Genetic Resources and Crop Evolution (2006) 53: 225–244
DOI 10.1007/s10722-004-6150-9

Patterns of phenotypic variation in a germplasm collection of *Carthamus tinctorius* L. from the Middle East

A.A. Jaradat[1,*] and M. Shahid[2]
[1]*USDA-Agriculture Research Service, 803 Iowa Avenue, Morris, MN 56267, USA;* [2]*Plant Genetic Resources Program, International Center for Biosaline Agriculture (ICBA), P.O. Box 14660, Dubai, United Arab Emirates;* **Author for correspondence (e-mail: jaradat@morris.ars.usda.gov; phone: 1-320-589-3411; fax: 1-320-589-3787)*

Received 28 October 2003; accepted in revised form 30 January 2004

Key words: Carthamus, Diversity, Middle East, Safflower

Abstract

Phenotypic diversity was assessed for quantitative and qualitative traits in a salt-tolerant subset of the international safflower (*Carthamus tinctorius* L.) germplasm collection originating from 11 countries in three regions (Central Asia, Southwest Asia and Africa) of the Middle East. Phenotypically, the germplasm, among and within regions, was highly variable, especially for rosette- and yield-related traits. Frequency of desirable variants of seven agronomically important traits ranged from 14% for long rosette period to 50% for no or few spines. Level of population differentiation was high for number of capitula per plant (30%), whereas most traits partitioned their diversity (82–87%) within populations. Region-specific nonrandom associations among sets of qualitative traits and the existence of broad morphological and phenotypic diversity in this germplasm were supported by the large number of log-linear models needed to describe qualitative trait associations, the high number of principal components needed to account for total variability, and the low discriminatory power of phenotypic traits among germplasm from regions and countries in the Middle East. These results suggest that adaptation of the species to the wide spatial and temporal variation in the Middle East resulted in a multitude of ecotypes and in enormous amount of local variation. A multivariate selection criterion for high biological and seed yield, long rosette period and no or few spines identified five accessions from Southwest Asia that can be introduced into subsistence farming systems as a multipurpose crop under saline agriculture.

Introduction

Safflower (*Carthamus tinctorius* L.) is an annual multipurpose crop of economic importance in a few countries around the world, especially in a band from the Mediterranean to the Pacific Ocean at latitudes between 20°S and 40°N. In particular, it thrives under the hot, dry climate of the Middle East (Knowles and Ashri 1995). Cultivated safflower belongs to a group of closely related diploids ($2n = 2x = 24$) that extends from central Turkey in the west to north-western India in the east (Li and Mündel 1996). The area delimited by Turkistan, southern Turkey, western Iran, Iraq, Syria, Jordan and Israel in the Middle East is the center of origin and diversity of the species. The crop was domesticated in the Middle East for its dried flowers and as a dye plant and subsequently very widely used, until recently, for this purpose in the Mediterranean basin and the subtropics (Harlan 1992). Safflower spread from this region to China, India, Ethiopia, Sudan, North Africa

and Europe, and by the early 19th century became one of the most important plant sources of dyestuffs (Li and Mündel 1996).

Safflower developed considerable diversity after it was taken into cultivation for a long time across huge and diverse regions in the Old World, and there is evidence of incipient genetic differentiation (Knowles 1989). However, very little is known about the level of phenotypic plasticity and its relationship with genetic diversity in the cultivated species in its center of origin (Khan et al. 2003). Knowles (1969) coined the term 'centers of similarity' for seven regions which are not centers of diversity or origin, but of remarkably similar safflower types. These are as follows: the Far East, India–Pakistan, Middle East (Israel, Iran, Iraq, Jordan, Syria and Turkey), Egypt, Sudan, Ethiopia and Europe. Germplasm from the Middle East is dominated by late maturing, spineless, tall, and red-flowered genotypes. Egyptian germplasm was found to be variable in maturity, with usually large heads, whereas the Sudanese germplasm is early in maturity, very spiny and yellow-flowered (Knowles 1969). Safflower dyes were particularly important to the carpet-weaving industries of the Middle East. Today domesticated safflower is prized for its oil-bearing seeds and for the yellow dye in its flowers. Moreover, safflower can be grazed or stored as hay or silage. The forage is palatable and its feed value and yields are similar to or better than oats or alfalfa (Li and Mündel 1996).

Considerable efforts in germplasm collection and evaluation culminated in assembling a large germplasm collection at the USDA Regional Plant Introduction Station, Pullman, WA, USA (Johnson et al. 2001). This collection was characterized and evaluated for many traits of agronomic importance, including tolerance to salinity (Li et al. 1993). A salt-tolerant germplasm subset consisting of 631 germplasm accessions from 11 countries in the Middle East was acquired from the Western Plant Introduction Station, Pullman, WA, USA (NPGS 2000) for evaluation under the subtropical growing conditions of the Persian Gulf, and to select salt-tolerant accessions suitable for oil, dye or forage production under subsistence farming systems. The germplasm subset was evaluated for morphological and agronomic traits under the subtropical climate of the Persian Gulf to (1) quantify the level of phenotypic variation and detect how genetic variation is partitioned among and within countries of origin, (2) identify germplasm with high level of variation for key agronomic traits and (3) select salt-tolerant accessions adapted to short growing season, with long rosette period and high potential for biomass, seed and dye production.

Materials and methods

A field experiment was conducted during the 2001/2002 growing season at the Experiment Station of The International Center for Biosaline Agriculture, Dubai, United Arab Emirates (25°13′N and 55°17′E). We planted 10 seeds from each of 631 accessions of safflower from 11 countries (referred to hereafter as populations) across three geographical regions (Central Asia, Southwest Asia and Africa) in the Middle East at a distance of 50 cm between rows and seeds within rows (Table 1). Characterization and evaluation data (Table 1) were recorded on single plants during the growing season and at harvest. Hull percent was estimated using 50 random seeds per accession. Data for eight quantitative traits of agronomic value were collected on a minimum of five plants per accession of only 591 accessions. Safflower descriptor list (Li et al. 1993) was used to score eight qualitative traits on a single representative plant from each accession.

The mean and standard deviation (s.d.) calculated for eight quantitative traits (Table 1) were used to categorize accessions into three discrete classes (i.e., (low) $\leq$ mean $-$ 1.0 s.d., (medium) $>$ mean $-$ 1.0 s.d. $<$ mean $+$ 1.0 s.d. and (high) $\geq$ mean $+$ 1.0 s.d.) according to Zar (1996). The polymorphic index (I), as described by Zhang and Allard (1986), was calculated for geographic regions, populations and all 16 traits from the relative phenotypic frequencies for each categorical trait as $-\sum p_i \ln p_i$ for $i = 1, 2, 3$, where p_i is the relative frequency in the ith category of the jth trait (Yeh et al. 2000) and was used as a measure of phenotypic diversity. Total (H_T), within populations (H_S) and among populations (D_{ST}) genetic diversity, were calculated for geographical regions and populations using frequencies of all categorical data, then the genetic differentiation coefficient

Table 1. Region and country of origin, number of samples, and categorical levels of eight quantitative and qualitative traits recorded on 591 accessions of safflower from 11 Middle Eastern countries.

Region	Country	Number of accessions	Traits (number of categories)	
			Quantitative	Qualitative
Central Asia	Afghanistan	21	1. Rosette period, days (RP) (1-short, 2-medium, 3-long)	1. Location of branches on stem (LB) (1-upper third, 2-upper two-thirds, 3-from base to apex)
	Iran	168	2. Days to maturity (DM) (1-early, 2-medium, 3-late)	2. Angle of branches (AB) (1-appressed, 2-intermediate, 3-spreading)
	Pakistan	100	3. Plant height, cm. (PH) (1-short, 2-intermediate, 3-tall)	3. Leaf shape (LS) (1-Ovate, 2-lanceolate, 3-Oblong)
	Subtotal	289	4. Capitalum diameter, mm. (CD) (1-small, 2-intermediate, 3-large)	4. Leaf margin (LM) (1-entire, 2-serrate or dentate, 3-parted)
Southwest Asia	Israel	33	5. Capitula per plant (CP) (1-small, intermediate, 3-high)	5. Spininess (SP)(1-no or few spines, 2-intermediate, 3-many spines)
	Jordan	8	6. Biological yield/plant, g. (BY) (1-low, 2-intermediate, 3-high)	6. Capitulum shape (CS) (1-conical, 2-oval, 3-flattened)
	Syria	18	7. Hull percent (HL%) (1-low, 2-intermediate, 3-high)	7. Corolla color (CC)(1-yellow, 2-yellow-orange, 3-red–orange to red)
	Turkey	127	8. Seed yield per plant, g. (SY) (1-low, 2-intermediate, 3-high)	8. Seed shape (SS) (1-oval, 2-conical, 3-crescent)
	Subtotal	186		
Africa	Egypt	72		
	Ethiopia	28		
	Morocco	8		
	Sudan	8		
	Subtotal	116		
Middle East	Total	591		

(G_{ST}) was calculated as D_{ST}/H_T, then a genetic distance among populations was calculated based on frequencies of the eight categorical traits (Yeh et al. 2000) and a squared Mahalanobis distance (SM) among populations was calculated based on eight quantitative traits (StatSoft 2001). Quantitative data were standardized prior to statistical analyses using the nursery mean for each trait (StatSoft 2001).

Variation among variants of each categorical trait (e.g., short, medium and long rosette period) for eight quantitative traits of agronomic value was estimated via one-way analysis of variance (ANOVA). Results of the mean separation are presented as number of pairwise significant differences among variants of each categorical trait (minimum of zero and maximum of 3).

The Kruskal–Wallis test (Zar 1996), a non-parametric alternative to one-way ANOVA when error variance can not be estimated in a non-replicated field experiment, and to test whether different germplasm samples (from different regions or different populations within regions) were drawn from the same distribution with the same median. Additionally, it was used to identify categorical traits that can significantly discriminate among regions and populations. We used the principal component analysis (PCA) on the standardized means (Rohlf 2000) of the whole germplasm set and separately on germplasm from each country (i.e., population), as a linear dimensionality reduction technique to identify orthogonal directions of maximum variance in the original data set and to project the data into a lower (two dimensions) dimensionality space formed of a subset (two PCs) of the highest-variance components. Correlations between the first two PCs and the initial eight quantitative traits were calculated to

Table 2. Polymorphic diversity indices (I, mean ± standard deviation) and GD analyses (total diversity, H_T, diversity among regions or among countries, H_S, and level of differentiation, G_{ST}) based on 16 categorical traits scored on 591 accessions of safflower from 11 Middle Eastern countries.

Region	Country	I	H_T	H_S	G_{ST}
Central Asia	Afghanistan	0.71 ± 0.28	0.42 ± 0.04	0.28 ± 0.02	0.35
	Iran	0.77 ± 0.26	0.44 ± 0.03	0.36 ± 0.02	0.18
	Pakistan	0.73 ± 0.32	0.46 ± 0.04	0.38 ± 0.03	0.17
	Subregion	0.80 ± 0.28	0.45 ± 0.03	0.41 ± 0.03	0.08
Southwest Asia	Israel	0.78 ± 0.32	0.45 ± 0.03	0.36 ± 0.03	0.20
	Jordan	0.61 ± 0.35	0.34 ± 0.04	0.24 ± 0.02	0.30
	Syria	0.70 ± 0.32	0.43 ± 0.03	0.39 ± 0.03	0.09
	Turkey	0.75 ± 0.26	0.42 ± 0.02	0.38 ± 0.02	0.10
	Subregion	0.78 ± 0.28	0.45 ± 0.03	0.42 ± 0.03	0.07
Africa	Egypt	0.75 ± 0.26	0.42 ± 0.03	0.32 ± 0.02	0.23
	Ethiopia	0.72 ± 0.30	0.41 ± 0.03	0.34 ± 0.02	0.17
	Morocco	0.45 ± 0.33	0.32 ± 0.04	0.26 ± 0.02	0.15
	Sudan	0.56 ± 0.25	0.45 ± 0.06	0.35 ± 0.03	0.14
	Subregion	0.78 ± 0.26	0.38 ± 0.03	0.34 ± 0.02	0.10
Middle East Region		0.81 ± 0.27	0.46 ± 0.03	0.42 ± 0.02	0.21

aid in interpretation of the analysis. In order to explore the structure of categorical variables in a matrix of 11 populations and the desirable variants of eight categorical traits of agronomic value, we used a correspondence analysis as a dimensional reduction and perceptual mapping statistical procedure (Agresti 1990; Zar 1996). A two-dimensional plot was developed where associations among populations and levels of descriptive traits can be identified.

Tests of the independence of pairs of categorical data were performed using the standard two-way chi-square tests. Discrete multivariate log-linear analysis was used to quantitatively describe the multitrait organization of phenotypic variation (Agresti 1990) within the safflower germplasm collection in the Middle East and in Central Asia, Southwest Asia and Africa. A preliminary selection of interaction terms was performed and a basic log-linear model consisting of only significant interaction terms was constructed. The likelihood ratio statistics (G^2) were then calculated for the respective log-linear models and tested for the goodness-of-fit of the models. Stepwise multivariate discriminant analysis was used on the data matrix of 16 categorical traits, 591 cases, 3 geographical regions and 11 populations to (1) select a subset of traits that maximized differences among populations from the three geographical regions, and among populations from the 11 countries in the Middle East and (2) verify the separation of accessions into their respective populations. Finally, and based on results of multivariate analyses procedures, we identified a subset of accessions with high biomass and seed yield potential (>mean + 1.96 s.d.), and another subset with long rosette period (>mean + 1.96 s.d.), then we combined both subsets to identify accessions with multiple traits of agronomic value for biomass, seed and dye production under subsistence farming systems. The Tukey HSD *post hoc* test was used to determine the significant differences ($P < 0.05$) among means of diversity indices (Table 3) and means of squared Mahalanobis and genetic distances (Table 6). All statistical analyses were carried out using several modules of STATISTICA (StatSoft 2001), unless otherwise specified.

Results

A total of 48 categorical variants were scored on the whole germplasm collection from 11 populations in three geographical regions of the Middle East (Table 1). The overall polymorphic diversity index calculated from frequencies of these variants was 0.81, however, it was associated with a high coefficient of variation (33%). The regional and country-based polymorphic indices were highly variable as indicated by their standard deviations (Table 2), hence, there were no significant differences between the phenotypic diversity estimates

for the three geographical regions, or between populations within or among regions. None of the 16 trait was monomorphic and most showed at least two relatively frequent phenotypic variants.

Diversity analysis

Total diversity (H_T) based on frequencies of 16 categorical traits for the Middle East was estimated as 0.46, whereas the average diversity within regions (H_S) was 0.42, consequently, almost one-fifth ($G_{ST} = 21\%$) of total variation was apportioned among geographical regions. Populations from Central and Southwest Asia had similar H_T estimates, albeit a slightly lower G_{ST}, as compared with populations from Africa. Although the polymorphic diversity index (I) was positively correlated with H_T ($r = 0.69$, $P < 0.32$), and with H_S ($r = 0.59$, $P < 0.45$), this positive relationship did not influence its association ($r = 0.06$, $P < 0.89$) with the level of population differentiation (G_{ST}). Nevertheless, H_T and H_S were negatively, but not significantly correlated with G_{ST} ($r = -0.22$, $P < 0.47$ and $r = -0.69$, $P < 0.07$, respectively). Consequently, variation in H_S alone, and in H_S and H_T combined, accounted for a medium ($R^2 = 0.42$), and a high ($R^2 = 0.83$) portion of variation in G_{ST}, respectively. G_{ST} can be estimated as a function of H_T and H_S as $G_{ST} = 0.208 + 1.92\,H_T - 2.44\,H_S$ ($R^2 = 0.83$). G_{ST} estimates for individual populations ranged from a low of 0.09 for Syrian germplasm, to a high of 0.35 for the Afghani germplasm. However, almost 50% of the G_{ST} estimates were within the 0.14–0.20 range.

The high phenotypic diversity indices for individual traits (Table 3) suggest that this is a highly polymorphic germplasm collection. Most phenotypic diversity estimates were >0.8, and a few (e.g., capitulum diameter, spininess, corolla color and branch angle) displayed extreme levels of phenotypic diversity. Seven and 15 traits were powerful enough to differentiate between populations from geographical regions, and from different countries, respectively, as indicated by the P-values of the Kruskal–Wallis test (Table 3). The categorical traits, leaf margin, spininess, biological yield and seed yield were the most differentiating traits among regions and population.

Patterns of variation among variants of qualitative traits

There were 59 (or 46%) pairwise significant differences among means of variants of 16 categorical traits for eight traits of agronomic value (Table 3). Categorical traits with highly significant P-Kruskal–Wallis test among populations differentiated more than others among populations. Phenotypic variants of spininess, days to maturity, capitulum diameter and shape were among the most differentiating traits, whereas, hull percent, corolla color and plant height variants were amongst the least. Mean rosette period (17.9 days, range from 10 to 27 days), with a phenotypic diversity index of 0.84, differed significantly among variants of seven categorical traits. On average it represented 15% of the plant ontogeny, and was positively correlated with days to maturity ($r = 0.10$, $P < 0.05$, $n = 591$). This relationship was positive and highly significant in germplasm from Pakistan ($r = 0.35$, $P < 0.05$), Israel ($r = 0.46$, $P < 0.03$) and Syria ($r = 0.61$, $P < 0.01$). Days to maturity (mean 119.2 days, and a range from 108 to 139 days) differed significantly among variants of 11 categorical traits and was positively and significantly correlated ($r = 0.21$, $P < 0.01$, $n = 591$) with plant height. Mean plant height (75.5 cm) was significantly different among variants of eight categorical traits, and ranged from 28 to 135 cm. Although there was a strong positive correlation between plant height and seed yield per plant for the whole germplasm collection ($r = 0.63$, $P < 0.01$, $n = 591$), plant height was negatively and significantly correlated with seed yield per plant ($r = -0.30$, $P < 0.05$) in germplasm from Afghanistan. The collection displayed a highly variable number of capitula per plant (average 20.5, and range from 13 to 55) with significant differences detected among variants of seven categorical traits. Capitulum diameter averaged 20.9 mm, and ranged from 10 to 33 mm. Significant differences in capitulum diameter were found among variants of 11 categorical traits. Capitulum diameter was positively and significantly ($P < 0.05$) correlated with rosette period ($r = 0.20$), days to maturity ($r = 0.37$) and plant height ($r = 0.32$). However, it was negatively and significantly ($P < 0.05$, $n = 591$) correlated with number of capitula per plant ($r = -0.28$), biological yield ($r = -0.10$) and seed yield

Table 3. Mean polymorphic diversity indices, number of pairwise significant differences among means of eight quantitative trait, along with results of the Kruskal–Wallis test among regions and populations of safflower germplasm from 11 countries in the Middle East.

Source of variation (categorical trait)	Diversity index (*I*)	Number of pairwise significant differences among categorical trait means								*P*-Kruskal–Wallis test among	
		Rosette period, days	Days to maturity	Plant height, cm	Capitula per plant	Capitulum diameter, mm	Biological yield per plant, g	Seed yield per plant, g	Hull percent	Regions	Populations
Rosette period	0.84	0	1	0	2	2	2	0	0	0.125	0.025
Days to maturity	0.83	2	0	2	0	2	2	2	1	0.563	0.024
Plant height	0.92	1	2	0	0	0	0	0	0	0.294	0.050
Branches on main stem	0.84	0	0	2	0	1	0	0	0	0.042	0.050
Branch angle	0.96	1	0	0	1	0	1	1	0	0.000	0.001
Leaf shape	0.79	0	0	1	1	0	0	0	0	0.620	0.001
Leaf margin	0.78	1	2	1	0	1	0	1	0	0.000	0.000
Spininess	1.00	1	2	2	1	1	1	1	0	0.000	0.000
Capitulum shape	0.54	2	2	2	0	1	1	0	0	0.112	0.050
Capitulum diameter	0.93	2	3	2	2	0	2	1	0	0.953	0.000
Corolla color	1.00	0	1	1	0	1	0	0	0	0.250	0.029
Seed shape	0.12	0	0	0	0	1	0	0	1	0.414	0.430
Biological yield per plant	0.88	0	2	0	0	2	0	0	0	0.000	0.000
Hull percent	0.84	0	1	0	1	0	0	0	0	0.035	0.030
Capitula per plant	0.87	0	2	0	0	2	0	3	1	0.002	0.002
Seed yield per plant	0.92	0	2	0	3	2	0	0	0	0.417	0.000
Mean (minimum–maximum)		17.9 (10–27)	119.2 (108–139)	75.5 (28–135)	20.5 (13–55)	20.9 (10–33)	63.6 (38–130)	29.3 (19–73)	51.8 (25–67)		
Number of categorical traits		7	11	8	7	11	6	6	3		

($r = -0.12$). Biological and seed yield per plant averaged 63.6 and 29.3 g per plant, and ranged from 38 to 130 and from 19 to 73 g, respectively with significant differences among means of six categorical traits. Finally, hull percent was the least to vary among variants of categorical traits, with a mean of 51.8%, and a range from 25% to 67%.

Phenotypic diversity in agronomic traits

Means and standard deviations of phenotypic diversity indices (I) of seven agronomically important traits, along with the level of differentiation (G_{ST}) and frequency of the desirable variant (f) of these traits are presented in Table 4. Phenotypic diversity indices, in decreasing order, were displayed by spininess (0.93 ± 0.17), seed yield per plant (0.85 ± 0.17), capitula per plant (0.81 ± 0.15), hull percent (0.79 ± 0.18), rosette period (0.72 ± 0.20), biological yield per plant (0.65 ± 0.26) and days to maturity (0.61 ± 0.32); the last trait being the most, whereas the first was the least variable, with coefficients of variation of 52.5% and 18.3%, respectively.

Frequency of the desired variant for the same traits ranged from a low of 13% for early maturity to a high of 50% for "no or few spines". Most traits displayed high standard deviations, and consequently high coefficients of variation in this germplasm collection, however, the variant "no or few spines" displayed the lowest (46%) coefficient of variation across the germplasm collection. Population differentiation (G_{ST}) in the germplasm collection was highest for number of capitula per plant (30%) and seed yield per plant (27%); the remaining traits displayed much lower (13–18%) G_{ST} values.

Phenotypic diversity at the regional level

Apart from rosette period and, to some extent, hull percent, regional estimates of the phenotypic diversity indices for the eight agronomic traits did not differ significantly from each other (Table 4). Southwest Asian germplasm displayed high phenotypic diversity indices for rosette period (21% above average), and for hull percent (14% above average). Germplasm from Central Asia displayed higher than average frequencies for the variants high seed yield per plant and high number of capitula per plant, with populations from Afghanistan having the highest frequencies (0.62 and 0.47, respectively). Southwest Asian germplasm displayed higher than average frequencies for long rosette period (28.6% higher than average), high biological yield per plant (65% higher than average) and low hull percent (79% higher than average), with germplasm from Syria, and from Jordan and Syria, exhibiting the highest frequencies for the variants, respectively. Finally, germplasm from Africa displayed above average frequencies for the variants "no or a few spines" (38% above average), and "early maturity" (31% above average); Ethiopian germplasm furnished the highest frequencies for both variants.

Low levels of population differentiation (G_{ST}) in the germplasm from Central Asia were displayed for days to maturity and hull percent (both 20% lower than their respective average). Southwest Asian germplasm displayed lower than average G_{ST} values for spininess (44% lower than average) and days to maturity (20% lower than average). African germplasm was the least differentiated for rosette period (45% lower than average), biological and seed yield (50% and 44%, lower than their respective average), number of capitula per plant (60% lower than average) and hull percent (27% lower than average).

Phenotypic diversity at the population level

Germplasm from Iran and Syria displayed the highest I estimates for rosette period, however, Syrian germplasm displayed a high frequency (33%) of the long rosette period variant. On the other hand, Jordanian germplasm displayed the highest level of differentiation (0.44), followed by germplasm from Afghanistan (0.22) and Syria (0.18). Phenotypic diversity indices for spininess were high (>0.90) in seven countries, mostly in Central and Southwest Asia. However, frequencies of variants with no or few spines were highest in populations from three African (Ethiopia, Morocco and Sudan) and one Central Asian country (Pakistan). Nevertheless, the highest levels of differentiation for this trait was found, in decreasing order, in populations from Afghanistan ($G_{ST} = 0.41$), Iran (0.29) and Jordan (0.24), as compared with an overall average of 0.18. Phenotypic

Table 4. Phenotypic diversity indices for seven traits (I) and frequency of desirable variant of each trait (f) along with level of population differentiation (G_{ST}) for each trait in a safflower germplasm collection from 11 countries in three regions (Central Asia, Southwest Asia and Africa) of the Middle East.

	Desirable variant of the trait																				
	Long rosette period			No or few spines			Early maturity			High biological yield per plant			High seed yield per plant			High number of capitula per plant			Low hull percent		
Country	I	f	G_{ST}	I	f	G_{ST}	I	f	G_{ST}	I	f	G_{ST}	I	f	G_{ST}	I	f	G_{ST}	I	f	G_{ST}
Central Asia	0.73	0.14	0.10	0.94	0.42	0.30	0.66	0.14	0.12	0.71	0.05	0.20	0.84	0.28	0.27	0.84	0.26	0.31	0.86	0.12	0.12
Afghanistan	0.49	0.19	0.22	0.97	0.57	0.41	0.19	0.0	0.10	0.83	0.04	0.40	0.87	0.62	0.40	0.94	0.47	0.35	0.86	0.14	0.12
Iran	0.93	0.13	0.02	0.98	0.22	0.29	0.78	0.16	0.13	0.83	0.06	0.08	0.77	0.16	0.20	0.92	0.23	0.39	0.86	0.07	0.09
Pakistan	0.76	0.11	0.06	0.94	0.67	0.19	0.96	0.25	0.13	0.47	0.04	0.11	0.87	0.07	0.22	0.67	0.09	0.18	0.87	0.14	0.14
SW Asia	0.87	0.18	0.19	0.94	0.38	0.10	0.60	0.08	0.12	0.70	0.51	0.14	0.90	0.18	0.27	0.85	0.15	0.31	0.90	0.25	0.15
Israel	0.82	0.15	0.09	0.99	0.48	0.08	0.51	0.00	0.19	0.86	0.54	0.23	0.84	0.12	0.49	0.86	0.18	0.73	0.79	0.09	0.15
Jordan	0.94	0.10	0.44	0.80	0.20	0.24	0.33	0.10	0.14	0.63	0.80	0.28	0.93	0.30	0.22	0.80	0.20	0.29	1.00	0.20	0.33
Syria	0.93	0.33	0.18	0.98	0.45	0.02	0.73	0.05	0.07	0.69	0.56	0.01	0.82	0.06	0.08	0.94	0.11	0.07	0.93	0.33	0.03
Turkey	0.79	0.14	0.06	0.97	0.38	0.05	0.82	0.18	0.07	0.57	0.11	0.04	0.95	0.19	0.26	0.75	0.09	0.13	0.83	0.11	0.04
Africa	0.56	0.09	0.07	0.79	0.69	0.11	0.59	0.17	0.17	0.59	0.30	0.09	0.79	0.12	0.15	0.77	0.13	0.12	0.64	0.11	0.11
Egypt	0.70	0.09	0.13	0.95	0.56	0.12	0.76	0.06	0.22	0.93	0.53	0.13	0.98	0.18	0.12	0.89	0.09	0.20	0.65	0.08	0.29
Ethiopia	0.71	0.07	0.09	0.80	0.75	0.11	0.99	0.35	0.25	0.70	0.04	0.07	0.74	0.04	0.32	0.95	0.14	0.05	0.50	0.07	0.05
Morocco	0.45	0.17	0.03	0.63	0.67	0.12	0.00	0.00	0.07	0.00	0.00	0.02	0.45	0.00	0.04	0.45	0.17	0.11	0.45	0.05	0.05
Sudan	0.37	0.00	0.02	0.74	0.75	0.09	0.56	0.25	0.14	0.66	0.63	0.11	0.94	0.25	0.10	0.74	0.12	0.07	0.90	0.25	0.03
Mean	0.72	0.14	0.13	0.93	0.50	0.18	0.61	0.13	0.15	0.65	0.31	0.18	0.85	0.18	0.27	0.81	0.17	0.30	0.79	0.14	0.15
Standard deviation	0.20	0.08		0.17	0.23		0.32	0.12		0.26	0.30		0.17	0.17		0.15	0.11		0.18	0.09	

diversity index for days to maturity averaged 0.61 with a high standard deviation (0.32). It was monomorphic and highly polymorphic in Moroccan and Pakistani germplasm, respectively. Above average (>0.13) frequencies for early maturity, as a desirable variant, were found in germplasm from Ethiopia, Pakistan and Sudan, however, only Ethiopian and Egyptian germplasm exhibited a higher than average level of population differentiation ($G_{ST} = 0.25$ and 0.22, respectively) for days to maturity. Number of capitula per plant, as a major seed yield component, displayed high polymorphic diversity, especially in germplasm from Ethiopia, Afghanistan, Syria and Pakistan. Frequency of the variant "high number of capitula per plant averaged 0.17 and ranged from 0.09 in populations from each of Pakistan, Turkey and Egypt, to a high of 0.47 in populations from Afghanistan. The highest level of population differentiation for this trait was found in germplasm from Israel ($G_{ST} = 0.73$), followed by Iran ($G_{ST} = 0.39$) and Afghanistan ($G_{ST} = 0.35$). Biological and seed yield per plant displayed medium ($I = 0.65$) and high ($I = 0.85$) phenotypic diversity indices, respectively, with the Egyptian germplasm being highly polymorphic for both traits. Frequencies for high biological and seed yield per plant did not correspond across countries, however, germplasm from Jordan and Sudan displayed comparably high polymorphic indices for biological (0.63 and 0.66, respectively) and seed yield (0.93 and 0.94, respectively). Levels of population differentiation for biological yield were highest in germplasm from Afghanistan (0.40), Jordan (0.28) and Israel (0.23). Germplasm from Israel displayed the highest level of population differentiation for seed yield per plant (0.49), followed by germplasm from Afghanistan (0.40), and from Ethiopia (0.32). Hull percentage was highly variable, with an average polymorphic index of 0.79 (range from 0.45 to 1.00). The germplasm from Jordan and Syria was highly polymorphic for this trait, and the Syrian germplasm exhibited the highest frequency (0.33) for the low hull percent variant in this germplasm collection, albeit with a very low level of population differentiation (0.03). The highly polymorphic germplasm from Jordan displayed above average frequency (0.20) for the low hull percent and exhibited the highest level of population differentiation (0.33) for this trait.

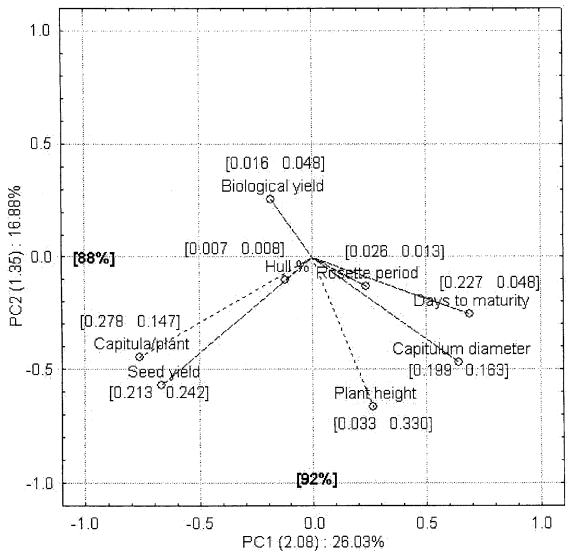

Figure 1. Principal components analysis plot for eight traits (and their contribution) to total variation in PC1 (right-side) and PC2 (left-side) in a germplasm collection of safflower from 11 countries in the Middle East. Bold face figures are percent total variation explained by traits in each of PC1 and PC2.

Multivariate structure based on quantitative traits

Principal component analyses were used to provide a reduced dimension model that would indicate measured differences among germplasm collections from different geographical regions. A minimum of four principal components were necessary to explain ~80.0% of total variation in the whole germplasm collection and in germplasm from each geographical region. The first and second principal components (Figure 1) accounted for 26.03% and 16.88% of total variation in the whole germplasm collection, respectively. Four traits, in decreasing order (number of capitula per plant, days to maturity, seed yield per plant and capitulum diameter) explained 92% of total variation in PC1, whereas plant height, seed yield per plant, capitula diameter and number of capitula per plant, in decreasing order, explained 88% of total variation in PC2. Hull percent, rosette period and biological yield per plant, in increasing magnitude and as indicated by their eigenvectors, had the lowest loadings on both PCs.

Accessions with a long rosette period tend to have short ($r = -0.08$, $P < 0.05$), late maturing

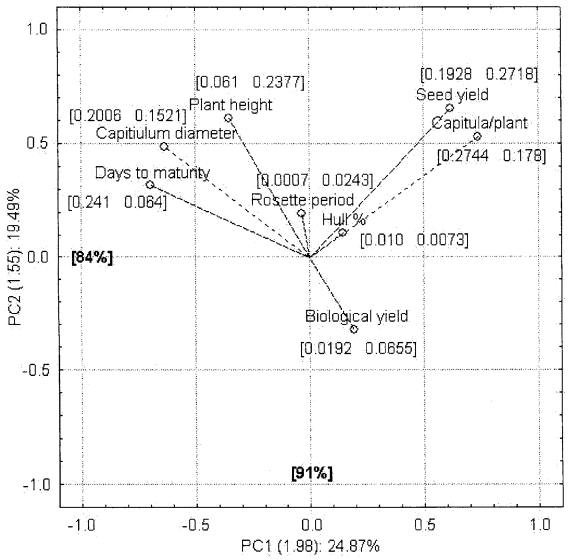

Figure 2. Principal components analysis plot for eight traits (and their contribution) to total variation in PC1 (right-side) and PC2 (left-side) in a germplasm collection of safflower from three countries in Central Asia. Bold face figures are percent total variation explained by traits in each of PC1 and PC2.

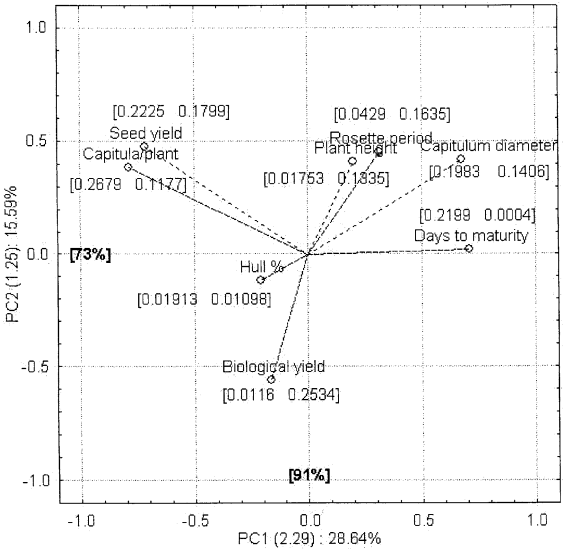

Figure 3. Principal components analysis plot for eight traits (and their contribution) to total variation in PC1 (right-side) and PC2 (left-side) in a germplasm collection of safflower from four countries in Southwest Asia. Bold face figures are percent total variation explained by traits in each of PC1 and PC2.

($r = 0.10$, $P < 0.05$) plants with large capitulum diameter ($r = 0.20$, $P < 0.05$) and relatively low biological yield ($r = -0.09$, $P < 0.05$). Late maturing accessions usually have tall plants ($r = 0.21$, $P < 0.05$), with large capitulum diameter ($r = 0.37$, $P < 0.01$) and lower number of capitula per plant ($r = -0.31$, $P < 0.01$). Accessions with high biological yield are characterized by short rosette period ($r = -0.09$, $P < 0.05$), short days to maturity ($r = -0.15$, $P < 0.05$) and small capitula ($r = -0.10$, $P < 0.05$). Forty percent of variation in seed yield was accounted for by variation in number of capitula per plant ($r = 0.64$, $P < 0.001$), however, seed yield was negatively impacted by longer days to maturity ($r = -0.23$, $P < 0.05$) and by large capitulum diameter ($r = -0.12$, $P < 0.05$).

Two principal components accounted for 44.36% of total variation in the germplasm collection from Central Asia (Figure 2). Capitula per plant, days to maturity, capitulum diameter and seed yield per plant, in decreasing order, explained 91% of variation in PC1, whereas seed yield per plant, plant height, number of capitula per plant and capitulum diameter, in decreasing order, explained 84% of variation in PC2. Accessions with high seed yield potential are characterized by having a large number of capitula per plant ($r = 0.68$, $P < 0.001$), slightly taller plants ($r = 0.13$, $P < 0.05$) and short days to maturity ($r = -0.18$, $P < 0.05$). A long rosette period is a characteristic of accessions with short plants ($r = -0.14$, $P < 0.05$), especially those with large capitulum diameter ($r = 0.14$, $P < 0.05$), however, such a long rosette period adversely impacted biological yield ($r = -0.12$, $P < 0.05$).

A more diverse pattern of multivariate structure was found in the germplasm from Southwest Asia (Figure 3). Two principal components accounted for 44.23% of total variation. Capitula per plant, seed yield, days to maturity and capitulum diameter, in decreasing order, explained 91% of variation in PC1, whereas biological yield, seed yield, rosette period and plant height, in decreasing order, explained 73% of variation in PC2. High seed yield was associated with a large number of capitula per plant ($r = 0.64$, $P < 0.001$), short maturity period ($r = -0.34$, $P < 0.01$) and smaller capitulum diameter ($r = -0.20$, $P < 0.01$), the latter was also negatively correlated with number of capitula per plant ($r = -0.32$, $P < 0.01$) and with hull

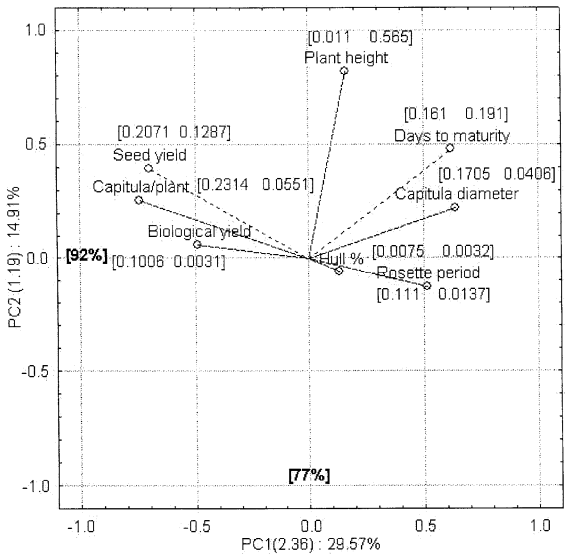

Figure 4. Principal components analysis plot for eight traits (and their contribution) to total variation in PC1 (right-side) and PC2 (left-side) in a germplasm collection of safflower from four countries in Africa. Bold face figures are percent total variation explained by traits in each of PC1 and PC2.

percent ($r = -0.16$, $P < 0.05$). Positive correlations were found among capitulum diameter and rosette period ($r = 0.26$, $P < 0.01$), days to maturity ($r = 0.40$, $P < 0.001$) and plant height ($r = 0.24$, $P < 0.01$), however, late maturing accessions and accessions with high biological yield tend to have smaller number of capitula per plant ($r = -0.41$, $P < 0.001$ and $r = -0.14$, $P < 0.05$, respectively).

The African germplasm displayed a different pattern of trait association (Figure 4), with capitula per plant, plant height, seed yield and days to maturity accounting for 44.48% of total variation. Capituala per plant, seed yield per plant, capitulum diameter and days to maturity, in decreasing order, explained 77% of total variation in PC1, whereas, plant height, days to maturity, seed yield per plant and capitulum diameter, in decreasing order, explained 92% of variation in PC2. Seed yield was positively associated with number of capitula per plant ($r = 0.58$, $P < 0.01$), biological yield ($r = 0.32$, $P < 0.01$), and negatively with rosette period ($r = -0.20$, $P < 0.05$), days to maturity ($r = -0.19$, $P < 0.05$) and capitulum diameter ($r = -0.21$, $P < 0.05$). In addition to its negative impact on seed yield, a longer rosette period was positively correlated with capitulum diameter ($r = 0.32$, $P < 0.01$), however, a large number of capitula per plant tends to shorten the rosette period ($r = -0.24$, $P < 0.05$). Finally, late maturing accessions in the African germplasm have tall plants ($r = 0.25$, $P < 0.05$), with large capitulum diameter ($r = 0.33$, $P < 0.01$), low biological yield ($r = -0.32$, $P < 0.01$) and low number of capitula per plant ($r = -0.31$, $P < 0.01$).

Multivariate structure based on qualitative traits

Test of independence of pairs of categorical traits suggest that 35, of the 120 contingency chi-square tests involving 16 categorical traits, are independent (data not presented). Eleven of the 35 independent associations involved the trait hull percent, whereas rosette period was involved in nine other independent associations. Non-random associations among different groupings of nine categorical traits are presented in Table 5 as log-linear models for the whole collection, and for each geographical region in the Middle East. Six different log-linear models were developed for the whole germplasm collection and included a maximum of three-trait terms. A variable number of models and trait combinations, including one-, two- and three-trait terms, were developed for each geographical region separately. Capitulum shape followed by spininess and rosette period were the most interactive qualitative traits, and were involved in 20, 19 and 15 trait combinations, respectively. The first and second log-linear models involving three categorical traits with a three-term combination each, were identical for populations from the Middle East and Central Asia, however, two-term trait combination models fit the data for populations from Southwest Asia and Africa. The first model involved rosette period, capitulum shape and spininess, whereas the second involved rosette period, days to maturity and capitulum shape. The third log-linear model for the Middle East involved capitulum shape, capiulum diameter and corolla color; three two-trait terms fit the data for the whole germplasm collection, however, two two-term combinations fit the data for Central Asian populations, and one two-term combination and a single trait model fit the data for Southwest Asian populations. Data for these traits did not fit any model for the African germplasm. The fourth model involved plant height, spininess and branch angle

Table 5. Log-linear models based on categorical traits scored on safflower germplasm collection from 11 countries in the Middle East.

Log-linear model	Region	Model terms			Likelihood ratio statistic (G^2)	d.f.	*P*
1	Middle East	[RP × SP]	[CS × RP]	[CS × SP]	12.2	16	0.73
	Central Asia	[RP × SP]	[CS × RP]	[CS × SP]	11.3	16	0.79
	Southwest Asia	[RP × SP]		[CS × SP]	8.3	16	0.93
	Africa	[RP × CS]		[CS × SP]	13.4	18	0.76
2	Middle East	[RP × DM]	[RP × CS]	[CS × DM]	9.07	8	0.34
	Central Asia	[RP × DM]	[RP × CS]	[CS × DM]	7.9	8	0.93
	Southwest Asia	[RP × DM]	[RP × CS]		13.2	12	0.35
	Africa	[RP × DM]	[RP × CS]		12.2	12	0.43
3	Middle East	[CS × CD]	[CC × CS]	[CC × CD]	23.3	20	0.28
	Central Asia	[CS × CD]	[CC × CS]		20.5	24	0.67
	Southwest Asia Africa		[CC × CS]	[CD]	27.4	28	0.49
4	Middle East	[PH × SP]	[AB × PH]	[SP × PH]	8.3	24	0.99
	Central Asia	[PH × SP]	[AB × SP]		18.8	30	0.94
	Southwest Asia	[PH × SP]	[AB]		27.0	33	0.76
	Africa	[PH × SP]	[AB]		42.3	39	0.33
5	Middle East	[AB × SP]	[LB × SP]	[AB × SP]	13.8	24	0.95
	Central Asia	[AB × SP]	[LB × SP]	[AB × LB]	20.2	24	0.69
	Southwest Asia Africa		[LB × SP]	[AB × LB]	14.6	27	0.97
6	Middle East	[RP × CS]	[CC × CS]		19.6	30	0.99

on main stem. Data for populations from Southwest Asia and Africa fit identical models, whereas the model for populations from Southwest Asia was different and differed from the model that fit data for the whole germplasm collection. The fifth model involved branch angle on main stem, spininess and location of branches on main stem. Log-linear models describing the data for the whole germplasm collection and for Central Asian populations were identical except for one trait. A reduced model (two two-term model) fit the data for Southwest Asian populations, however, data for the same traits recorded on African populations did not fit any model. Finally, a model with two two-trait combinations (rosette period, capitulum shape and corolla color) fit the data for the whole germplasm collection but not for any of the individual geographical regions.

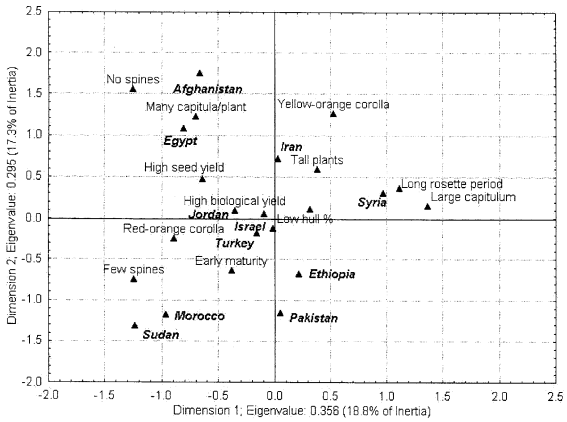

Figure 5. Correspondence analysis plot of populations and desirable variants of agronomic traits in a germplasm collection of safflower from the Middle East.

Correspondence analysis

Each dimension of the correspondence analysis separated the 11 populations and 12 desirable variants of eight agronomic traits into two groups (Figure 5). The first dimension accounted for 18.8% of inertia, with an eigenvalue of 0.356, and separated germplasm from Iran, Syria, Ethiopia and Pakistan, along with desirable variants of five agronomic traits from the remaining populations and variants. The second dimension, accounted for 17.3% of inertia with an eigenvalue of 0.295, and separated germplasm from Afghanistan, Egypt, Iran, Syria and Jordan, along with eight desirable variants of agronomic traits from the rest. Germplasm from Afghanistan, Egypt and Jordan were associated with four desirable variants of

agronomic traits (no spines, many capitula per plant, high seed yield and high biological yield per plant). Germplasm from Iran and Syria were associated with four desirable variants of agronomic traits (yellow–orange corolla color, tall plants, long rosette period, large capitulum and low hull percent). Germplasm from Israel and Turkey, and to a lesser extent germplasm from Morocco and Sudan, were closely associated with few spines, red–orange corolla color and early maturity. Finally, germplasm form Ethiopia and Pakistan was grouped at a farther distance together with germplasm from Turkey, Israel, Morocco and Sudan, and was loosely associated with the same variants (i.e., few spines, red–orange corolla color and early maturity).

Distances among regions and among populations

Genetic diversity (GD) distances (Table 6) based on frequencies of eight categorical traits (above diagonal) and SM distances (below diagonal) based on eight quantitative traits among safflower populations from 11 Middle Eastern countries reflect different relationships among germplasm populations. Mean GD between any two populations was 0.11, whereas mean SM distance was 6.98. SM pairwise distances separated more populations from each other than did GD distances. A higher percentage (45.5%) of SM pairwise distance was higher than the mean SM distance; the respective value for GD distances was 36.4%.

The correlation between GD and SM distances among countries was negative, but not significant ($r = -0.24$, $P < 0.09$). Genetic distances among populations did not correspond with the geographical location of their country of origin. On average, the most distant populations, based on genetic distance estimates, were those from Jordan, Ethiopia, Morocco and Sudan, whereas the least distant were those from Egypt, Pakistan and Israel. Populations from Egypt and Turkey were closest (GD = 0.03), whereas populations from Ethiopia and Turkey were the most distant (GD = 0.52). A wide range of SM distances were found among germplasm populations (Table 6, below diagonal). SM distances ranged from a non-significant 0.3 between populations from Israel and Syria, to 23.0 ($P < 0.001$) between populations from Iran and Egypt. Mean GD distance among and within regions were 0.126 and 0.102, respectively. The respective mean SM distances were 8.53 and 5.81. Populations from Southwest Asia and Africa were most distant (GD = 0.18), followed by populations from Central Asia and Africa (GD = 0.109), and then by populations from Central and Southwest Asia (GD = 0.088). The respective SM values for the same pairwise distances were 7.5, 9.2 and 8.9. Mean GD distances within populations from Central Asia and Africa were equal (0.09), however, GD within populations from Southwest Asia (0.12) was 33% higher. SM distance within populations from Central Asia (9.2) was more than twice the SM distances within populations from Southwest Asia (SM = 4.2), and from Africa (SM = 4.02).

Discriminant analysis among regions

Discriminating power of five phenotypic traits (rosette period, leaf shape, branch angle, number of capitula per plant and biological yield per plant) was strong enough to differentiate among populations from the three geographical regions in the Middle East. Based on partial Wilk's lambda (λ), biological yield contributed most to the overall discrimination ($\lambda = 0.88$), followed by number of capitula per plant ($\lambda = 0.97$), and then by the remaining three traits ($\lambda = 0.98$). Slightly more than half (52.9%) the accessions from the three geographical regions were correctly classified based on discriminant analysis between regions using all 16 traits; however, a higher percentage (81.3%) of Central Asian accessions was correctly classified, followed by accessions from Southwest Asia (46.8%). Less than one-third (30.7%) of accessions from Africa was correctly classified (Figure 6).

Discriminant analysis among populations

Ten phenotypic traits were powerful enough to discriminate among countries of origin; four of these traits (rosette period, branch angle, number of capitula per plant and biological yield per plant) contributed to the discrimination among regions. Biological yield per plant contributed most to the discrimination among countries ($\lambda = 0.72$). The remaining nine traits were grouped, based on the magnitude of their λ values, into four groups. The first group consisted of capitulum diameter

Table 6. Genetic distances based on frequencies of eight categorical traits (above diagonal), and SM distances (below diagonal) based on eight quantitative traits among safflower germplasm collections from 11 countries in the Middle East.

	Central Asia			Southwest Asia				Africa				
	Afghanistan	Iran	Pakistan	Israel	Jordan	Syria	Turkey	Egypt	Ethiopia	Morocco	Sudan	Mean SM distance
Afghanistan		0.06	0.11	0.07	0.15	0.12	0.05	0.08	0.11	0.16	0.16	7.41[c#]
Iran	4.9*		0.11	0.05	0.08	0.09	0.03	0.08	0.12	0.15	0.20	12.35a
Pakistan	9.9**	12.9***		0.07	0.18	0.11	0.05	0.07	0.04	0.07	0.07	9.76b
Israel	9.0**	12.7***	8.8**		0.16	0.15	0.12	0.02	0.08	0.12	0.09	5.98d
Jordan	8.6**	9.6**	7.8**	2.2 n.s.		0.09	0.11	0.08	0.20	0.20	0.16	6.16d
Syria	9.4**	10.6**	8.8**	0.3 n.s.	3.6*		0.11	0.04	0.11	0.14	0.15	6.36d
Turkey	4.5*	8.8**	8.3**	5.9*	6.3*	6.9*		0.03	0.52	0.09	0.12	6.20d
Egypt	5.3*	23.0***	15.9***	1.91 n.s.	2.7 n.s.	3.9*	10.0**		0.06	0.13	0.09	5.55e
Ethiopia	5.3*	4.78*	11.5***	6.3*	7.9**	6.8*	3.1*	8.9**		0.10	0.08	5.94d
Morocco	12.5***	7.9*	10.8**	8.4**	10.4***	8.8**	5.1*	4.1*	1.9 n.s.		0.11	7.38c
Sudan	4.7*	5.5*	2.9 n.s.	4.3*	2.5 n.s.	4.5*	3.1*	2.4 n.s.	2.9 n.s.	3.9*		3.67f
Mean GD distance	0.11b[c#]	0.10bc	0.09c	0.09c	0.14a	0.11b	0.12ab	0.07d	0.14a	0.13a	0.12ab	

*,** Significant at the 5% and 1% levels of probability; n.s., not significant.

[#] SM or GD mean distances followed by the same letter do not differ significantly ($P < 0.05$).

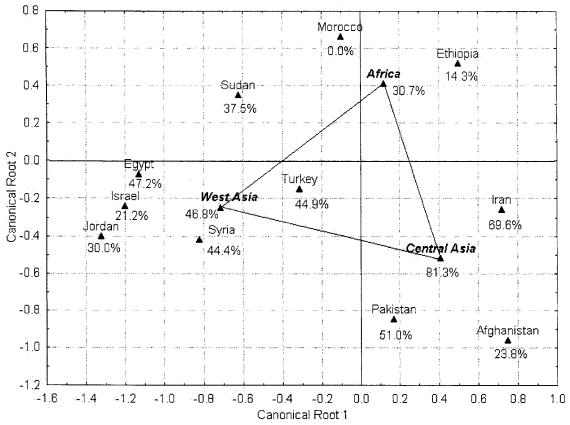

Figure 6. Canonical discriminant analysis and relative distances plot of three geographical regions and 11 countries of origin, and percent correct classification of a safflower germplasm collection from the Middle East.

with $\lambda = 0.89$. The second group consisted of plant height, branch angle and seed yield per plant ($\lambda = 0.92$). The third group consisted of rosette period and spininess ($\lambda = 0.95$). The fourth, and last group, consisted of days to maturity, leaf margin and number of capitula per plant ($\lambda = 0.98$). Percent correct classification (Figure 6) of accessions from each of the 11 countries in the Middle East ranged from a low of 0.0 for Moroccan germplasm to a high of 69.6 for Iranian germplasm. Most African (~70%) and almost half of Southwest Asian (~53%) germplasm were misclassified as being from Central Asia, with Iran, Afghanistan and Pakistan, in decreasing order, implied as the country of origin of the misclassified accessions.

Selected germplasm

A multitrait selection procedure resulted in identifying 87 accessions (14.7% of total germplasm collection) with higher biological (45%) and seed yield (23.9%), taller plants (4.5%), higher number of capitula per plant (4.6%) and lower hull percent (4.8%) as compared with the whole germplasm collection (Table 7). Populations from Central Asian, Southwest Asian and African countries contributed 11, 40 and 36 accessions to the selected germplasm, respectively. Egyptian germplasm exhibited the highest frequency (35.63%) for the desirable trait combinations, followed by germplasm from Israel (19.54%), Turkey (11.49%) and Iran (10.34%). Another selection procedure was carried out to identify germplasm accessions with long rosette period. We identified 76 accessions (12.86% of total collection) with longer rosette period (28.0%), higher number of capitula per plant (15.0%), higher biological yield (14.0%) and slightly higher days to maturity (5.0%) than the whole germplasm collection. Populations from Central and Southwest Asian countries contributed the highest number of accessions (36 and 30, respectively), whereas, a small number (10 accessions) were contributed by populations from African countries. Germplasm from Iran (21 accessions), Turkey (18 accessions) and Pakistan (11 accessions) contributed most (65.8%) accessions in this subset. Finally, we combined both selection criteria and were able to identify five accessions (one from each of Israel, Jordan and Turkey, and two from Syria) having desirable variants of both sets of traits.

Discussion

During its process of domestication, cultivated safflower was dispersed beyond its center of origin and was compelled to adapt to different ecological conditions (Harlan 1992). Adaptation of the species to the wide spatial and temporal variation in the Old World resulted in a multitude of morphological and physiological ecotypes (Knowles (1989) and in enormous amount of local variation of the domesticated crop (Li and Mündel 1996). The high levels of highly variable phenotypic diversity indices (coefficients of variation ranged from 33% to 47%) among regions (mean and standard deviation $I = 0.81 \pm 0.27$)) and among populations (I ranged from 0.56 to 0.78) detected in safflower germplasm from the Middle East indicate the presence of reasonably balanced frequency classes for individual traits, and a high level of genetic diversity (Polignano et al. 1999).

Estimates of total "functional genetic diversity" (based on morphological traits deliberately selected by farmers) and its components are in line with values cited for other (Hernandez-Verdugo et al. 2001; Upadhyaya 2003) landraces of domesticated crops. Our results indicate that 79% and 21% of total diversity were partitioned within and among populations, respectively. Although a relatively high G_{ST} estimate was found for the whole

Table 7. Frequency of germplasm accessions selected.

Region	Country	High biological (>63.6 g/plant) and seed yield (>29.3 g/plant)[a]		Long rosette period (>17.9 days)[b]		Combined selections
		Number	Frequency	Number	Frequency	Number[c]
Central Asia	Afghanistan	1	1.15	4	5.26	
	Iran	9	10.34	21	27.63	
	Pakistan	1	1.15	11	14.47	
	Subtotal	11	12.64	36	47.37	
Southwest Asia	Israel	17	19.54	5	6.58	1
	Jordan	6	6.90	1	1.32	1
	Syria	7	8.05	6	7.89	2
	Turkey	10	11.49	18	23.68	1
	Subtotal	40	45.98	30	39.47	5
Africa	Egypt	31	35.63	7	9.21	
	Ethiopia	1	1.15	2	2.63	
	Morocco	0.0	0.0	1	1.32	
	Sudan	4	4.60	0.	0.0	
	Subtotal	36	41.38	10	13.16	
Middle East	Number of accessions	87		76		5

[a]High biological and seed yield.
[b]Long rosette period.
[c]Both sets of traits combined.

germplasm collection (0.21), regional values were relatively smaller (0.07 to 0.10); however, three populations (Afghanistan, Jordan and Israel) displayed relatively high levels of population differentiation, whereas germplasm from Turkey and Syria partitioned most (~90%) of their diversity within populations. The large differences in the level of population differentiation among populations across the Middle East may reflect different histories of these populations and the consequent different allocation of genetic variation among and within populations (Lahane et al. 1999; Hernandez-Verdugo et al. 2001). The exceptionally high within population variation in this germplasm collection is expected to be maintained or even maximized by a combination of genetic, environmental (Patel et al. 1994) and management factors (Assefa et al. 1999) across its geographical region of adaptation.

Patterns of variation

The highly variable phenotypic structures in Middle Eastern safflower populations found in this study reflect their long-term response to strong spatial and temporal selective pressure (Ashri 1971; Abebe and Bjornstad 1996) and to farmers' deliberate targeted selection of preferred phenotypes (vom Brocke et al. 2003), thus leading to changes in their morphology and genetic structures. Changes in morphology are more perceptible in plant parts which are of human interest (Patel et al. 1989; Johnson et al. 2001) such as plant size, spininess, corolla color and capitulum shape, all of which displayed high levels of phenotypic polymorphism within and among populations. Likewise, high phenotypic and genetic variation was observed for seed yield and number of capitula per plant in safflower germplasm from Pakistan (Chowdhury et al. 1999) and for number of seed per plant, plant height and corolla color in germplasm from Iran (Bagheri et al. 2001).

Variation among agronomic traits

Morphological variation in safflower (Knowles 1969) and other domesticated crops (Rabbani et al. 1998; Upadhyaya 2003) of the Old World is the result of the combined natural and farmers' selection for certain phenotypes adapted to the prevailing climatic and edaphic conditions throughout the species' adaptation range. It can be postulated (e.g., Abebe and Bjornstad 1996; Alemayehu and Becker 2002) that selection pressure under

domestication in safflower and other crop species could have been similar in different geographical regions, and resulted in larger within population diversity compared to diversity across regions.

Plant height, being strongly associated with flowering (Johnson et al. 2001) and with days to maturity, post-anthesis period and anthocyanin (Kotecha 1979) was highly variable ($I = 0.92$) and ranged from 28 to 135 cm. There were significant differences among variants of plant height (i.e., short, medium and tall) in rosette period and days to maturity. However, plant height as a quantitative trait varied tremendously among variants of eight qualitative traits (Table 2), including four traits reported by Kotecha (1979) and by Johnson et al. (2001).

Branch angle, number and location of branches on main stem, and spininess displayed high levels of variability in this germplasm collection. If desirable variants of these traits are combined in one genotype, they could result in higher biological and seed yield provided dense stand can be maintained with high number of capitula per unit area (Lahane et al. 1999). High broad sense heritability was reported for number of branches per plant (Kavani et al. 2000) which is controlled by additive gene action (Ashri 1971). Ashri (1971) and Senapati et al. (1999) concluded that capitula per plant is the most important yield component; it appeared to be controlled by four groups of genes in a 10-parent incomplete diallel cross, with mainly non-additive gene action. Although high broad sense heritability was reported for capitula per plant, it is influenced by field management and environment (Kavani et al. 2000). Variation in number of capitula per plant accounted for about 40% of total variation in seed yield, except for African germplasm ($R^2 = 0.34$). It was influenced by variants of six categorical traits (rosette period, branch angle, leaf shape, spininess, capitulum diameter and hull percent), and variants with low, medium and high number of capitula per plant differed significantly in days to maturity (Table 3).

Sources of earliness or lateness were identified in germplasm from the same country of origin throughout the Middle East (Tables 3 and 4). Early maturing genotypes would compete well with other crops such as wheat and would permit double cropping and production of safflower on marginal lands. The wide range in days to maturity (108–139) offers great flexibility for the development of new varieties suitable for various agroecological zones and cropping systems (Assefa et al. 1999).

Duration of rosette period although genetically controlled (Kotecha 1979), is subjected to environmental factors, mainly temperature and daylength. The daylength- and temperature-insensitive accessions, identified earlier (Li and Mündel 1996) in germplasm from the Middle East, have a high potential for wide geographic adaptation. A high level of phenotypic diversity ($I = 0.84$) was found for rosette period (range 10–27 days), especially in germplasm from Syria (frequency of long rosette period = 33%) where florets are used for food coloring and as a source of vegetable dye. All corolla colors were represented in this germplasm collection ($I = 1.0$) and accessions with different corolla colors differed in days to maturity, plant height and capitulum diameter (Table 3). However, in search for specific corolla color types, it was possible to identify (Bradley et al. 1999) only four accessions from Iran with the yellow corolla color variant in the ornamental type of the world safflower collection. Accessions with high frequency (>50%) of spineless plants were found in germplasm from Afghanistan, Pakistan, and in germplasm from all four African countries. Spines are considered a handicap especially where manual harvest is involved (Chaudhry 1986; Li and Mündel 1996). Spinelessness is also an important phenotypic trait for the development of fresh-cut and dried flowers (Uher 1997).

Sources of reduced hull (minimum of 25%) were identified, in decreasing order, in safflower germplasm from Syria, Sudan and Jordan. High frequency of low hull percent variants in Syrian, Sudanese and Jordanian germplasm was associated with spineless plants, fewer capitula per plant and high biological yield. Variation in this trait was powerful enough to differentiate among regions and populations (Table 3). Well-developed hulls usually result in lower seed oil content. Accessions with reduced hull percentage (<30%) may produce higher oil (i.e., >50%) (Urie 1986).

Association among traits

Multivariate analyses when applied to the quantitative and qualitative traits at the population, regional or Middle Eastern levels, were successful

in identifying pronounced non-random associations among sets of traits at each level. However, not all traits, as was the case for *Brassica juncea* L. (Rabbani et al. 1998), *Eragrostis tef* (Assefa et al. 1999) and chickpea (Upadhyaya 2003) contributed substantially to the apparent variation in safflower germplasm from the Middle East. Number of capitula per plant contributed the most, whereas hull percent contributed the least to the variation in the whole collection and in germplasm from each of the three geographical regions.

The existence of broad morphological and phenotypic diversity in this germplasm collection is further substantiated by PC analyses. A minimum of four PCs (50% of the number of quantitative traits in the study) were required to explain about 80% of total variation, with only about 44% of total variation being explained by the first two PCs. Similar results were reported for *E. tef* from Ethiopia (Assefa et al. 1999) and a world germplasm collection of chickpea (*Cicer arietinum*) (Upadhyaya 2003). The different patterns of variation, as elucidated by PC analyses were also supported by the highly significant correlation coefficients among traits (Alemayehu and Becker 2002), and by the traits' variable loadings (Patel et al. 1989) on the first two PCs. Moreover, the standard two-way contingency chi-square method and log-linear models were successful in elucidating the organization of categorical traits and indicated that safflower has developed specific first- and second-order associations among phenotypic traits during its long history throughout the Middle East.

Genetic and phenotypic distances

Results of uni- and multi-variate analyses suggest that germplasm from Southwest Asia, the center of origin and a primary center of diversity, was most distant from Central Asian germplasm, thus confirming earlier findings based on a core collection (Johnson et al. 2001). However, contrary to earlier findings (Johnson et al. 2001) African germplasm was found to be more closely related to the Southwest Asian germplasm; this discrepancy could be attributed to a larger number of accessions (19.6%) and number of traits in this study as compared to the study by Johnson et al. (2001). Nevertheless, GD and SM distances (Table 6) suggest the presence of highly complex relationships among populations, both within and among regions.

Discrimination among regions and populations

Results of discriminant analyses among regions and populations suggest that the *a priori* assignment of populations to geographical regions or to countries within these regions did not correspond strictly with the phenotypic groupings. This could be attributed to historic (Ladizinsky 1998) or recent (Polignano et al. 1999) and intensive formal or informal germplasm exchange among farmers or users of germplasm that may have overlapped the previous diversity distribution patterns.

The loose association between morphological traits and geographical origin of germplasm accession supports earlier findings (Patel et al. 1989; Khan et al. 2003) and suggests that geographical isolation is not the only factor causing genetic diversity in safflower. However, genetic drift and natural selection forces under diverse environmental conditions within a country could cause considerable diversity as reported by Rabbani et al. (1998) for landrace *B. juncea* L. in Pakistan and by Assefa et al. (1999) for landrace *E. tef* in Ethiopia.

Conclusions

A germplasm collection of safflower from 11 countries in three geographical regions of the Middle East proved to be highly diverse for 16 quantitative and qualitative traits of agronomic value to breeders and end-users. The collection provides a large amount of variation for traits to contribute toward suitable plant morphology and canopy structure to breed new cultivars well adapted to different growing conditions, especially in the Persian Gulf. Although our results indicate that the geographical regions have played a crucial role in the species evolution process under domestication, the *a priori* classification of the countries of origin according to their geographical region does not strictly correspond to the phenotypic groupings of origin. This is probably the result of a recent and intensive germplasm exchange among germplasm users that may have overlapped the previous diversity distribution pattern of the domesticated species. A small number of accessions with a combination of

desirable traits for high biological and seed yield, or for floret production, were identified (Table 7), mostly in germplasm from Southwest Asia, the center of origin and a major center of diversity of safflower. It is expected that these accessions will contribute toward sustainable seed, forage and dye production under the hot climate and short growing season of the Persian Gulf.

References

Abebe D. and Bjornstad A. 1996. Genetic diversity in Ethiopian barley in relation to geographical regions, altitude range and agro-ecological zones as an aid to germplasm collection and conservation strategy. Hereditas 124: 17–29.

Agresti A. 1990. Categorical Data Analysis. Wiley International, New York, USA, pp. 558.

Alemayehu N. and Becker H. 2002. Genotypic diversity and patterns of variation in a germplasm material of Ethiopian mustard (*Brassica carinata* A. Braun). Genet. Resour. Crop Evol. 49: 573–582.

Ashri A. 1971. Evaluation of the world collection of safflower, *Carthamus tinctorius* L. I. Reaction to several diseases and associations with morphological characters in Israel. Crop Sci. 11: 253–257.

Assefa K., Ketama S., Tefera H., Nguyen H.T., Blum A., Ayele M., Bai G., Simane B. and Kefyalew T. 1999. Diversity among germplasm lines of the Ethiopian cereal tef [*Eragrostis tef* (Zucc.) Trotter]. Euphytica 106: 87–97.

Bagheri A., Yazdi-Samadi B., Taeb M. and Ahmadi M.R. 2001. A study of genetic diversity in landrace populations of safflower in Iran. Iran. J. Agric. Sci. 32: 447–456.

Bradley V.L., Guenthner R.L., Johnson R.C. and Hannan R.M. 1999. Evaluation of safflower germplasm for ornamental use. In: Janick J. (ed.), Perspectives on New Crops and New Uses, ASHS Press, Alexandria, VA, USA, pp. 433–435.

Chaudhry A.H. 1986. Evaluation and culture of sunflower and safflower in dobari lands of Sind. First Annual Report. PL480 Program of USDA, Project No. PK-ARS-226, Grant No. FG.PA 395, 25 p.

Chowdhury B., Mandal A.B. and Banerjee S.P. 1999. Assessment of variability and cause and effect relationships in safflower (*Carthamus tinctorius* L.). Ann. Agric. Res. 20: 278–281.

Harlan J. 1992. Crops and Man, 2nd edn. ASA, CSSA, Inc., Madison, WI, USA, 284 pp.

Hernandez-Verdugo S., Luna-Reyes R. and Oyama K. 2001. Genetic structure and differentiation of wild and domesticated populations of *Capsicum annuuum* (Solanaceae) from Mexico. Plant Syst. Evol. 226: 129–142.

Johnson R., Ghorpade P. and Bradley V. 2001. Evaluation of the USDA core safflower collection for seven quantitative traits. International Safflower Conference, July 23, 2001.

Kavani R.H., Shukla P.T. and Madariya R.B. 2000. Analysis of variability for seed yield and related characters in safflower (*Carthamus tinctorius* L.). Madras Agric. J. 87: 449–452.

Khan M.A., von Witzke-Ebrecht S., Maass B.L. and Becker H.C. 2003. Evaluation of a worldwide collection of safflower for morphological diversity and fatty acid composition. Deutscher Tropentag, October 8–10, 2003, Gőttingen, Technological and Institutional Innovations for Sustainable Rural Development.

Knowles P.F. 1969. Centers of plant diversity and conservation of crop germplasm: Safflower. Econ. Bot. 23: 324–329.

Knowles P.F. 1989. Safflower. In: Rébbelen G., Downey R.K. and Ashri A. (eds), Oil Crops of the World, McGraw-Hill, New York, USA, pp. 363–374.

Knowles P.F. and Ashri A. 1995. Safflower *Carthamus tinctorius* (Compositae). In: Smartt J. and Simmonds N.W. (eds), Evolution of Crop Plants, 2nd edn., Longman Scientific and Technical, pp. 47–50.

Kotecha A. 1979. Inheritance and association of six traits in safflower. Crop Sci. 19: 523–527.

Ladizinsky G. 1998. Plant Evolution Under Domestication. Kluwer, Dordrecht, pp. 254.

Lahane P.S., Mukewar A.M., Zope J.S., Kalpande H.V. and Kalpande V.V. 1999. Genetic variability for different traits in safflower (*Carthamus tinctorius* L.). J. Soil Crops 9: 130–132.

Li D. and Mündel H.H. 1996. Safflower, *Carthamus tinctorius* L. Promoting the conservation and use of underutilized and neglected crops, 7. Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany/International Plant Genetic Resources Institute, Rome, Italy, 83 pp.

Li D., Zhou M. and Rao R. (eds), 1993. Characterization and Evaluation of Safflower Germplasm. Geological Publishing House, Beijing, China, 260 pp.

National Plant Germplasm System (NPGS), 2000. Germplasm Resources Information Network (GRIN). Database Management Unit (DBMU), NPGS, USDA, Beltsville, Maryland.

Patel M.Z., Reddy M.V., Rana B.S. and Reddy B.J. 1989. Genetic divergence in safflower (*Carthamus tinctorius* L.). Indian J. Genet. Plant Breed. 49: 113–117.

Patel V.D., Reddy V.S. and Nerkar Y.S. 1994. Efficiency of early generation selections for yield and related characters in safflower (*Carthamus tinctorius* L.). Theor. Appl. Genet. 89: 293–296.

Polignano G.B., Elba E., Uggenti P. and Scippa G. 1999. Geographical patterns of variation in Bari faba bean germplasm collection. Genet. Resour. Crop Evol. 46: 183–192.

Rabbani M.A., Iwabuchi A., Murakami Y., Suzuki T. and Takayanagi K. 1998. Phenotypic variation and the relationships among mustard (*Brassica juncea* L.) germplasm from Pakistan. Euphytica 101: 357–366.

Rohlf F.J. 2000. NTSYSpc Numerical taxonomy and multivariate analysis system. Version 2.1 User guide 38p. Exeter Software.

Senapati N., Samal K.M., Mohanta I.C. and Dhal A. 1999. Performance, variability and character association in safflower (*Carthamus tinctorius* L.). Indian J. Agric. Res. 33: 254–258.

StatSoft Inc. 2001. STATISTICA (data analysis software systems), version 6, www.statsoft.com.

Uher J. 1997. Safflower in European agriculture. In: Corleto A. and Mündel H.-H. (eds), Safflower: A Multipurpose Species with Unexploited Potential and World Adaptability.

Proceedings of the Fourth International Safflower Conference, Bari, Italy, 2–7 June, 1997, Adriatica Editrice, Bari, Italy.

Upadhyaya H.D. 2003. Geographical patterns of variation for morphological and agronomic characteristics in the chickpea germplasm collection. Euphytica 132: 343–352.

Urie A.L. 1986. Inheritance of partial hull in safflower. Crop Sci. 26: 493–498.

vom Brocke K., Christnck A., Weltzien E., Presterl R.T. and Geiger H.H. 2003. Farmers' seed systems and management practices determine pearl millet genetic diversity patterns in semiarid regions of India. Crop Sci. 43: 1680–1689.

Yeh F.C., Yang R.-C., Boyle T.B.J., Ye Z.-H. and Mao J.X. 2000. POPGENE, the User-Friendly Shareware for Population Genetic Analysis. Molecular Biology and Biotechnology Center, University of Alberta, Canada.

Zar J.H. 1996. Biostatistical Analysis, 3rd edn. Prentice Hall, NJ, USA, p. 662.

Zhang Q. and Allard R.W. 1986. Sampling variance of the genetic diversity index. J. Hered. 77: 54–55.