64.

Balakrishnan R. and Suresh K.K. 2001a. Strategies for developing core collections of safflower (*Carthamus tinctorius* L.) germplasm – Part II. Using an information measure for obtaining a core sample with predetermined diversity levels for several descriptors simultaneously. Indian J. Plant Genet. Resour. 14: 32–42.

Balakrishnan R. and Suresh K.K. 2001b. Strategies for developing core collections of safflower (*Carthamus tinctorius* L.) germplasm – Part III. Obtaining diversity groups based on an information. Indian J. Plant Genet. Resour. 14: 342–349.

Bernatsky R. and Tanksley S.D. 1989. Restriction fragments as molecular markers for germplasm evaluation and utilization. In: Brown A.H.D., Frankel O.H., Marshall D.R. and Williams J.T. (eds), The Use of Plant Genetic Resources. Cambridge University Press, Cambridge, England, pp. 353–362.

Bisht I.S., Mahajan R.K., Loknathan T.R. and Agrawal R.C. 1998. Diversity in Indian sesame collection and stratification of germplasm accessions in different diversity groups. Genet. Resour. Crop Evol. 45: 325–335.

Brown A.H.D. 1989a. Core collections: a practical approach to genetic resources management. Genome 31: 818–824.

Brown A.H.D. 1989b. The case for core collections: a practical approach to genetic resources management. In: Brown A.H.D., Frankel O.H., Marshall D.R. and Williams J.T. (eds), The Use of Plant Genetic Resources. Cambridge University Press, Cambridge, England, pp. 136–155.

Cordeiro C.M.T., Morales E.A.V., Ferreira P., Rocha D.M.S., Costa I.R.S., Valois A.C.C. and Silva S. 1995. Towards a Brazilian core collection of cassava. In: Hodkin T., Brown A.H.D., van Hintum Th.J.L. and Morales B.A.V. (eds), Core Collection of Plant Genetic Resources, International Plant Genetic Resources Institute (IPGRI). John Wiley and Sons, New York, USA, pp. 155–168.

Damodaram T. and Hegde D.M. 2002. Oilseeds Situation: A Statistical Compendium 2002. Directorate of Oilseeds Research, Rajendranagar, Hyderabad 500030, India, p. 471

Danisova C. and Sarooka I. 1994. Yellow pigments of the flowers of *Carthamus tinctorius*. Biologia Bratislava 49: 887–892.

Erskine W. and Muehlbauer F.J. 1991. Allozyme and morphological variability, outcrossing rate and core collection formation in lentil germplasm. Theor. Appl. Genet. 83: 119–125.

Frankel O.H. 1984. Genetic perspective of germplasm conservation. In: Arber W., Llimensee K., Peacock W.J. and Starlinger P. (eds), Genetic Manipulations: Impact on Man and Society. Cambridge University Press, Cambridge, England, pp. 161–170.

Frankel O.H. and Brown A.H.D. 1984. Current plant genetic resources – a critical appraisal. In: Chopra V.L., Joshi B.C., Sharma R.P. and Bansal H.C. (eds), Genetics: New Frontiers, Vol IV. Oxford and IBH Publishing Company, New Delhi, India, pp. 1–63.

Gilbert J.E., Levis R.V., Wilkinson M.J. and Caligari P.D.S. 1999. Developing an appropriate strategy to assess genetic variability in plant germplasm collections. Theor. Appl. Genet. 98: 1125–1131.

Huamán Z., Aguilar C. and Ortiz R. 1999. Selecting a Peruvian sweet potato core collection on the basis of morphological, ecogeographical, and disease and pest reaction data. Theor. Appl. Genet. 98: 840–844.

Johnson R.C., Bergman J.W. and Flynn C.R. 1999. Oil and meal characteristics of core and non-core safflower accessions from the USDA collection. Genet. Resour. Crop Evol. 46: 611–618.

Johnson R.C., Stout D.M. and Bradley V.L. 1993. The US Collection: a rich source for safflower germplasm. In: Dajue L. and Yuanzhou H. (eds), Proceedings of the Third International Safflower Conference, Beijing Botanical Garden, Institute of Botany. Chinese Academy of Sciences, Beijing, China, pp. 202–208.

Keuls M. 1952. The use of "Studentized range" in connection with an analysis of variance. Euphytica 1: 112–122.

Knüpffer H. and van Hintum Th.J.L. 1995. The barley core collection: an international effort. In: Hodkin T., Brown A.H.D., van Hintum Th.J.L. and Morales B.A.V. (eds), Core Collection of Plant Genetic Resources, International Plant Genetic Resources Institute (IPGRI). John Wiley and Sons, New York, USA, pp. 171–178.

Lachover D. and Kostrinski J. 1965. Essais et observations sur la valeur des verts de carthame comme fourrage et ensilage. Qualitas Pl. Mater. Veg. 12: 363–375.

Levene H. 1960. Robust tests for equality of variances. In: Olkin I. (ed.), in Honour of Harold Hotelling. Stanford University Press, Stanford, pp. 278–292.

Milligan G.W. and Cooper M.C. 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika 50: 159–179.

Nagaraj G. 2001. Nutritional characteristics of three Indian safflower cultivars. In: Bergman J.W., Mundel H.H., Jensen J.L., Flynn C.R., Grings E.E., Tanaka D.L., Riveland N.R., Johnson R.C. and Hill A.B. (eds), Proceedings pf the Fifth International Safflower Conference. Williston, North Dacota and Sidney, MT, USA, pp. 303–305.

Newman D. 1939. The distribution of range in samples from a normal population expressed in terms of an independent estimate of standard deviation. Biometrika 31: 20–30.

Ortiz R., Ruiz-Tapia E.N. and Mujica-Sanchez A. 1998. Sampling strategy for a core collection of Peruvian quinoa germplasm. Theor. Appl. Genet. 96: 475–483.

Rudometova N.V., Pasovskij A.P. and Blohina E.A. 2001. Method of isolation and identification of carthamin in safflower: application's perspectives in Russian food products. In: Bergman J.W., Mundel H.H., Jensen J.L., Flynn C.R., Grings E.E., Tanaka D.L., Riveland N.R., Johnson R.C. and Hill A.B. (eds), Proceedings of the Fifth International Safflower Conference. Williston, North Dakota and Sidney, MT, USA, pp. 309–313.

SAS Institute 1989. SAS/STAT User Guide Version 6, 4th edn. SAS Institute. Inc., Cary, NC, USA.

Spagnoletti Zeuli P.L. and Qualset C.O. 1987. Geographical diversity for quantitative spike characters in a world collection of durum wheat. Crop Sci. 27: 235–241.

Suresh K.K. and Balakrishnan R. 2001. Strategies for developing core collection of safflower (*Carthamus tinctorius* L.) germplasm – Part 1. Sampling from diversity groups of quantitative morphological descriptors. Indian J. Plant Genet.Resour. 14: 22–31.

Tohme J., Jones P., Beebe S. and Ibanga M. 1995. The combined use of agroecological and characterization data to establish the CIAT *Phaseolus vulgaris* core collection. In: Hodgkin T., Brown A.H.D., van Hintum Th.J.L. and Morales B.A.V. (eds), Core Collection of Plant Genetic Resources, International Plant Genetic Resources Institute (IPGRI). John Wiley and Sons, New York, USA, pp. 95–108.

Upadhyaya H.D., Bramel P.J., Ortiz R. and Singh S. 2002. Developing a mini core of peanut for utilization of genetic resources. Crop Sci. 42: 2150–2156.

Upadhyaya H.D., Bramel P.J., Ortiz R. and Singh S. 2003. Development of a groundnut core collection using taxonomical, geographical and morphological descriptors. Genet. Resour. and Crop Evol. 50: 139–148.

Upadhyaya H.D., Bramel P.J. and Singh S. 2001. Development of a chickpea core subset using geographic distribution and quantitative traits. Crop Sci. 41: 206–210.

Upadhyaya H.D. and Ortiz R. 2001. A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. Theor. Appl. Genet. 102: 1292–1298.

Varma M., Shukla Y.N., Ram M., Jain S.P. and Kumar S. 1997. Chemistry and biology of the oil and dye crop *Carthamus tinctorius*: a review. J. Med. Aromat. Pl. Sci. 19: 734–744.

Virk P.S., Ford B.V.L., Jackson M.T., Pooni H.S. and Newbury H.J. 1995. Use of RAPD for the study of diversity within plant germplasm collections. Heredity 74: 170–179.

Ward J. 1963. Hierarchical grouping to optimize an objective function. J. Amer. Stat. Assoc. 38: 236–244.

Wichman D.M., Welty L.E., Strang L.M., Bergman J.W., Westcott M.P., Stallknecht G.F., Riveland N.R. and Ditterline R.L. 2001. Assessing the forage production potential of safflower in the Northern Great Plains and inter-mountain regions. In: Bergman J.W., Mundel H.H., Jensen J.L., Flynn C.R., Grings E.E., Tanaka D.L., Riveland N.R., Johnson R.C. and Hill A.B. (eds), Proceedings of the Fifth International Safflower Conference. Williston, North Dakota, Sidney, MT, USA, pp. 269–273.

Yazdi-Samadi B., Amiri R.M., Ghannadha M.R. and Abd-Mishani C. 2001. Detection of DNA polymorphism in landrace populations of safflower in Iran using RAPD-PCR technique. In: Bergman J.W., Mundel H.H., Jensen J.L., Flynn C.R., Grings E.E., Tanaka D.L., Riveland N.R., Johnson R.C. and Hill A.B. (eds), Proceedings of the Fifth International Safflower Conference. Williston, North Dakota and Sidney, MT, USA, p.163.

Zeven A.C. and Zhukovsky P.M. 1975. Dictionery of Cultivated Plants and their Centers of Diversity, Center for Agricultured Publication and Documentation (Pudoc). Wageningen, Netherlands, p.219

Zhang Z. 2001. Genetic diversity and classification of safflower (*Carthamus tinctorius* L.) germplasm by isozyme techniques. In: Bergman J.W., Mundel H.H., Jensen J.L., Flynn C.R., Grings E.E., Tanaka D.L., Riveland N.R., Johnson R.C. and Hill A.B. (eds),

Proceedings of the Fifth International Safflower Conference. Williston, North Dakota and Sidney, MT, USA, pp. 157–162.

Zhaomu W. and Lijie D. 2001. Current situation and prospects of safflower products development in China. In: Bergman J.W., Mundel H.H., Jensen J.L., Flynn C.R., Grings E.E., Tanaka D.L., Riveland N.R., Johnson R.C. and Hill A.B. (eds), Proceedings of the Fifth International Safflower Conference. Williston, North Dakota and Sidney, MT, USA, pp. 315–319.