

Research Paper

Selection tools for oil content and fatty acid composition in safflower (*Carthamus tinctorius* L.)

Alfonso Cerrotta*¹⁾, Lilia Ivone Lindström²⁾ and Viviana Echenique^{1,2)}

¹⁾ Departamento de Agronomía, Centro de Recursos Naturales Renovables de la Zona Semiárida (CERZOS-CONICET, CCT, Bahía Blanca), Universidad Nacional del Sur, San Andrés 800, Bahía Blanca, 8000, Argentina

²⁾ Departamento de Agronomía, Universidad Nacional del Sur, San Andrés 800, Bahía Blanca, 8000, Argentina

Agricultural expansion requires the deployment of stress-tolerant crops like safflower (*Carthamus tinctorius* L.). In safflower breeding, oil improvement in early generations requires indirect selection through simply inherited traits. The oil quality is mostly related to the fatty acid profile, which is determined by the OL locus. The aim of this research was to identify simple easy-to-measure traits that indirectly explain oil content variation and its interaction with yield components, and also to generate an effective tool for genotyping the OL locus. A field experiment with F₅ and pure lines was carried out to correlate the oil content with 18 traits including yield components, and phenological and morphological characteristics. KASP technology using primers designed according to the ctFAD2-1 gene sequence was applied for OL locus genotyping and validated through fatty acids phenotyping. Hull content, the length:width ratio of the grain, and plant height were identified as the most promising selection tools for increasing oil content, and grains per capitulum was the best yield component for increasing yield without decreasing the oil content. KASP genotyping successfully worked as a MAS tool, identifying oleic and linoleic genotypes. These tools enhance options for improving oil content and quality for safflower breeding.

Key Words: safflower, oil content, fatty acids, yield components, indirect selection tools.

Introduction

Safflower (*Carthamus tinctorius* L.) is an underutilized and neglected annual oilseed crop belonging to the Asteraceae family. It has been cultivated since ancient times in Egypt and the Near East, where the center of origin is most probably located. It is a diploid species (Patel and Narayana 1935) mainly self-pollinated, with a variable rate of outcrossing (5–40%). The deep taproot and the xerophytic attributes of the plant contribute to heat and drought-stress tolerance. Originally its colorful florets were used in the dye industry, as a food coloring and for medicinal purposes (Li and Mündel 1996). Nowadays it is produced mostly for its high-quality edible oil (Knowles 1989), but it is also used as birdseed, industrial oil, and more recently in molecular farming (Markley *et al.* 2006). Although genetic and structural male sterility have been discovered in safflower, commercial hybrid seed production is still rarely used due to the cost of hybrid breeding (Baydar *et al.* 2003).

In traditional safflower breeding, visual selection is the

most utilized tool for choosing the best genotypes (Udall 2003). Breeders gather the best gene combinations available, with the objective of improving one or more desired quantitative traits (yield, protein, oil content). Grain and oil yield are complex traits with low heritability (Parmeshwarappa *et al.* 1984), usually with a high genotype by environment interaction. In self-pollinated crops, early selection for yield (F₂–F₃) is highly inefficient (Knott 1972). In some cases, breeders do the selection for two or more highly heritable simple traits in order to identify superior genotypes. High correlation between a complex trait and an indirect selection tool enhance the expected genetic gain. Indirect selection can be applied in early generations when direct selection for complex traits in replicated trials is inefficient due to segregation and operational inconveniences. At early stages of a breeding program selection is done on individual plants level, so seed for conducting replicated yield trials is limited. Breeding by indirect selection has been utilized for improving grain yield, the grain oil content, and both combined (oil yield) (Golparvar 2011). Some simple closely related attributes have been recommended for indirect yield selection. For example, the number of capitula per plant, capitulum diameter and grains per capitulum (Arslan 2007). Capitula weight, capitula per plant, and hull (pericarp) content can be successfully used

Communicated by Ebrahiem M Babiker

Received May 11, 2020. Accepted July 30, 2020.

First Published Online in J-STAGE on October 28, 2020.

*Corresponding author (e-mail: acerrotta@cerzos-conicet.gov.ar)

for indirect selection for oil yield (Rao *et al.* 1977).

Grain yield and grain oil content in safflower also depend on crop phenology, which is related to the genotype and the environmental conditions. Depending on the type of cultivar (winter or spring) and the sowing date (autumn or winter), the grain and oil yield may differ widely due to the climatic conditions during the critical developmental stages. In 17 safflower genotypes evaluated in Iran, Vafaei *et al.* (2010) observed that when the flowering date was earlier the length of the reproductive period and grain yield both increased.

Highly heritable qualitative traits, with a single or a low number of genes involved, are fairly easy to manipulate and marker-assisted selection (MAS) is the most useful breeding approach for this task (Ribaut and Hoisington 1998). The fatty acid profile is the most important commercial oil quality trait in safflower. Linoleic types contain 70–75% linoleic acid whereas oleic types contain 75–80% oleic acid (Fernandez-Martinez *et al.* 1993). Both fatty acids reduce LDL cholesterol levels to help prevent cardiovascular disease (Lunn and Theobald 2006), but oleic acid has higher oxidative stability for frying at high temperature and for prolonged storage. This attribute makes it suitable for several chemical reactions, increasing its value in the oleochemical industry (Gunstone 2001). High oleic type is widely studied in safflower (Hamdan *et al.* 2012, Liu *et al.* 2013) and is caused by a mutation that consists of a cytosine deletion in the position 727 of the coding region of the FAD2-1 gene (HM165274.1) (Guan *et al.* 2012). It is expressed as a recessive monogenic (olol) trait with simple Mendelian inheritance, but also present are minor modifying QTLs (Hamdan *et al.* 2012). In any germplasm collection or crossing block, when both linoleic and oleic materials are present, a MAS approach using a KASP (Kompetitive Allele Specific PCR, LGC Genomics) protocol could be used to genotype the OL locus. When applied in early generations of a breeding program this would optimize time, space and economic resources. The technology is based on allele-specific oligo extension and fluorescence resonance energy transfer (FRET) for signal generation (Kumpatla *et al.* 2012).

Since early selection for oil content is highly inefficient due to its low heritability (Camas and Esendal 2006), and multi-environmental trials are often impractical, easy to measure and diagnostic traits for indirect selection are needed. The relatedness between the target breeding population and the assayed materials is crucial because the association between traits may be specific for each genetic material. For this reason, a field assay with F₅ segregating lines and cultivars was performed. Utilizing multivariate analysis, simple traits closely related to the complex trait (grain oil content) were identified. Some yield components gathered from multiple bibliography sources were added to the analysis. The objectives of the present research were: (i) to evaluate the genetic variability in oil content and related traits; (ii) to identify morphological and phenological traits

that best explain the oil content variation among genotypes, including yield components, in order to avoid raising the oil content while decreasing grain yield. Another objective was (iii) to obtain a MAS method to easily and reliably identify the oleic acid allele.

Materials and Methods

Plant material and experimental design

In autumn of 2017, nine genetically related F₅ lines and two test cultivars of safflower were grown at the Asociación de Cooperativas Argentinas (ACA) experimental field, located in Cabildo, Buenos Aires, Argentina (−38.602, −61.974). The segregating F₅ lines were derived from simple crosses: WSRC01 (PI651878) × Ole (PI537695) (lines 2, 3.2, 6.1 and 6.2); WSRC02 (PI651879) × Ole (PI537695) (lines 11 and 12); and WSRC03 (PI651880) × Ole (PI537695) (lines 17.1 and 17.2). The test cultivars were Montola 2000 (PI 651880) and WSRC03 (PI651880).

The experiment was arranged in three randomized complete blocks, in plots (experimental units) with three 4 m rows. The rows were spaced 0.4 m apart, and the distance between plants within rows was 0.2 m, resulting in a plant density of 12 plants per m². Each plot was 0.8 m apart and the blocks were orientated in a northwest-southeast direction, blocking over a natural gradient produced by the predominant wind direction and a minimal field slope (<1%).

Measurements and observations

Crop phenology was recorded according to the BBCH scale proposed by Flemmer *et al.* (2015). The phenological stages were defined when half of the plants of each experimental unit reached a given growth stage.

Data was collected from two plants per plot and then averaged. Morphological traits (plant height, number of primary branches, number of capitula per plant, number of capitula per branch, capitulum diameter) were measured 15 days after anthesis when these traits were already fixed (Flemmer *et al.* 2015). The number of grains and grain weight per capitulum were determined at harvest maturity.

After harvest, grain dimensions were measured with an electronic caliper (mm). Mean geometric diameter and sphericity were calculated as given in Ada (2014). Hull content was determined by manual dehulling of 20 grains expressed as a percentage of the grain weight. The average weight of grain was recorded as the mean weight of three subsamples of 100 and presented as 1000-grain weight. Oil content was determined by milling 150 g grain samples, processed by the Soxhlet method (AOCS Ag 1-65, AOCS 2017), and expressed as a percentage of the dry weight. Then, the fatty acid profile was determined by the quantification of the corresponding methyl esters (FAMES) by GLC (AOCS Ce 2-66, AOCS 2017).

Temperature data was recorded with an automatic weather station (EasyWeather, version 2.0) placed in the

experimental field, and the growing degree days, accumulated in each stage, were calculated as given in Dwyer and Stewart (1986) $((\max T + \min T)/2) - \text{base } T (10^\circ)$ (Tanaka *et al.* 1997).

The experimental results were subjected to an analysis of variance (ANOVA). The relationships between the oil content, and plant and grain variables were assessed using correlation, path, and principal components analysis (INFOSTAT, Di Rienzo *et al.* 2014).

OL allele perfect molecular marker validation

To distinguish oleic (olol), linoleic (OLOL) and heterozygous (OLOl) lines or plants, a set of three primers were designed according to the *FAD2-1* gene sequence (HM165274.1) and its mutant. Hybrids (OLOL × olol: WSRCO3 × Montola 2000) were sown to check the allele calling of heterozygous lines at the OL locus. DNA from F_5 , tests, and hybrids (OLOL × olol: WSRCO3 × Montola 2000), was extracted by the CTAB method (Doyle 1990). 30 ng/ul dilutions were required by the KASP (Kompetitive Allele Specific PCR) assay from KBioscience or LGC Genomics (<http://www.lgcgenomics.com>). This assay was carried out in the GENeTyC lab (CONICET, Bahia Blanca). Primer sequences were: F1 (5'→3') 5'-GAAGG TGACCAAGTTCATGCTGAAAGTTGCAGAGACCTTC T-3'; F2 5'-GAAGGTCGGAGTCAACGGATTGAAAGTT

GCAGAGACCTTCTG-3' and R 5'-GCAGGCGAAACGG TTGTAGG-3'.

Results

Oil content and selection traits

The average oil content (F_5 + test cultivars) was highly variable between genotypes (Table 1). The extreme values of oil content corresponded to WSRCO3 (20.87%) and Montola 2000 (40.33%) whereas the intermediate values were observed in segregating genotypes (Fig. 1). The standard deviation of cultivars (0.64 and 0.85 for WSRCO3 and Montola 2000, respectively) was lower than that of the breeding lines (between 1.17 and 1.76), due to the residual variability within each experimental line (Fig. 1).

The *F* value for genotypic variance was significant for all traits except the number of capitula per plant and duration of anthesis, indicating a strong genetic component (Table 1).

The grain oil content was negatively correlated with hull content, grain thickness, mean geometric diameter and sphericity of the grain (Table 2). On the other hand, positive correlations between the grain oil content and the length:width and the length:thickness ratios of grain indicated that longer and thinner grains tend to have higher oil content. The opposite occurred with more rounded grains

Table 1. Maximum, minimum, mean and standard deviation of the oil content and other 18 traits measured in nine genetically related F_5 lines and two cultivars of safflower sown in ACA-Cabildo (−38.602, −61.974) in 2017. *F*-value and the significance of the genetic component (G) of ANOVA. GDD EA, growing degree days between emergence and anthesis; GDD ESE, growing degree days between emergence and stem elongation.

TRAITS	Min	Max	Mean	SD	<i>F</i> -test for G in ANOVA
Oil content	20.5	41.3	28.9	5.66	41.0**
Morphometric traits of the plant					
Plant height (cm)	88.00	133.00	110.86	9.02	3.7**
No. of primary branches (first order)	8.00	24.50	18.17	3.47	4.5**
No. of capitula per branch	3.10	9.60	4.32	1.20	4.4**
No. of capitula per plant	45.50	133.50	77.59	18.20	1.8 ns
Capitulum diameter (cm)	2.30	3.10	2.62	0.20	6.9**
No. of grains per capitulum	7.00	52.50	33.30	10.70	4.8**
Grain weight per capitulum (g)	0.2	1.9	1.2	0.38	2.7*
Morphometric traits of the grain					
1000-grain weight (g)	27.70	53.70	40.58	7.00	30.5**
Hull content (%)	34.40	55.80	47.65	6.20	29.8**
Length (mm)	6.20	8.30	7.36	0.64	10.2**
Thickness (mm)	2.90	4.20	3.64	0.37	8.5**
Length:width ratio (mm/mm)	1.60	2.50	1.70	0.21	12.5**
Length:thickness ratio (mm/mm)	1.70	2.40	2.03	0.14	2.6*
Geometric mean diameter (mm)	4.10	5.50	4.47	0.40	16.1**
Sphericity	0.60	0.70	0.65	0.03	5.2**
Phenological traits					
Duration of anthesis (days)	10.00	24.00	16.79	2.68	1.0 ns
GDD EA	464.8	605.2	522.75	31.40	3.6**
GDD ESE	38.7	114.4	79.7	20.3	3.3*

ns: non-significant, *: significant at $p < 0.05$; **: significant at $p < 0.01$.

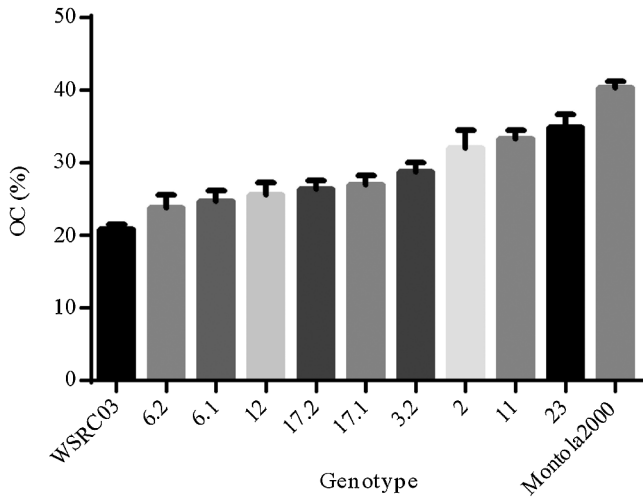


Fig. 1. Oil content (OC) of nine genetically related F₅ lines and two cultivars of safflower (WSRCO3 and Montola 2000) sown in the ACA experimental field in 2017. Error bars represent the magnitude of the standard deviation.

possessing a greater proportion of hull, indicated by a negative correlation with grain sphericity (Table 2). In other words, the grain dimensions and shape were the most meaningful features that affected the hull proportion and thus the oil content. No significant relationship between phenology and oil content was observed.

The analysis of the interactions between traits was necessary because indirect selection for higher grain oil content could result in reduced grain yield, and vice versa. From the correlation analyses it appeared that a higher 1000-grain weight was related to an increased grain geometric diameter and hull content, which in turn had a negative impact on oil content (Table 2). The capitulum diameter was posi-

tively related to the number of grains and grain weight per capitulum. On the other hand, the capitulum diameter showed important negative effects on oil content due to its positive association with hull content (Table 2). The lack of correlation between oil content and number of grains per capitulum suggested that grain yield may be enhanced in this way without decreasing the oil content. Although there was no statistical significance, plant height was almost negatively correlated ($p < 0.076$) with the oil content, meaning that shorter genotypes tend to increase or at least not reduce the oil content.

In turn, among the traits that were significantly correlated with the oil content, hull content ($r_{(partial)} = -0.77$), grain thickness ($r_{(partial)} = -1.22$), and the length:width ratio of the grain ($r_{(partial)} = 0.50$) had the highest direct effect on oil content and in the same direction as the global correlations (Table 3). For the remaining traits, indirect effects through the hull content prevailed, and in other cases, the direct effect was substantial, but with the opposite sign of the global correlation (r). Grain weight per capitulum had a global negative correlation with oil content, predominantly through the indirect effect of hull content. Its potentially negative effect on oil content led to it being discarded as a selection tool for yield improvement, at least in the genetic material used in this study (Table 3).

The principal component analysis (PCA) was carried out with oil content as a dependent attribute and hull content, grain thickness, grain length:width ratio, number of grains per capitulum and plant height as independent attributes. Six traits were reduced to two principal components that explained more than 78% of the total variability (Table 4). The six traits were well represented in the model (sum of squares >0.55) (Table 4), meaning that each vector had enough magnitude to draw conclusions from the graphs.

Table 2. Pearson’s correlation matrix for phenotypic traits. OC, oil content; PH, plant height; 1°B, number of primary branches; CD, capitulum diameter; CPB, capitula per branch; GWC, grain weight per capitulum; GPC, grains per capitulum; 1000GW, 1000-grain weight; HC, hull content; GL, grain length; GT, grain thickness; L:T, grain length:thickness ratio; L:W, grain length:width ratio; MGD, mean geometric diameter; SPH, sphericity; GDD EA, growing degree days between emergence and anthesis; GDD ESE, growing degree days between emergence and stem elongation.

	OC	PH	1°B	CD	CPB	GWC	GPC	1000GW	HC	GL	GT	L:T	L:W	MGD	SPH	GDDEA
PH	-0.31 ns	1														
1°B	-0.09 ns	-0.23 ns	1													
CD	-0.53**	0.19 ns	0.02 ns	1												
CPB	0.27 ns	0.32 ns	-0.43**	-0.38*	1											
GWC	-0.42*	-0.04 ns	0.04 ns	0.52**	-0.45**	1										
GPC	-0.07 ns	-0.02 ns	0.0041 ns	0.52**	-0.41*	0.75**	1									
1000GW	-0.56**	0.04 ns	0.3 ns	0.13 ns	0.0032 ns	0.02 ns	-0.42*	1								
HC	-0.95**	0.26 ns	0.15 ns	0.6**	-0.34 ns	0.52**	0.21 ns	0.43**	1							
GL	-0.22 ns	-0.36*	0.19 ns	-0.01 ns	-0.0027 ns	-0.16 ns	-0.16 ns	0.7**	0.13 ns	1						
GT	-0.53**	-0.11 ns	0.16 ns	-0.01 ns	-0.02 ns	-0.05 ns	-0.05 ns	0.75**	0.42**	0.75**	1					
L:T	0.51**	-0.33 ns	0.02 ns	0.02 ns	-0.02 ns	-0.11 ns	0.2 ns	-0.2 ns	-0.46**	0.16 ns	-0.52**	1				
L:W	0.53**	-0.54**	-0.01 ns	-0.17 ns	-0.09 ns	-0.27 ns	-0.18 ns	-0.25 ns	-0.48**	0.39*	0.04 ns	0.48**	1			
MGD	-0.57**	-0.09 ns	0.2 ns	0.06 ns	0.02 ns	-0.03 ns	-0.03 ns	0.88**	0.45**	0.84**	0.91**	-0.3 ns	-0.11 ns	1		
SPH	-0.58**	0.48**	0.03 ns	0.14 ns	0.001 ns	0.3 ns	0.3 ns	0.27 ns	0.54**	-0.36*	0.2 ns	-0.75**	-0.88**	0.18 ns	1	
GDD EA	0.25 ns	-0.18 ns	0.19 ns	-0.0012 ns	-0.17 ns	0.12 ns	0.12 ns	-0.44**	-0.18 ns	-0.32 ns	-0.34*	0.1 ns	0.02 ns	-0.37**	-0.02 ns	1
GDD ESE	-0.1 ns	-0.14 ns	0.14 ns	-0.14 ns	-0.01 ns	-0.25 ns	-0.25 ns	0.06 ns	0.07 ns	0.17 ns	0.18 ns	-0.09 ns	0.08 ns	0.16 ns	-0.14 ns	-0.11 ns

ns: non-significant, *: significant at $p < 0.05$; **: significant at $p < 0.01$.

Table 3. Path analysis for oil content. Direct (in the diagonal) and indirect effects (out of the diagonal) of each significantly correlated trait. At the bottom, the global *r* coefficient and the error probability. CD, capitulum diameter; GWC, grain weight per capitulum; 1000GW, 1000-grain weight; HC, hull content; GT, grain thickness; L:T, grain length:thickness ratio; L:W, grain length:width ratio; MGD, mean geometric diameter; SPH, sphericity.

	CD	GWC	1000GW	HC	GT	L:T	L:W	MGD	SPH
CD	-0.063	-0.033	-0.008	-0.038	0.001	-0.001	0.011	-0.004	-0.007
GWC	0.033	0.062	0.001	0.032	-0.004	-0.007	-0.017	-0.002	0.017
1000GW	0.003	0.000	0.025	0.011	0.018	-0.005	-0.006	0.022	0.006
HC	-0.465	-0.401	-0.333	-0.769	-0.323	0.356	0.367	-0.347	-0.409
GT	0.018	0.069	-0.908	-0.514	-1.222	0.640	-0.048	-1.117	-0.225
L:T	-0.010	0.060	0.109	0.247	0.279	-0.534	-0.253	0.158	0.415
L:W	-0.086	-0.134	-0.126	-0.239	0.020	0.237	0.500	-0.057	-0.453
MGD	0.044	-0.022	0.698	0.356	0.722	-0.234	-0.090	0.790	0.144
SPH	-0.008	-0.020	-0.018	-0.039	-0.013	0.057	0.066	-0.013	-0.073
<i>r</i> _{total}	-0.533	-0.417	-0.561	-0.953	-0.523	0.509	0.530	-0.570	-0.585
<i>P</i> _{error}	0.00139	0.01566	0.00069	<0.00001	0.00181	0.0024	0.00151	0.00053	0.00035

Table 4. Representation of each trait in the PC system (eigenvectors). The sum of squares (SS) of each relevant trait reveals the relative magnitude of the vector represented in the graph ((EigenvectorPC1)² + (EigenvectorPC2)²). HC, hull content; OC, oil content; GPC, grains per capitulum; L:W, grain length:width ratio; GT, grain thickness; PH, plant height.

	PC1	PC2	SS
HC	-0.91	0.13	0.85
OC	0.95	-0.23	0.96
GPC	-0.19	-0.77	0.63
L:W	0.80	0.33	0.75
GT	-0.35	0.92	0.97
PH	-0.64	-0.38	0.55
Variation explained (%)	49	29	78

WSRC03 and Montola 2000 were very influential data points (more distance to 0) over the PC1, displaying extreme characteristics due to their different genetic backgrounds (Fig. 2). Oil content and the length:width ratio of the grain were mainly represented by PC1, and the opposite was the case with hull content and plant height (Fig. 2). A right angle between number of grains per capitulum and oil content vectors demonstrated the lack of correlation, and therefore indirect selection for grain yield improvement should not trigger a decrease in the oil content (Fig. 2).

OL allele perfect molecular marker validation

The fatty acid profile of the assayed F₅ lines and the test cultivars was perfectly correlated with the genotypic data from the KASP reactions. Allele calling was consistent with the phenotypic data. The lines that were homozygous for the mutant allele were unequivocally high oleic types (olol) and the lines homozygous for the wild type allele were linoleic types (OLOL) (Fig. 3).

All the F₅ lines assayed were scored as homozygous for the OL locus, so some hybrids (WSRC03 × Montola 2000) were added to the assay to check the efficiency of the protocol in genotyping heterozygous individuals. The hybrids

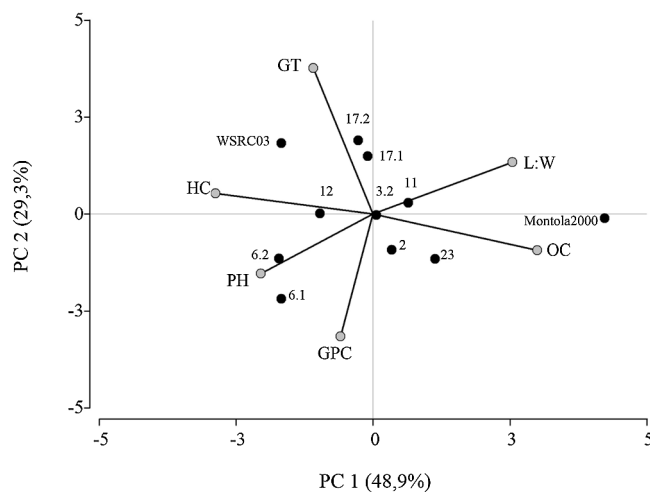


Fig. 2. Biplot of the two more explicative components of the PCA, showing the interrelationships among traits (grains per capitulum (GPC), plant height (PH), grain thickness (GT), grain length:width ratio (L:W), oil content (OC) and hull content (HC)) and genotype data of nine lines and two cultivars (WSRC03 and Montola 2000) of safflower.

were accurately scored with the heterozygous state at the OL locus (Fig. 4).

Discussion

Through the evaluation of F₅ lines and test genotypes, this study generated indirect selection tools for oil content. Among the 18 traits initially analyzed, the successive multivariate analyses identified hull content, grain length:width ratio, number of grains per capitulum, and plant height as the best diagnostic characteristics. Furthermore, a MAS tool was developed and successfully validated for early characterization of the fatty acid profile.

It has been already reported that grain oil content is negatively related to hull content (Ada 2014, Rao *et al.* 1977, Sai Santhosh 2015). The grain (botanically fruit) is

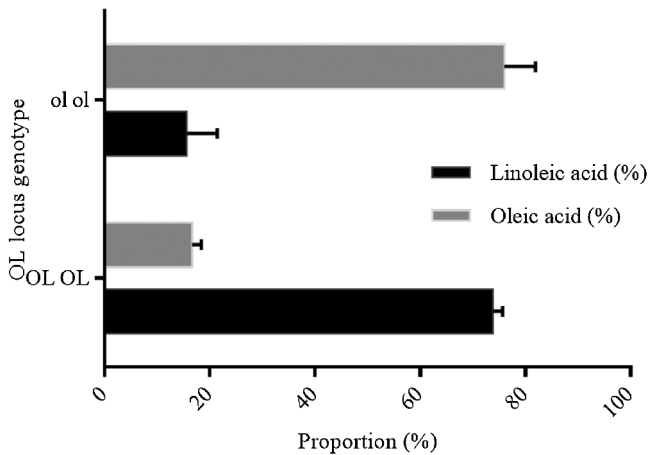


Fig. 3. Genotype/phenotype. Allelic composition on the OL locus versus its fatty acid proportions.

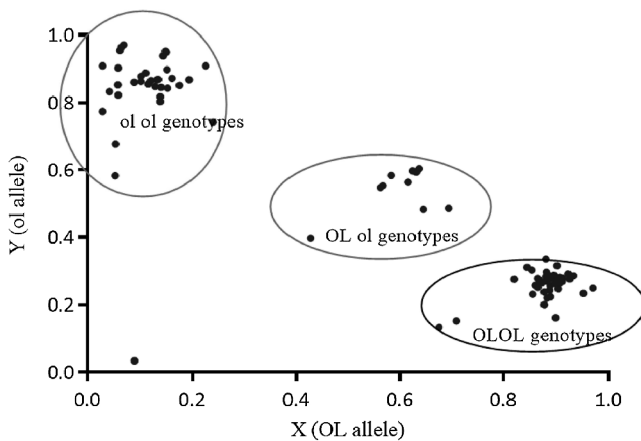


Fig. 4. KASP screening of the F_5 lines, test cultivars and hybrids (heterozygosity tests). X and Y axis represent the FAM and HET dyes fluorescent signals, each representing one allele.

composed of the hull, seed teguments, and the embryo, and in the absence of variability in the oil content of the embryo itself, a higher proportion of hull results in a lower proportion of embryo, where the fatty compounds are stored. The hull of the grain is high in fibers whereas the embryo is rich in lipids and proteins, resulting in a negative relationship between the hull and the oil contents (Urie 1986). Franchini *et al.* (2014) studied grain development in safflower and determined that the hull growth and lignification end eight days after flowering. At this point, the potential seed size is fixed, because the inner parenchyma of the pericarp is highly lignified and cannot be compressed by the developing seed. Between that critical point and the grain filling process, the hull content is finally fixed (Franchini *et al.* 2014). As stated by Li and Mündel (1996) a commercial genotype should not exceed 50% hull content. Rao *et al.* (1977) also found that hull proportion has a strong and negative relationship with the oil content ($r = -0.83$) in a set of 215 safflower genotypes. According to our results, out of a set of 18 traits, the hull content was the most promising

characteristic to be used for indirect selection for oil content in segregating populations. In other safflower studies, the grain size and geometry had a great effect on hull content (Ada 2014). In the present work, oil content was negatively related to the mean geometric diameter, but positively related with the length:width ratio of the grain. Consistent with our results, Sai Santhosh (2015) and Wichman (1983) found a strong negative relationship between grain size and oil content. But contrary to our findings, Wichman (1983) established that longer grains have lower oil content. The explanation for this contradiction is the existence of huge genetic variability for grain shape and density (Sai Santhosh 2015). In accordance with this, Wichman (1983) found varying patterns in grain shape, and its relationship with oil content varied among genotypes.

If a yield-related trait is used for indirect selection without analyzing its relationships with other aspects of the crop, undesired effects may be carried into the breeding process (Bagavan and Ravikumar 2001). Arslan (2007), Jagtap *et al.* (2012) and Rao *et al.* (1977) observed negative correlations between the oil content and yield. In the same way, some yield components are negatively correlated with oil content (Jagtap *et al.* 2012). Grain weight is one of the most important yield components (Bidgoli *et al.* 2006) and in this and other studies, it was negatively associated with oil content (Rao *et al.* 1977). In this context, enhancing yield without decreasing the oil content or increasing the hull content appears challenging. From the present study, the relationships between yield components and oil content were recognized. In this set of genotypes, 1000-grain weight, grain weight per capitulum, and capitulum diameter seemed to be unpromising selection tools for improving the yield in early generations, due to their negative potential impact on the grain oil content (Table 2). Branch and capitula number per plant were seen to be positively associated with yield (Moghaddasi and Omidi 2010), but some authors have emphasized their low heritability and high genotype by environment interaction (Rao *et al.* 1977). In addition, selection for branch and capitula number per plant may result in extended and desynchronized flowering period (Bellé *et al.* 2012, Singh and Nimbkar 2016, Tanaka *et al.* 1997). Also, the grain shape and size would be more irregular and the harvest losses much greater, making it agronomically non-viable (Pascual-Villalobos and Albuquerque 1995, Singh and Nimbkar 2016). Moreover, these two features can change radically when the crop is taken from experimental fields to commercial ones, due to plasticity in plant density and yield in response to environmental conditions.

Some authors found a positive relationship between grain yield and the number of grains per capitulum (Arslan 2007, Golkar *et al.* 2011), and in our dataset, the number of grains per capitulum did not show any significant negative correlation with the oil content (Table 2). Thus it appeared to be the yield component that best fits the objectives of this work. On the other hand, selection for shorter

genotypes is suggested by the negative relationship ($p < 0.076$) between plant height and oil content, although not significant at $p < 0.05$ (Table 2). Pahlavani (2005) observed no relationship between grain yield and plant height while significant positive relationships have been reported by other authors (Hussain *et al.* 2014, Pavithra *et al.* 2016). If the tendency towards breeding dwarf genotypes that improves the harvest index in almost all the major crops is considered, it may be valuable to select for reduced height genotypes in safflower (Ashri *et al.* 1975, Singh and Nimbkar 2016).

Seker and Serin (2004) established that a close relationship between traits does not ensure selection success, so many authors have used path analysis to distinguish between direct and indirect effects of the independent variables on the important dependent trait (Wright 1921). Identifying simple characteristics to be used as indirect selection tools in breeding programs is crucial (Bahmankar *et al.* 2014, Mahasi *et al.* 2006). Among the statically correlated traits, hull content, grain thickness, and the length:width ratio of the grain were recognized as the most promising selection tools, given their direct effect on the main trait (grain oil content) (Table 3). In similar studies but in contrast to our results, Golparvar (2011) and Karimi *et al.* (2013) found that the 1000-grain weight was one of the most reliable selection tools for oil yield improvement.

The biplot of principal components gives a graphical representation that explains the more relevant features of the analysis and displays the interrelations between characters and genotypes (Gabriel 1971). The main characteristic was the grain oil content, and the independent traits were hull content, grain length:width ratio, grain thickness, plant height, and number of grains per capitulum (Fig. 2). This model indicated that selecting short plants or lines with long, thin and low hull content grains could be promising for oil content improvement, and with a high number of grains per capitulum, the grain yield may be increased at the same time. Although grain thickness seemed to be a promising tool, the high variability between grains from branches of different order and between differently positioned grains of the same capitulum (pers. Obs.) led to it being discarded. A similar situation was reported by Urie and Zimmer (1970), who found high heterogeneity in the hull thickness in grain of different branch orders, and different positions in the capitulum. Accordingly, using grain dimensions as a selection criterion could be inaccurate. Instead, the use of dimension ratios such as the length:width ratio of the grain could properly represent genetically determined and stable grain attributes.

Environmental effects on the quantitative traits such as oil content are well known (Li 1993). The environmental component plus the genetic variability and the associated interaction can modify the relation between the indirect selection tools and the target complex traits. Therefore, in large scale breeding projects, generating information in a preliminary population in the same environment (location)

and using genotypes of similar genetic background may enhance the accuracy of the decisions regarding breeding material. The simple field assay may include a few experimental lines or genotypes that represent the genetic resources selected in the breeding plan. This procedure potentially recognizes the best tools for indirect selection of complex traits in early generations (F_2 - F_4). In these stages, progeny could not be evaluated in replicated field trials due to the limited amount of seed and the large number of lines.

The main genetic resources used in safflower breeding activities belong to germplasm collections in gene banks, where linoleic and oleic accessions are conserved (Fernandez-Martinez *et al.* 1993). Breeders often include linoleic and oleic accessions in their crossing blocks depending on the oil type required by the specific industrial application (Li and Mündel 1996). As a result, most breeding populations segregate for the OL locus. So a practical tool for detecting OL alleles in early generations is highly useful. Oleic types were distinguished from linoleic (homozygous and heterozygous) types by the design and phenotypic validation of a KASP protocol (LGC Genomics), resulting in a valuable tool for the recurrent selection for the OL allele.

In conclusion, this study contributed to the knowledge of the most relevant traits related to oil content and grain yield in safflower. Simple traits were identified that could be utilized in the characterization of germplasm accessions, as indirect indicators of complex traits that could not be phenotyped in a large-scale collection or breeding project. These procedures may serve as a model to identify selection tools for complex traits in safflower and other commercial crops. At the same time, an innovative molecular tool for distinguishing the fatty acid profile was validated.

Author Contribution Statement

LLI contributed with the study conception and design. EV worked on the data acquisition and critical revision. The corresponding author (CA) contributed to the acquisition, analysis and interpretation of the data, and wrote the manuscript.

Acknowledgments

We gratefully acknowledge Dr. Richard Johnson from the USDA-ARS Plant Germplasm Unit at Pullman (WA, USA) for providing the initial plant material, and CONICET for the author's fellowship. The study was funded by the "Secretaría General de Ciencia y Tecnología, Universidad Nacional del Sur" (24/A234) and grants from Oleaginosa Moreno Hermanos S.A., Bahía Blanca, Argentina.

Literature Cited

Ada, R. (2014) Dimension, geometric, agricultural and quality characteristics of safflower seeds. *Turkish J. Field Crops* 19: 7–12.

- AOCS (2017) Official Methods and Recommended Practices of American Oil Chemist.
- Arslan, B. (2007) The path analysis of yield and its components in safflower (*Carthamus tinctorius* L.). *J. Biol. Sci.* 7: 668–672.
- Ashri, A., D.E. Zimmer, A.L. Urie and P.F. Knowles (1975) Evaluation of the germ plasm 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.
- Bagavan, I. and R.L. Ravikumar (2001) Strong undesirable linkages between seed yield and oil components—A problem in safflower improvement. *Proceeding of the V International Safflower Conference, USA*, pp. 103–107.
- Bahmankar, M., D.A. Nabati and M. Dehdari (2014) Correlation, multiple regression and path analysis for some yield-related traits in safflower. *J. Biodivers. Environ. Sci.* 4: 111–118.
- Baydar, H.A.S.A.N., O.Y. Gökmen and W. Friedt (2003) Hybrid seed production in safflower (*Carthamus tinctorius*) following the induction of male sterility by gibberellic acid. *Plant Breed.* 122: 459–461.
- Bellé, R.A., E.K.D. Rocha, F.A.A.L. Backes, M. Neuhaus and N.T. Schwab (2012) Safflower grown in different sowing dates and plant densities. *Cienc. Rural* 42: 2145–2152.
- Bidgoli, A.M., G.A. Akbari, M.J. Mirhadi, E. Zand and S. Soufizadeh (2006) Path analysis of the relationships between seed yield and some morphological and phenological traits in safflower (*Carthamus tinctorius* L.). *Euphytica* 148: 261–268.
- Camas, N. and E. Esendal (2006) Estimates of broad-sense heritability for seed yield and oil components of safflower (*Carthamus tinctorius* L.). *Hereditas* 143: 55–57.
- Di Rienzo, J.A., F. Casanoves, M.G. Balzarini, L. Gonzalez, M. Tablada and C.W. Robledo (2014) Software InfoStat 2014. Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina.
- Doyle, J.J. (1990) Isolation of plant DNA from fresh tissue. *Focus* 12: 13–15.
- Dwyer, L.M. and D.W. Stewart (1986) Leaf area development in field-grown maize. *Agron. J.* 78: 334–343.
- Fernandez-Martinez, J., M. Del Rio and A. de Haro (1993) Survey of safflower (*Carthamus tinctorius* L.) germplasm for variants in fatty acid composition and other seed characters. *Euphytica* 69: 115–122.
- Flemmer, A.C., M.C. Franchini and L.I. Lindström (2015) Description of safflower (*Carthamus tinctorius*) phenological growth stages according to the extended BBCH scale. *Ann. Appl. Biol.* 166: 331–339.
- Franchini, M.C., A.C. Flemmer, L.I. Lindström, M.A. David and P.A. Fernandez (2014) Fruit development of two high oleic safflower (*Carthamus tinctorius* L.) cultivars. *Phyton, International Journal of Experimental Botany* 83: 379–388.
- Gabriel, K.R. (1971) The biplot graphic display of matrices with application to principal component analysis. *Biometrika* 58: 453–467.
- Golkar, P., A. Arzani and A.M. Rezaei (2011) Determining relationships among seed yield, yield components and morpho-phenological traits using multivariate analysis in safflower (*Carthamus tinctorius* L.). *Ann. Biol. Res.* 2: 162–169.
- Golparvar, A.R. (2011) Genetic improvement of oil yield in spring safflower cultivars under drought and non-drought stress conditions. *Electronic Journal of Biology* 7: 40–43.
- Guan, L.L., Y.B. Wang, H. Shen, K. Hou, Y.W. Xu and W. Wu (2012) Molecular cloning and expression analysis of genes encoding two microsomal oleate desaturases (FAD2) from safflower (*Carthamus tinctorius* L.). *Plant Mol. Biol. Rep.* 30: 139–148.
- Gunstone, F.D. (2001) Basic oleochemicals, oleochemical products, and new industrial oils. *In: Oleochemical manufacture and application*, Sheffield Academic Press, Sheffield, pp. 1–22.
- Hamdan, Y.A.S., M.J. Garcia-Moreno, J.M. Fernández-Martínez, L. Velasco and B. Pérez-Vich (2012) Mapping of major and modifying genes for high oleic acid content in safflower. *Mol. Breed.* 30: 1279–1293.
- Hussain, T., M.A. Tariq, I. Ahmad, M. Saghir, M. Batool, M. Safdar, A. Sher and M. Tariq (2014) Characters association analysis in safflower (*Carthamus tinctorius* L.). *J. Biol. Agric. Healthc.* 4: 63–65.
- Jagtap, P.K., A.A. Chavan, B.N. Joshi and A.M. Misal (2012) Nature of association and path analysis among different quantitative traits in early segregating generations of safflower. *J. Oilseeds Res.* 29: 57–59.
- Karimi, M., A.R. Golparvar and M. Shams (2013) Multivariate regression and path coefficient analysis of seed and oil yield in spring safflower (*Carthamus tinctorius* L.) cultivars. *Applied Science Reports* 4: 184–186.
- Knott, D.R. (1972) Effects of selection for F₂ plant yield on subsequent generations in wheat. *Can. J. Plant Sci.* 52: 721–726.
- Knowles, P.F. (1989) Safflower. *In: Röbbelen, G., R.K. Downey and A. Ashri (eds.) Oil Crops of the World*, McGraw-Hill, New York, pp. 363–374.
- Kumpatla, S.P., R. Buyyarapu, I.Y. Abdurakhmonov and J.A. Mammadov (2012) Genomics-assisted plant breeding in the 21st century: technological advances and progress. *In: Abdurakhmonov, I. (ed.) Plant breeding*. InTech publishers, <http://www.intechopen.com/books/plantbreeding/genomics-assisted-plant-breeding-in-the-21stcenturytechnological-advances-and-progress>.
- Li, D. (1993) Progress of safflower (*Carthamus tinctorius* L.) research and production in China. *In: Third interactional safflower conference*, Beijing China, pp. 35–46.
- Li, D. and H.H. Mündel (1996) Safflower *Carthamus tinctorius* L. Promoting the conservation and use of underutilized and neglected crops. Institute of Plant Genetics and Crop Plant Research, Gatersleben. *Int Plant Genetic Resources Institute*, Rome, pp. 1–83.
- Liu, Q., S. Cao, X.R. Zhou, C. Wood, A. Green and S. Singh (2013) Nonsense-mediated mRNA degradation of CtFAD2-1 and development of a perfect molecular marker for olol mutation in high oleic safflower (*Carthamus tinctorius* L.). *Theor. Appl. Genet.* 126: 2219–2231.
- Lunn, J. and H.E. Theobald (2006) The health effects of dietary unsaturated fatty acids. *Nutr. Bull.* 31: 178–224.
- Mahasi, M.J., R.S. Pathak, F.N. Wachira, T.C. Riungu, M.G. Kinyua and J.W. Kamundia (2006) Correlations and path coefficient analysis in exotic safflower (*Carthamus tinctorius* L.) genotypes tested in the arid and semi arid lands (asals) of Kenya. *Asian J. Plant Sci.* 5: 1035–1038.
- Markley, N., C. Nykiforuk, J. Boothe and M. Moloney (2006) Producing proteins using transgenic oil body-oleosin technology. *Biopharm. Int.* 19: 34–46.
- Moghaddasi, M.S. and A.H. Omid (2010) Study of some morphological and phenological traits in exotic and native safflower genotypes using multivariate statistical methods. *Adv. Environ. Biol.* 4: 350–352.
- Pahlavani, M.H. (2005) Some technological and morphological characteristics of safflower (*Carthamus tinctorius* L.) from Iran.

- Asian J. Plant Sci. 4: 234–237.
- Parmeshwarappa, K.G., K.M. Aradhya and K.S. Prakash (1984) Inheritance of seed yield, hull and oil content in safflower. SABRAO J. Breed. Genet. 16: 129–134.
- Pascual-Villalobos, M.J. and N. Alburquerque (1995) Genetic variation of a safflower germplasm collection grown as a winter crop in southern Spain. Euphytica 92: 327–332.
- Patel, J. and G. Narayana (1935) Chromosome numbers in safflower. Curr. Sci. 4: 412.
- Pavithra, K.P., R.S. Patil, Y. Harijan and G.K. Nishanth (2016) Correlation and path analysis studies in safflower (*Carthamus tinctorius* L.) germplasm. Research Journal of Agricultural Sciences 7: 428–432.
- Rao, V.R., M. Ramachandram and V. Arunachalam (1977) An analysis of association of components of yield and oil in safflower (*Carthamus tinctorius* L.). Theor. Appl. Genet. 50: 185–191.
- Ribaut, J.M. and D. Hoisington (1998) Marker-assisted selection: new tools and strategies. Trends Plant Sci. 3: 236–239.
- Sai Santhosh, K. (2015) Studies on seed morphological, physiological and biochemical traits of safflower (*Carthamus tinctorius* L.) genotypes in relation to oil content. M. Sc. (Ag.) thesis, submitted to Professor Jayashankar Telangana State Agricultural University, Hyderabad, Telangana State, India.
- Seker, H. and Y. Serin (2004) Explanation of the relationships between seed yield and some morphological traits in smooth brome grass (*Bromus inermis* Leys.) by path analysis. Eur. J. Agron. 21: 1–6.
- Singh, V. and N. Nimbkar (2016) Safflower. In: Gupta, S.K. (ed.) Breeding Oilseed Crops for Sustainable Production, Elsevier Inc, Jammu, India, pp. 149–167.
- Tanaka, D.L., N.R. Riveland, J.W. Bergman and A.A. Schneiter (1997) Safflower plant development stages. Proceedings of IVth International Safflower Conference, Bari, pp. 179–180.
- Udall, J.A. (2003) Breeding for quantitative traits in plants. Crop Sci. 43: 1578–1580.
- Urie, A.L. and D.E. Zimmer (1970) A reduced-hull seed character of safflower. Crop Sci. 10: 371–372.
- Urie, A.L. (1986) Inheritance of partial hull in safflower. Crop Sci. 26: 493–498.
- Vafaei, S.N., A. Tobeh, A. Tae and S. Jamaati-e-Somarin (2010) Study of phenology, harvest index, yield, yield components and oil content of different cultivars of rain-fed safflower. World Appl. Sci. J. 8: 820–827.
- Wichman, D.M. (1983) Evaluation of safflower seed size, length (shape), and density in relation to seed vigor and oil content. Doctoral dissertation, Montana State University-Bozeman, College of Agriculture.
- Wright, S. (1921) Correlation and causation. J. Agric. Res. 20: 557–585.