

# Relationships among different geographical groups, agro-morphology, fatty acid composition and RAPD marker diversity in Safflower (*Carthamus tinctorius*)

Muhammad A. Khan · Sabine von Witzke-Ehbrecht ·  
Brigitte L. Maass · Heiko C. Becker

Received: 31 January 2008 / Accepted: 14 April 2008 / Published online: 10 May 2008  
© The Author(s) 2008

**Abstract** Patterns of geographical diversity, and the relationship between agro-morphological traits and fatty acid composition were assessed for 193 safflower (*Carthamus tinctorius*) accessions representing forty countries. Accessions were assigned to eight groups based on geographical proximity. Cluster and Principal Component analyses were performed to assess patterns of diversity among the accessions and to select the most distant accessions from each of eight groups for analysis of randomly amplified polymorphic DNA (RAPD) markers. There was a large amount of diversity for agro-morphological traits, fatty acid composition, and RAPD markers. Most correlations among different traits were rather low. Plant height showed a positive correlation with days to flowering ( $r = 0.63^{**}$ ). Palmitic acid was positively correlated with stearic acid and oleic acid values, and negatively correlated with linoleic acid ( $P < 0.01$ ). Oleic acid and

linoleic acid showed a strong negative correlation ( $r = -0.89^{**}$ ). The first three principal components together explained 59% of the variation, however, neither principal component analysis (PCA) nor marker analysis revealed a clear relationship between diversity pattern and geographical origin. Accessions from some geographical regions tended to group together, such as accessions from South Western Asia, Central Western Europe, and the Mediterranean region. The correlation between the morphological matrix and the genetic matrix based on RAPD markers was not significant ( $r = 0.027$ ). Wide diversity in safflower germplasm indicates a considerable potential for improving this crop for both agronomic and quality traits.

**Keywords** Accessions · *Carthamus tinctorius* · Fatty acids · Genetic diversity · Geographical groups · Multivariate analysis · Polymorphic bands · Safflower

M. A. Khan · S. von Witzke-Ehbrecht · H. C. Becker  
Department of Crops Sciences: Plant Breeding,  
Georg-August-Universität Göttingen, von Sieboldstr. 8,  
37075 Göttingen, Germany

M. A. Khan (✉)  
Plant Pathology Group, Institute of Integrative Biology  
(IBZ), ETH Zürich, Zürich 8092, Switzerland  
e-mail: awais.khan@agrl.ethz.ch

B. L. Maass  
Department of Crop Sciences: Agronomy in the Tropics,  
Georg-August-Universität Göttingen, Grisebachstr. 6,  
37077 Göttingen, Germany

## Introduction

Despite the large volume of globally traded vegetable oil, 75% of the production is from only four crops: soybean, oil palm, rapeseed, and sunflower. Because of the domination of these four (Murphy 1999), many other oil crops are now underutilized or neglected, though these species provide opportunities through

their great genetic diversity and diverse agro-ecological adaptation (Hammer and Heller 1998; Padulosi et al. 1999; Thies 2000). This monopoly of few crops has not only resulted in increased vulnerability of agricultural production to sudden disease outbreaks but also impoverished the human diet. Further domestication and development of underutilized crops is a possible solution for the growing and diversified nutritional needs of humankind.

Safflower (*Carthamus tinctorius* L.), a member of the family *Asteraceae*, is one of the important underutilized oil seed crops. It is a highly branched, herbaceous, thistle-like annual or winter annual, usually with many long sharp spines on the leaves. Plants are 30–200 cm tall and have strong taproots, which enable them to thrive in dry climates. It is a multi-purpose crop species grown for oil, medicinal and industrial uses, from the Mediterranean region to the Pacific Ocean, at latitudes between 45° N and 45° S. Traditionally, the crop was grown for its yellow-orange flowers, used for coloring and flavoring foods, and making dyes (Ashri 1975; Weiss 1983; Knowles 1989). Safflower oil is thought to be one of the highest quality vegetable oils, containing oleic acid and linoleic acid. There are two types of safflower; high linoleic types that have an approximate range of linoleic acid from 3.1% to 88.8%, and high oleic types with an approximate range of oleic acid from 3.9% to 90.6% (Fernández-Martínez et al. 1993).

For the effective use of underutilized crops, it is critical to understand the extent and distribution of genetic diversity within species (Padulosi et al. 1999). In the past, genetic diversity studies in crop plants mostly relied on the evaluation of morphological and agronomic traits (Upadhyaya et al. 2002; Atta-Krah et al. 2004). Currently, there are many molecular marker systems routinely used to evaluate genetic diversity in plants. These include randomly amplified polymorphic DNA (RAPDs), amplified fragment length polymorphisms (AFLPs) and ISSRs (intersimple sequence repeats) (Patzak 2001). RAPD markers are frequently used in genetic divergence studies of crops because they do not require prior genomic information, and are simpler, less costly, and less labor-intensive than other DNA marker techniques (Williams et al. 1990). RAPDs have been used successfully with many other underutilized crop species, for example, lentil (Ferguson et al. 1998), sesame (Bhat et al. 1999), and flax (Fu 2005).

Characterization of safflower germplasm so far included agro-morphological studies (Khan et al. 2004; Dwivedi et al. 2005; Jaradat and Shahid 2006), biochemical analyses (Fernández-Martínez et al. 1993; Pascual-Villalobos and Albuquerque 1996; Zhang 2001) and, recently, molecular markers have been applied (Sehgal and Raina 2005; Johnson et al. 2007; Yang et al. 2007). Nevertheless, the RAPD study by Sehgal and Raina (2005) included 14 Indian cultivars only. Also, genetic diversity is best estimated if agro-morphological, biochemical and marker studies are used together (Vollmann et al. 2005) along with appropriate multivariate analysis (Mohammadi and Prasanna 2003). Despite the limited information available that integrates agro-morphological, biochemical and genetic methods in diversity studies of safflower (e.g., Johnson et al. 2007), core collections have been proposed by Johnson et al. (1999) and Dwivedi et al. (2005).

The objectives of this research were to evaluate the diversity in a comprehensive safflower germplasm collection at agro-morphological, biochemical and molecular levels, to determine relationships among these factors and, consequently, identify patterns of geographical diversity.

## Materials and methods

Seed material of 747 accessions from all over the world was obtained from several genebanks and sown in April 2002 in the experimental field at Reinshof of Georg-August-University Göttingen (Latitude, 51°32' North and Longitude, 9°57' East). The experiment was laid out in a simple lattice design with two replications of each accession. In every plot, there were three rows and in each row 25 seeds were sown. Each row was 120 cm long, and the path between rows was 80 cm. All accessions were provided with the same agronomical practices and weeds were manually controlled. The design was equal to similar field trials carried out at the German Hohenheim University within the frame of a larger project "Screening, selection and suitability for cultivation of safflower (*Carthamus tinctorius*) and false flax (*Camelina sativa*) for the extraction of edible oil under conditions of ecological agriculture" (Reinbrecht et al. 2003) financed by StollVITA-Stiftung, Waldshut, Germany.

Data previously collected by Khan et al. (2004) for the traits flower color, spines, branching pattern, head size, disease, days to flowering, plant height, palmitic acid, stearic acid, oleic acid, and linoleic acid were analyzed. Methodology adopted for the collection of data was as follows. *Flower color*: Color of fresh flowers on the plants was recorded as white = 1, pale yellow = 2, yellow = 3, yellow orange = 4, orange = 5, orange red = 6, red = 7. *Spines*: Presence and absence of spines was observed as spineless = 1, spiny = 2. *Branching pattern*: Three levels of branching pattern were scored, basal = 1, medium = 2, upper = 3. *Head size*: Scored as small = 1, intermediate = 2, and large = 3. *Disease*: Infection of any kind of disease was noted as tolerant = 1, intermediate = 2, infected = 3. *Days to flowering*: Number of days from the date of sowing to the date when at least three plants showed open flowers. *Plant height*: Mean height (cm) at maturity was measured when plants were in full flower from ground level to the tip of the main stem. *Fatty acid analysis*: Gas Chromatography (GC) was used for fatty acid analysis after Velasco et al. (1998).

Seed was harvested manually at maturity, about 6 months after sowing. Because of disease incidence and unfavorable climate, only 193 out of 747 accessions produced enough seeds, to be used for further analysis. These 193 accessions were divided into geographical groups according to the information received from the genebanks, similar to Johnson et al. (1999), and one group with accessions from unknown origin (Table 1).

**Table 1** Safflower accessions divided into geographical groups based on origin after geographical information provided by genebanks

Geographical group	Abbreviation	No. of accessions
East Asia	Ea	13
South Western Asia	Swa	34
Eastern Europe	Eeu	14
Central Western Europe	Cwe	42
Southern Europe	Seu	15
Mediterranean	Med	22
Africa	Af	10
North America	Na	24
Unknown	Un	19
Total		193

## Analysis of agro-morphological and biochemical data

Pearson correlation coefficient was calculated for all agro-morphological and biochemical variables. The means of each geographical group for all agro-morphological traits and fatty acid composition were then used for cluster analysis to study the relationship between geographical groups. Euclidean distance was used as the resemblance coefficient for cluster analysis with the Unweighted Pair Group Arithmetic Means method (UPGMA) using NTSYS-pc software version 2.1 (Rohlf 2001).

## Selection of accessions for marker study

Agro-morphological and biochemical data was standardized before using in multivariate analysis by applying the YBAR and STD options in NTSYS-pc software version 2.1 (Rohlf 2001). Principal component analysis (PCA) was then performed on all these variables. In order to cover broad agro-morphological and biochemical variation, accessions were chosen for further analysis on the basis of the first two PCA scores. Four accessions, two with minimum values and two with maximum values, were selected from each of the nine geographical groups for randomly amplified polymorphic DNA (RAPD) analysis. Single plants of selected accessions were raised from seed produced in the field in a greenhouse of University of Göttingen; however, one accession from South Western Asia did not germinate.

## RAPD marker analysis

Genetic diversity was estimated in safflower accessions using RAPD marker analysis. Young, emerging leaves from a single plant randomly chosen per accession were harvested from plants three weeks after emergence. Genomic DNA was extracted from 90 to 120 mg of young leaf tissue of 35 accessions by using the Nucleon PhytoPure DNA Extraction Kit (Amersham Pharmacia) protocol. DNA concentration was quantified by using a Fluorometer (Versa Fluor-TM, System 170-2402; Bio-Rad Lab. Inc.). For RAPD analysis of genomic DNA, 10-decamer primers (Operon DNA Technologies, Alameda, California, USA) were chosen randomly and others recommended by RC Johnson (personal communication). The

polymerase chain reaction (PCR) was performed in a volume of 25  $\mu$ l containing 50 ng of genomic DNA, 10 $\times$  polymerase buffer, 0.2 mM of each dNTP, 3.0 mM MgCl<sub>2</sub>, 0.4  $\mu$ M primer, and 1 unit of *Taq* polymerase. PTC-100 thermocycler (MJ Research, Watertown, Massachusetts, USA) was used for PCR amplification reactions with the PCR conditions: 94°C denaturation step (30 s) followed by 45 cycles of 1 min at 92°C, 1 min at 35°C, and 2 min at 72°C, and a final extension step of 5 min at 72°C. After amplification, 5  $\mu$ l of loading buffer was added in each reaction tube and the amplification products were separated on 1.5% agarose gels in TBE buffer. The gels were stained with ethidium bromide (1 mg/l H<sub>2</sub>O) for 10–15 min and washed in distilled water for 25–30 min. The RAPD bands were then visualized through UV-Transilluminator (UV-Rays at  $\lambda = 254$  nm) and photographed.

#### Analysis of RAPD marker data

Sharp and highly reproducible polymorphic RAPD bands were visually scored to develop a 1, 0 matrix. Jacquard's similarity coefficient was calculated and used to produce an UPGMA-based dendrogram. For the similarity analysis, NTSYS-pc software version 2.1 (Rohlf 2001) was applied.

#### Comparison of morphological and marker diversity

Finally, for 35 accessions an Euclidean distance matrix for agro-morphological and biochemical data and a Jaccard's similarity matrix for marker data were compared. Correlations between the distance and similarity matrices were performed by using NTSYS-pc software version 2.1 (Rolf 2001).

## Results

#### Agro-morphological and biochemical diversity

The safflower germplasm evaluated exhibited large variation in all characters studied (Table 2). Fresh flower color varied from white to orange red, and both spiny and spineless accessions were present in all geographical groups. Under the field conditions tested, South Western Asian (mean 1.5 SD  $\pm$  0.4) and East Asian (1.6 SD  $\pm$  0.4) accessions were less

susceptible to diseases compared to the remainder of the groups. The range of days to flowering in material from Africa, East Asia and South Western Asia was slightly larger than those of the other groups. The range of plant height in East Asian accessions was lower than in the other groups, while it was higher in accessions from Eastern Europe. Mediterranean and African material had a relatively narrow range of plant height compared to other groups.

Material from North America exhibited a larger range of palmitic acid as compared to material from the other groups, while for stearic acid, accessions from Central Western Europe had the widest range. For oleic acid, germplasm from North America, East Asia, and Africa displayed high maximum values. Linoleic acid ranged from 63% to 83%, with Eastern European accessions showing the greatest range (68.3–83.3%).

There was a significant ( $P > 0.01$ ) correlation between days to flowering and plant height as well as plant height and disease susceptibility (Table 3). Palmitic acid was positively correlated with both stearic acid and oleic acid, and negatively with linoleic acid ( $P < 0.01$ ). A highly significant negative correlation existed between oleic acid and linoleic acid.

#### Patterns of agro-morphological and biochemical diversity

The cluster analysis based on means of all agro-morphological and biochemical variables from each geographical group showed that the accessions from Eastern and Southern Europe clustered together with the Mediterranean group, and African accessions were also close to this cluster. The East Asian group was the most distant (Fig. 1).

When applying PCA on all agro-morphological and biochemical variables together, the first four principal components accounted for 68.4% of the variation (Table 4). While the first principal component (PC1) contributed about 25% of the variation, showing highest contributions from the proportions of linoleic and oleic acid, PC2 received higher contributions from agro-morphological traits. However, palmitic acid and disease susceptibility were also important for PC1. PC2 contributed 23% to the total variation, with days to flowering and plant height contributing the most.

**Table 2** Agro-morphological and biochemical variation in 193 safflower germplasm accessions evaluated at Reinshof, Germany, according to different geographical regions of origin

Variable	Geographical group							
	East Asia	South Western Asia	Eastern Europe	Central Western Europe	Southern Europe	Mediterranean	Africa	North America
Flower color								
<i>n</i>	11	33	14	40	15	22	10	22
Mean	3.9	4.0	4.0	3.7	4.1	4.1	4.0	3.2
SD	0.5	0.7	0.6	0.6	0.6	0.6	0.6	0.7
Min.	3.0	2.5	3.5	3.0	3.0	3.0	3.0	1.0
Max.	4.5	6.0	5.0	6.0	5.0	5.0	5.0	4.0
Spines								
<i>n</i>	13	32	14	41	15	20	9	24
Mean	1.9	1.8	1.4	1.6	1.6	1.5	1.7	1.7
SD	0.3	0.3	0.4	0.5	0.4	0.5	0.4	0.5
Min.	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Max.	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Branching pattern								
<i>n</i>	13	33	14	42	15	21	10	24
Mean	1.8	1.7	2.1	2.0	2.2	2.1	2.0	1.9
SD	0.4	0.5	0.4	0.3	0.4	0.3	0.2	0.3
Min.	1.0	1.0	1.0	1.0	1.5	1.5	1.5	1.5
Max.	2.0	2.5	2.5	2.5	3.0	3.0	2.0	2.5
Head size								
<i>n</i>	13	34	14	42	15	21	10	24
Mean	2.0	1.9	2.1	2.0	2.2	2.4	2.1	2.0
SD	0.4	0.7	0.5	0.4	0.6	0.5	0.6	0.4
Min.	1.0	1.0	1.0	1.0	1.0	1.5	1.5	1.0
Max.	2.5	3.0	3.0	3.0	3.0	3.0	3.0	3.0
Disease								
<i>n</i>	13	34	14	42	15	22	10	24
Mean	1.6	1.5	2.3	2.7	2.3	1.8	2.0	1.9
SD	0.4	0.4	0.5	0.4	0.5	0.5	0.6	0.6
Min.	1.0	1.0	1.5	1.5	1.5	1.0	1.0	1.0
Max.	2.0	2.5	3.0	3.0	3.0	3.0	3.0	3.0
Days to flowering								
<i>n</i>	13	34	14	42	15	22	10	24
Mean	205.3	209.9	212.5	209.7	212.0	213.3	213.5	212.0
SD	5.8	5.0	5.8	2.8	3.6	3.2	5.7	3.5
Min.	197.5	203.0	207.0	204.0	206.5	206.0	208.0	206.5
Max.	218.0	221.0	228.5	217.5	220.0	218.0	224.0	220.0
Plant height (cm)								
<i>n</i>	13	34	14	42	15	22	10	24
Mean	77.7	85.9	97.9	88.9	94.3	94.3	94.0	90.4
SD	12.9	12.3	16.4	12.0	11.9	9.3	16.0	9.3
Min.	55.0	65.0	80.0	60.0	80.0	80.0	75.0	70.0
Max.	100.0	115.0	130.0	115.0	115.0	115.0	120.0	115.0

**Table 2** continued

Variable	Geographical group							
	East Asia	South Western Asia	Eastern Europe	Central Western Europe	Southern Europe	Mediterranean	Africa	North America
Palmitic acid (%)								
<i>n</i>	13	34	14	42	15	22	10	24
Mean	5.8	6.8	6.3	6.1	6.1	6.4	6.3	6.5
SD	0.8	0.9	1.1	0.7	0.8	0.6	0.8	1.1
Min.	4.6	5.3	3.8	4.1	4.9	5.6	5.0	4.8
Max.	7.1	9.3	7.8	7.9	7.4	8.5	7.4	10.4
Stearic acid (%)								
<i>n</i>	13	34	14	42	15	22	10	24
Mean	1.7	2.1	1.9	2.0	1.7	2.0	2.0	1.7
SD	0.5	0.5	0.6	0.6	0.3	0.5	0.6	0.7
Min.	0.8	1.0	0.9	0.6	1.2	1.4	0.4	0.0
Max.	2.4	3.1	2.7	3.7	2.4	3.0	2.5	2.9
Oleic acid (%)								
<i>n</i>	13	34	14	42	15	22	10	24
Mean	14.8	14.0	12.6	11.9	13.1	13.3	14.7	14.4
SD	4.2	3.0	2.7	1.3	2.5	2.3	3.7	4.1
Min.	11.3	9.5	9.8	9.5	10.1	10.4	10.7	11.0
Max.	25.5	21.9	20.1	14.4	18.2	18.7	23.7	27.6
Linoleic acid (%)								
<i>n</i>	13	34	14	42	15	22	10	24
Mean	75.5	74.9	77.2	78.4	76.5	76.0	74.2	75.1
SD	4.9	4.3	4.0	2.2	3.9	3.2	4.6	4.2
Min.	64.0	65.7	68.3	71.4	67.7	67.4	63.3	64.5
Max.	80.5	81.9	83.3	81.3	81.2	80.4	80.1	80.4

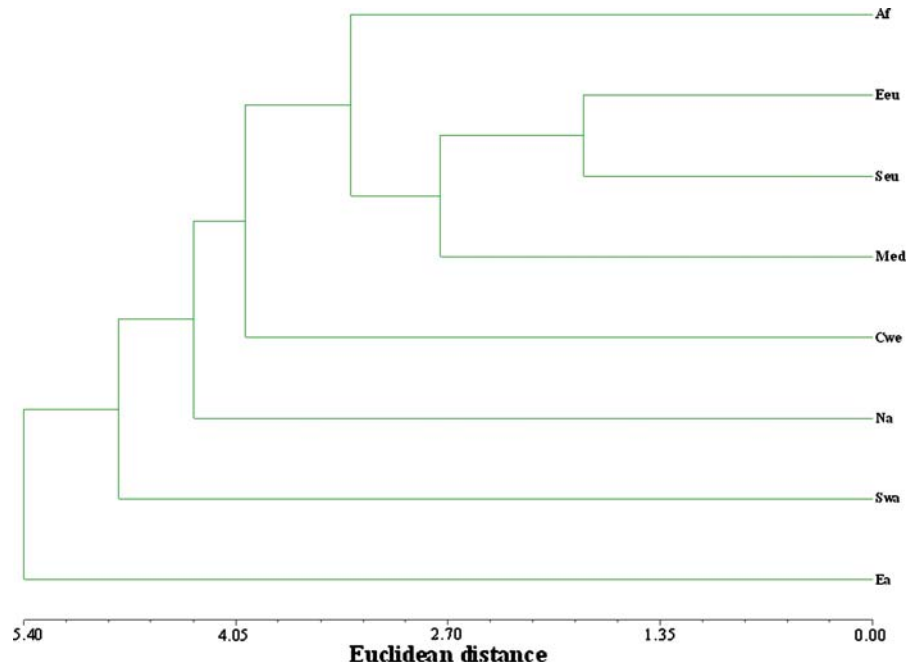
Number of observations, *n*; minimum, Min.; maximum, Max.; Mean, and standard deviation, SD

**Table 3** Pearson correlation coefficients for agro-morphological traits and fatty acid composition of 193 safflower germplasm accessions evaluated at Reinshof, Germany

Variable	Flower color	Branching pattern	Head size	Disease	Days to flowering	Plant height	Palmitic acid	Stearic acid	Oleic acid
Branching pattern	-0.02								
Head size	-0.02	-0.01							
Disease	-0.09	-0.01	-0.12						
Days to flowering	-0.16*	0.05	-0.01	0.14*					
Plant height	-0.15*	0.03	0.05	0.19**	0.63**				
Palmitic acid	0.06	0.05	0.06	-0.26**	0.00	-0.07			
Stearic acid	0.11	0.02	0.14	-0.05	-0.13	-0.17*	0.25**		
Oleic acid	0.02	0.03	-0.01	-0.34**	0.08	-0.003	0.27**	-0.10	
Linoleic acid	-0.06	-0.03	-0.05	0.39**	-0.08	-0.03	-0.50**	-0.14	-0.89**

\*, \*\* Means significant at  $P > 0.05$  and  $0.01$  respectively

**Fig. 1** Dendrogram produced by UPGMA clustering based on means of 7 agro-morphological traits and fatty acid composition (4 acids) of 193 safflower accessions from eight geographical groups. (East Asia, Ea; South Western Asia, Swa; Eastern Europe, Eeu; Central Western Europe, Cwe; Southern Europe, Seu; Mediterranean, Med; Africa, Af; North America, Na; Unknown, Un.)



**Table 4** Eigenvectors and eigenvalues generated by Principal Component Analysis applied on agro-morphological traits and fatty acid composition of 193 safflower accessions

Variable	Eigenvectors			
	PC1	PC2	PC3	PC4
Flower color	0.14	0.48	0.19	0.68
Branching pattern	0.45	0.42	-0.14	-0.25
Head size	0.23	0.47	-0.06	0.43
Disease	0.63	-0.05	-0.25	-0.22
Days to flowering	0.28	0.74	-0.20	-0.13
Plant height	0.41	0.70	-0.17	-0.17
Palmitic acid	-0.61	0.16	-0.50	-0.07
Stearic acid	-0.27	-0.23	-0.78	0.30
Oleic acid	-0.71	0.52	0.27	-0.16
Linoleic acid	0.80	-0.52	0.01	0.10
	Eigenvalues			
Eigenvalue	2.51	2.29	1.11	0.93
Percent of total variance explained	25.11	22.87	11.11	9.35
Cumulative percent of total variance explained	25.11	47.98	59.08	68.43



**Table 5** Operon primers with number of scorable polymorphic bands, with 5′–3′ sequence used to assess genetic diversity of safflower

Operon code	Sequence 5′–3′	Polymorphic bands produced (no.)
OP A-1	CAGGCCCTTC	4
OP A-9	GGGTAACGCC	2
OP AA-04 <sup>a</sup>	AGGACTGCTC	7
OP AA-10 <sup>a</sup>	TGGTCGGGTG	6
OP AA-11 <sup>a</sup>	ACCCGACCTG	4
OP AA-12 <sup>a</sup>	GGACCTCTTG	8
OP AA-16 <sup>a</sup>	GGAACCCACA	6
OP Al-1	TGTGACGAGG	1
OP B-03 <sup>a</sup>	CATCCCCCTG	5
OP B-05	TGCGCCCTTC	3
OP B-06	TGCTCTGCC	4
OP B-13	TTCCCCGCT	10
OP B-19 <sup>a</sup>	ACCCCGAAG	8
OP F-02 <sup>a</sup>	GAGGATCCCT	5
OP F-05 <sup>a</sup>	CCGAATTCCC	5
Total polymorphic bands		78

<sup>a</sup> Recommended by RC Johnson (personal communication)

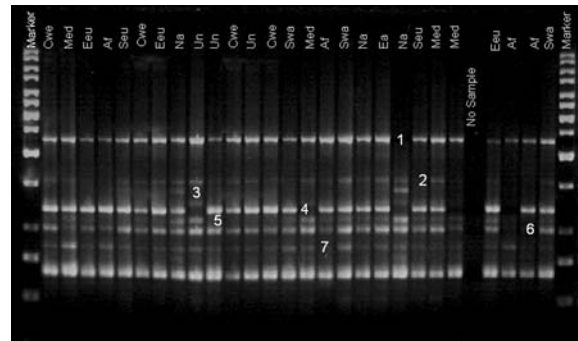
Despite the large diversity in germplasm for fatty acid composition, days to flowering and plant height, no clear separation between accessions from different geographical groups was achieved by PCA. However, a general tendency was observed that South Western Asian accessions and accessions from Central Western Europe lay close to each other (data not shown).

#### Genetic diversity

A total of 78 polymorphic bands were produced by 15 primers (Table 5). The primers produced a maximum of 10 (OP A-9, OP B-13) and a minimum of 1 (OP Al-1) polymorphic band. Operon primer AA-04 produced seven polymorphic bands (Fig. 2).

The cluster analysis based on marker data showed no clear grouping of accessions with regard to geographical regions of origin reported (Fig. 3). Two of the African and one of the Mediterranean accessions were most distant in the cluster.

The correlation between the Euclidean distance matrix for agro-morphological traits and fatty acid composition, and Jacquard's similarity matrix for



**Fig. 2** RAPD analysis of 27 safflower accessions from eight different geographical regions with Operon primer AA-04 (see Table 5). (East Asia, Ea; South Western Asia, Swa; Eastern Europe, Eeu; Central Western Europe, Cwe; Southern Europe, Seu; Mediterranean, Med; Africa, Af; North America, Na; Unknown, Un.)

marker data was low and not significant ( $r = 0.027$ ). This shows that agro-morphological and biochemical variation was generally independent of RAPD variation.

#### Discussion

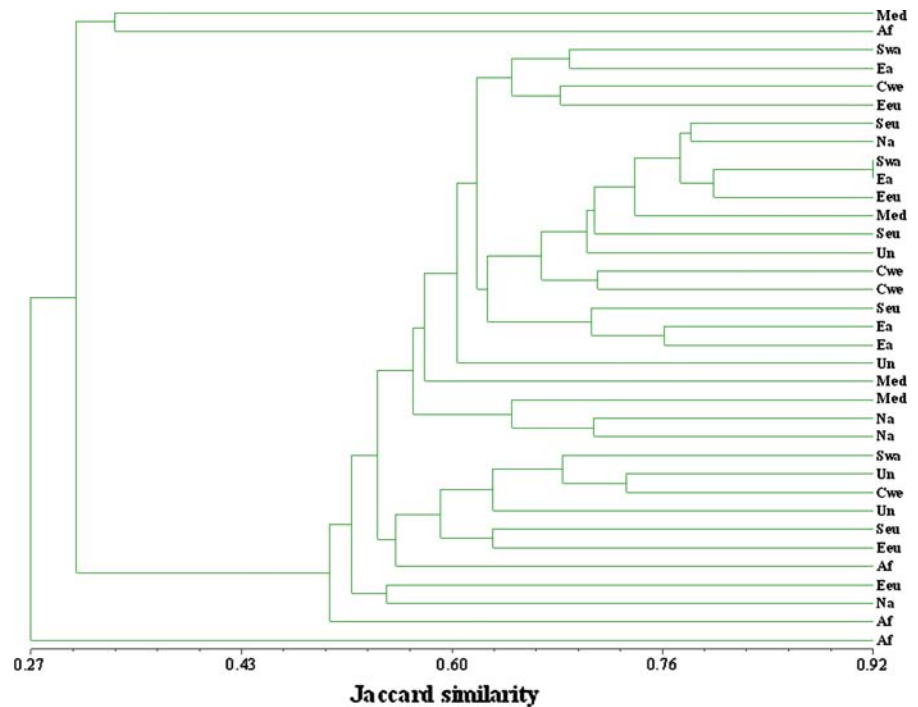
##### Agro-morphological, biochemical and genetic diversity

A large amount of diversity was found in safflower accessions at agro-morphological, biochemical, and genetic levels. Among the agro-morphological and biochemical traits studied were oil content (oleic and linoleic acid), plant height and days to flowering. However, the present germplasm did not contain high oleic acid types (Table 2) and, therefore, it is not directly comparable to the core and non-core accessions described by Johnson et al. (1999). Other studies (Ashri 1975; Jaradat and Shahid 2006) reported comparable results, showing a wide range of variation in safflower collections for agro-morphological traits as a result of human and natural selection. This suggests good potential for selection in a breeding program (Pascual-Villalobos and Albuquerque 1996).

PCA identified linoleic and oleic acid as the most important traits responsible for variation in our material. Variation in linoleic and oleic acid content of safflower oil indicates the possibility of improving oil quality through breeding (Fernández-Martínez



**Fig. 3** Dendrogram for 35 safflower accessions generated by UPGMA clustering using Jaccard's coefficient of similarity on RAPD data. (East Asia, Ea; South Western Asia, Swa; Eastern Europe, Eeu; Central Western Europe, Cwe; Southern Europe, Seu; Mediterranean, Med; Africa, Af; North America, Na; Unknown, Un.)



et al. 1993). Consequently, the most diverse accessions could be selected on the basis of these first two principal components. Even though a small number of RAPD primers was used in this study, these still produced a reasonable number of polymorphic bands, indicating large genetic diversity contained in the materials studied, especially compared to the few cultivars researched by Sehgal and Raina (2005).

#### Relationship between diversity patterns and geographical origin

Neither cluster analysis nor PCA revealed a clear relationship between diversity pattern and geographical origin. For most of the traits studied, similarity among accessions was independent of their origin, and there was no trait in the present study, which could clearly separate the accessions on the basis of geographical origin.

Ashri et al. (1975) stated that the variation for days to flowering within a region was so great that the difference between regions was not apparent. Most regional groups in the comprehensive study by Johnson et al. (1999) strongly overlapped. Our results (Fig. 3) also showed large variation within regions and, most likely, the lack of clear clustering of the

accessions according to geographical regions of origin could partly have been for this reason. On the other hand, the fact that only one single plant per accession was studied by RAPDs, disregarded any possible variability that might be contained in the accessions. Johnson et al. (2007) stated that the high AFLP marker uniformity within safflower accessions was likely associated with predominant self-pollinating, although the species also has potential for substantial outcrossing. They could even illustrate that eight bulk samples studied by AFLPs were representative for most individuals from these populations.

Results presented here may also point to the practice of genebanks across the world, which, in their effort to assemble diverse samples, may have acquired various duplicates of the accessions from other institutions. For example, Li and Mündel (1996) and Dwivedi et al. (2005) reported elimination of duplicate entries from safflower genebank material. This conclusion is also supported by the fact that there are a few closely related accessions from otherwise diverse regions in the cluster analysis (Fig. 3). Ashri (1975) also reported that divergence was not always obvious. He reasoned that this could be due to unspecificities of variety origins, either due

to inclusion of recently naturalized varieties, the discrepancy between phytogeographical regions and political boundaries, or unreported exchange of seeds between regions by nomads, armies, or merchants. Broad diversity of safflower accessions within large regions is reasonable given the variability of climatic and agricultural conditions, as well as safflower's traditional mode of cultivation, in small plots, which could have fragmented the gene pool into small populations, in which natural and human selection, and genetic drift could have operated (Ashri 1975; Ashri et al. 1975; Jaradat and Shahid 2006).

Cluster analysis based on geographical groups showed that Eastern European and Southern European accessions were fairly similar and clustered together with the Mediterranean group. African accessions also lay in the same main cluster, which supports the studies by Johnson et al. (1999) and Jaradat and Shahid (2006) that reported the proximity of African accessions to the Mediterranean group. The East Asian group stood out as the most distant. Isozyme study of 89 safflower accessions has shown that East Asian accessions had the highest mean gene diversity (Zhang 2001). Whereas, Johnson et al. (2007) attributed higher gene diversity to the American group due to past efforts of crop improvement by introducing exotic materials, especially from China. Genetic diversity showed no relationship to geographical origin (Fig. 3), as in the case of morphological diversity; accessions from unknown origin were also randomly distributed over the whole dendrogram.

#### Relationship between agro-morphological and biochemical, and genetic diversity

There was no association between agro-morphological and biochemical diversity and molecular diversity. Similar disparity between morphological traits and RAPDs was reported in different studies (Vollmann et al. 2005; Johnson et al. 2007). There could be many reasons for the lack of correlation between RAPDs and morphological distances. One reason could be that RAPDs detect polymorphisms in coding as well as non-coding regions of the genome, of which only a small portion is coding, therefore, it is very likely that the polymorphism found is in a non-coding region. The relationship between molecular markers and phenotypic traits could be significant if the markers were linked to selected loci (Persson and Gustavsson 2001).

Also, plants that are morphologically similar are not necessarily genetically so, as different gene pools could result in similar phenotypes. Comparisons may correctly display similarities in neither, both or either morphological or molecular phenotypes, making the relationship between groups complex to interpret. Therefore, consistency should only be expected if plants had shared genetic resources and parallel breeding objectives or, conversely, were very different in both gene-pool source and selection targets (Roldán-Ruiz et al. 2001; Archak et al. 2003).

#### Conclusions

The diversity encountered in safflower germplasm points out that there is a large potential for the improvement of safflower for both agronomic and quality traits. This wide variation, however, is probably only a proportion of the variation that exists in safflower worldwide, and further evaluation appears warranted to use and conserve this potentially valuable genetic resource. Breeding strategies need to exploit existing variation within the safflower germplasm to broaden the genetic base of currently used cultivars.

**Acknowledgements** We would like to thank Dr. Carsten Reinbrecht from the Institute of Plant Production and Grassland Research, University of Hohenheim for providing seeds of the safflower germplasm, his informative discussions and valuable suggestions during field studies. Further, we are grateful to Dr. Richard C. Johnson of USDA-ARS Plant Introduction Station Washington State University, for providing the primer sequences and useful comments on an earlier version of this manuscript. The valuable help of laboratory assistants Uwe Ammermann, Rosemarie Clemens, and Nicole Ritgen-Homayounfar during gas chromatography and molecular studies is also acknowledged.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

#### References

- Archak S, Gaikwad B, Gautam D, Rao E, Swamy K, Karihaloo J (2003) DNA fingerprinting of Indian cashew (*Anacardium occidentale* L.) varieties using RAPD and ISSR techniques. *Euphytica* 230:397–404

- Ashri A (1975) Evaluation of the germplasm collection of safflower, *Carthamus tinctorius* L. V. Distribution and regional divergence for morphological characters. *Euphytica* 24:651–659
- Ashri A, Zimmer DE, Urie AL, Knowles PF (1975) Evaluation of the germplasm collection of safflower *Carthamus tinctorius* L. VI. Length of planting to flowering period and plant height in Israel, Utah and Washington. *Theor Appl Genet* 46:359–364
- Atta-Krah K, Kindt R, Skilton JN, Amaral W (2004) Managing biological and genetic diversity in tropical agroforestry. *Agroforest Syst* 61:183–194
- Bhat KV, Babrekar PP, Lakhanpaul S (1999) Study of genetic diversity in Indian and exotic sesame (*Sesamum indicum* L.) germplasm using random amplified polymorphic DNA (RAPD) markers. *Euphytica* 110:21–34
- Dwivedi SL, Upadhyaya HD, Hegde DM (2005) Development of core collection using geographic information and morphological descriptors in safflower (*Carthamus tinctorius* L.) germplasm. *Genet Resour Crop Evol* 52:821–830
- Ferguson ME, Robertson LD, Ford-Lloyd BV, Newbury HJ, Maxted N (1998) Contrasting genetic variation amongst lentil landraces from different geographic origins. *Euphytica* 102:265–273
- Fernández-Martínez J, Rio MD, Haro AD (1993) Survey of safflower (*Carthamus tinctorius* L.) germplasm for variants in fatty acid composition and other seed characters. *Euphytica* 69:115–122
- Fu YB (2005) Geographic patterns of RAPD variation in cultivated flax. *Crop Sci* 45:1084–1091
- Hammer K, Heller J (1998) Promoting the conservation and use of underutilized and neglected crops. *Schriften Genet Resour* 8:223–227
- Jaradat AA, Shahid M (2006) Patterns of phenotypic variation in a germplasm collection of *Carthamus tinctorius* L. from the Middle East. *Genet Resour Crop Evol* 53:225–244
- Johnson RC, Bergman JW, Flynn CR (1999) Oil and meal characteristics of core and non-core safflower accessions from the USDA collection. *Genet Resour Crop Evol* 46:611–618
- Johnson RC, Kisha TJ, Evans MA (2007) Characterizing safflower germplasm with AFLP molecular markers. *Crop Sci* 47:1728–1736
- Khan MA, von Witzke-Ehbrecht S, Maass BL, Becker HC (2004) Evaluation of a worldwide collection of safflower for morphological diversity and fatty acid composition. *Vorträge für Pflanzenzüchtung* 62:104–106
- Knowles PF (1989) Safflower. In: Röbbelen G, Downey RK, Ashri A (eds) *Oil crops of the world, their breeding and utilization*. McGraw Hill, Inc., New York, pp 363–374
- Li D, Mündel HH (1996) Safflower. *Carthamus tinctorius* L. Promoting the conservation and use of underutilized and neglected crops. 7. Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben/International Plant Genetic Resources Institute (IPGRI), Rome, Italy, pp 8–66
- Mohammadi SA, Prasanna BM (2003) Analysis of genetic diversity in crop plants-salient statistical tools and considerations. *Crop Sci* 43:1235–1248
- Murphy DJ (1999) The future of new and genetically modified oil crops. In: Janick J (ed) *Perspectives on new crops and new uses*. ASHS Press, Alexandria, Virginia, pp 216–219
- Padulosi S, Eyzaguirre P, Hodgkin T (1999) Challenges and strategies in promoting conservation and use of neglected and underutilized crop species. In: Janick J (ed) *Perspectives on new crops and new uses*. ASHS Press, Alexandria, Virginia, pp 140–145
- Pascual-Villalobos MJ, Albuquerque N (1996) Genetic variation of a safflower germplasm collection grown as a winter crop in southern Spain. *Euphytica* 92:327–332
- Patzak J (2001) Comparison of RAPD, STS, ISSR and AFLP molecular methods used for assessment of genetic diversity in hop (*Humulus lupulus* L.). *Euphytica* 121:9–18
- Persson HA, Gustavsson BA (2001) The extent of clonality and genetic diversity in lingonberry (*Vaccinium vitis-idaea* L.) revealed by RAPDs and leaf-shape analysis. *Mol Ecol* 10:1385–1397
- Reinbrecht C, Barth S, von Witzke-Ehbrecht S, Khan MA, Becker HC, Kahnt G, Claupein W (2003) Selektion anbauwürdiger Saflorformen für den ökologischen Anbau aus einem zweijährigen Screening-Experiment. *Mitteilungen der Gesellschaft für Pflanzenwissenschaften* 15:23–26
- Rohlf FJ (2001) Comparative methods for the analysis of continuous variables: geometric interpretations. *Evolution* 55:2143–2160
- Roldán-Ruiz I, Van Eeuwijk FA, Gilliland TJ, Dubreuil P, Dillmann C, Lallemand J, De Loose M, Baril CP (2001) A comparative study of molecular and morphological methods of describing relationships between perennial ryegrass (*Lolium perenne* L.) varieties. *Theor Appl Genet* 103:1138–1150
- Sehgal D, Raina SN (2005) Genotyping safflower (*Carthamus tinctorius*) cultivars by DNA fingerprints. *Euphytica* 146:67–76
- Thies E (2000) Promising and underutilized species: crops and breeds. *Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH. Managing agrobiodiversity in rural areas*. Eschborn, Germany
- Upadhyaya HD, Bramel PJ, Ortiz R, Singh S (2002) Geographical pattern of diversity for morphological and agronomic traits in the groundnut germplasm collection. *Euphytica* 128:191–204
- Velasco L, Schierholt A, Becker HC (1998) Performance of near-infrared reflectance spectroscopy (NIRS) in routine analysis of C18 unsaturated fatty acids in intact rapeseed. *Fett/Lipid* 100:44–48
- Vollmann J, Grausgruber H, Stift G, Dryzhyruk V, Lelley T (2005) Genetic diversity in *Camelina* germplasm as revealed by seed quality characteristics and RAPD polymorphism. *Plant Breed* 124:446–453
- Weiss EA (1983) Safflower, Chapter 6. In: *Oilseed crops. Tropical agriculture series*. Longman, London, pp 216–281
- Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* 18:6531–6535

- Yang YX, Wu W, Zheng YL, Chen L, Liu RJ, Huang CY (2007) Genetic diversity and relationships among safflower (*Carthamus tinctorius* L.) analyzed by inter-simple sequence repeats (ISSRs). *Genet Resour Crop Evol* 54:1043–1051
- Zhang ZW (2001) Studies on genetic diversity and classification of safflower (*Carthamus tinctorius* L.) germplasm by isozyme techniques. In: Bergman J, Mündel HH (eds) *Proceedings of the 5th International Safflower Conference*, Williston, ND and Sidney, MT, 23–27 July 2001, pp 157–162