Research Article

Katerina Vedmedeva*

Genetic affinity of sunflower lines and cluster analysis by morphological traits

https://doi.org/10.1515/helia-2020-0014
Received May 23, 2020; accepted September 23, 2020;
published online October 1, 2020

Abstract: The purpose of the study is to identify the use of qualitative and quantitative morphological traits to ascertain genetic affinity and identification of sunflower lines. A collection of 39 sunflower lines was studied according to morphological qualitative traits described in the method of establishing differences, homogeneity, stability and quantitative traits (plant organ size, 1000 seeds weight, oil content). Several lines of unknown origin were identified to each other. The material of the collection of lines proved the possibility of clustering by the method of link analysis on a set of morphological features. Generalized data show that only three lines out of 39 do not correspond to known lineages in their clusters, which is 92% of the correct cluster definition. The results of clustering, identification and breeding records were compared. To be consistent with the breeding records classification, the number of distinguished traits that are not similar must exceed 20 names given to the score.

Keywords: cluster analysis; genetic affinity; line; morphological trait; sunflower.

Introduction

Maintaining sunflower collections requires determining their value and affinity. It is necessary to reduce the number of sample of plants to optimize performance work. Collectors make a description according to all possible traits and are guided to preserve them by their opinion and preference for a certain set of traits. More than one study of the morphological characteristics of plants of wild species is known to establish phylogenetic and kinship relationships. So, for example, when studying the morphology of *Helianthus anomalus* in comparison with its putative

***Corresponding author: Katerina Vedmedeva**, Institute of Oilseed Crops of NAAS, Institutskaya st., 1. Solnechnyi, Zaporizhzhia 70417, Ukraine, E-mail: vedmedeva.katerina@gmail.com; vedmedeva_k@mail.ru

ancestors: *Helianthus annuus* and *Helianthus petiolaris* confirmed the mosaic nature of the *H. anomalus* species. *H. anomalus* was morphologically intermediate for one trait (2.4%), parental-like for 23 traits (56.1%), and transgressive for 17 traits (41.5%,) (Schwarzbach et al. 2001).

To clarify the presence of gene flow between wild and cultivated sunflower, morphological traits and molecular technologies are sometimes used simultaneously (Gutierrez et al. 2010).

The modern establishment of genealogies and genetic proximity of lines, varieties and species of plants in the world is based on the application of molecular marking technologies (Darvishzadeh et al. 2010; Jan et al. 1998; Liu et al. 2001; Sivolap and Solodenko 1998). The study of sunflower breeding material for a specific purpose using molecular markers often shows a high uniformity of genotypes (Kholghi et al. 2012).

In some studies, the data obtained using the latest methods were compared with the known origin of the genetic material and their conformity was established (Whankaew et al. 2014; Yang et al. 2018). Studies of other species and their groups for morphological traits and molecular markers sometimes show significant inconsistencies between molecular markings and historical species distribution (Wood and Nakazato 2009).

Therefore, it becomes clear that the accuracy of establishing family ties does not so much depend on the modernity of the selected technology as on the correctly selected set of markers. Under specific conditions, a set of morphological markers can cover a larger part of the active genome than a set of molecular markers.

When describing and calculating molecular mapping methods, methods of cluster analysis are used. When calculating morphological traits in intraspecific collections, cluster analysis will also be objective when expressing traits in a point scale. For example, Zeinalzade-Tabrizi et al. (2018) used the arithmetic mean clustering algorithm. This method allows you to weigh the number of line differences and create a dendrogram.

Until now, there are practically no studies of cultivated sunflower lines using a comprehensive assessment of a large number of morphological characters using the link analysis method.

We assumed that the dendrogram of the collection lines obtained using the assessment of morphological characters would reflect their real relationship and help to more objectively select samples for maintenance in collections. To test the hypothesis, a set of lines of different morphotypes with a known relationship between individual lines was taken in the study.

In the breeding programs very careful attention is given to the qualitative and quantitative morphological traits, usually breeders focus on them in their

selections. Morphological traits are a display of the same genes or they're complex and they can also establish the genetic affinity of the plant material.

The purpose of the study is to identify the use of qualitative and quantitative morphological traits to ascertain genetic affinity and identification of sunflower lines.

Materials and methods

A collection of 39 sunflower lines was used as the study material. A part of sunflower lines that were included in the experiment during 2016–2018 was created as an analogue line by back-crossing to introduce into the genome a gene of morphological marker-trait, and part of lines by crossing similar in morphology and conducted by crossing identification, and part of known lines by crossing with their parent form created by selection. The lines were evaluated by 34 morphological features, the weight of 1000 seeds, oil content, seed size (UPOV TG/81/5 2000).

The studies were conducted in the scientific rotation planting of the Institute of Oil Crops of NAAS. The technology of cultivation is classical, planting by hand, observation in all phases of plant development is classical, measuring of oil content was conducted using NMR.

Statistical analysis of the results was carried out by methods of link analysis (Vorob'yev et al. 2006). For cluster analysis, Statistica used the full range of features, except those with no differences. The program allows you to create dendrograms in different ways: Ward's method, single, full, unweighted pair mean, weighted pair mean, unweighted and weighted centroid methods.

Results and discussion

According to previous molecular studies (Vedmed'va et al. 2010), the difference between the genetic distance calculated by molecular markers and the number of significant differences in morphological features in the group of KLV80/1 breeding line analogues was established. A greater genetic distance from the baseline of the M10 mutant was detected. For further research, the scope of observation features and the number of lines were expanded.

Thirty-nine lines included in the experiment were divided into groups of known origin. Lines of unknown origin, but similar in morphotype were crossed for genetic identification. A total number of groups were 7 (Table 1).

The first group includes two lines KG16 and InK1589, they both had raised leaf veins and after crossing gave first-generation hybrids similar in all morphological characters with the parental forms, but with a slightly larger expression of quantitative features (head diameter, plant height). In the second generation, splitting by some contrasting features was not observed. Samples from the KG16 collection

Table 1: Affinity of the sunflower lines and their specific morphological characteristics.

No.	Line name	Number of the registration in NCofGR Ukraine	Origin, identification	Morphological description	
1	КG16	UE0100523	Direct and backcrossings	Basal branch, elongated leaves with enlarged vein	
2	InK1589				
3	LD72/1	UE0100567	Collected from VIR collection in 1991	Basal branch, elongated leaves, long petioles, similar colour of the seeds	
4	LD72/2	UE0100568			
5	LD72/3	UE0100569			
6	LD4/2	UE0100566			
7	LD4	UE0100566			
8	InK404	UE0100533	VIR collection		
9	I2K224-2	UE0100595			
10	VIR199	UE0100549	VIR collection	Dwarfness, short internode, big head, late maturity	
11	VIR501	UE0101066			
12	InDH47	UE0100546	Derived from Donskoy nyzkoroslyi		
13	DH47-8	UE0100616			
14	DH47-2	UE0100613			
15	DH47-9	UE0100617			
16	ZL169/431		Backcross of ZL169	lb light brown leaves	Early maturity, light yellow colour of the ray flowers
17	MV1	UE0101003	Mutant lines ZL169	y yellow terminal bud	
18	MV3	UE0101004	(physical mutagenesis)	dw dwarfness	
19	ZL169	UE0100700	Initial line	Initial line	
20	VA1B	UE0100456	Identification		
21	VA2B	UE0100519			
22	ZL22B	UE0100696	Identification	Selected line	
23	HA89B	UE0100460		Initial line for ZL22B	Dwarfness with big head, late maturity
24	ZL22/319	UE0101290	Backcross of ZL22B	lb light-brown leaves	
25	ZL22/320	UE0101291		Fr fringed leaf margin	
26	ZL22Бахр				
27	ZL22/434	UE 0101292		Dw dwarfness	
28	KLV80/1	UE0101175		Selected line	
29	VK580	UE0101266	Initial KLV80/1		Branching up — part of the plant
30	M791	UE0101279	Backcross of KLV80/1	lb light-brown leaves,	
31	M790	UE0101244		lb, Dw light-brown leaves, dwarfness	
32	M10	UE0101254°	Natural mutant of	Orange ray flowers	
33	M17	UE0101245	КЛВ80/1 on the site of reproduction	shc striped-shaped ray flowers	
34	M19	UE0101242		l lemon colour ray flowers	
35	M23	UE0101255		ly light-yellow ray flowers	
36	M17/1	UE0101243		tu2 long tubular ray flowers	

Table 1: (continued)

No.	Line name	Number of the registration in NCofGR Ukraine	Origin, identification	Morphological description
37	InK235	UE0100919	Identification	Top branching of the plant, similar colour of seeds and flowers
38	APS 56	UE0100882		
39	InK103	UE0100885		

were isolated from the VIR collection, which is why InK self-pollination of many generations could lead to the same genetic set.

The second group consisted of seven lines that had a basal branch, elongated leaves, long petioles, similar colouration of seeds. The ancestors of the first five lines were collected from the VIR sites about 30 years ago. InK404 and I2K224-2 lines were obtained by self-pollination of known VIR collection numbers. The identification of lines by crossing was conducted between three lines InK404, I2K224-2 and LD72/2.

The third group includes six lines, four of which are of the same origin from the variety Donskoy nizkoroslyi, and two are derived from the VIR, but we do not know the origin of them. Crossing between lines VIR, VIR501 and MV1, confirmed their identity. They all have a low height, short internodes, late maturity and large head.

The fourth group consist of six lines, four of which are related in origin: one is the source line, one of them created by backcrossing and two by chemical mutagenesis. The VA1B and VA2B lines are identified between each other, however, identification with the whole group has not been made and they are classified to this group only by general appearance.

The fifth group includes six lines of common origin and were created from each other by backcrossing. The ZL22B and HA89B lines were identified by crossing each other and did not show significant differences in the first and second generation of hybrids.

The sixth group consist of lines of common origin, some of which were included in the study using molecular markers. This group includes breeding lines created one on the other, two created by backcrossing, four by natural mutagenesis and another by repeated natural mutagenesis.

The seventh group consists of three lines named different, but as such, they have only a few slight differences in the traits. They were identified and the differences were determined for individual genes, so the branching feature is represented in all lines by the dominant state of the genes, but their number in each line

is different (from one to three). The InK235 line is characterized by the appearance of tubular ray flowers that do not have complete penetrance of the trait.

The formed collection of lines was studied by the combination of all morphological traits described in the methodology for Distinctness, Homogeneity and Stability (UPOV TG/81/5 2000). Table 2 provides a list of qualitative traits, their gradations, and the number of differences found for each trait in the collection of investigated lines.

Signs of quantitative nature: plant height, number of leaves, leaf size and petiole, oil content, the weight of 1000 seeds were calculated in units of measurement and used in the mean values over the next three years. Conducting further analysis using absolute averages, quantitative traits have a greater impact, as they have a greater absolute expression.

All qualitative traits had expression from one to nine, while for example, the height of plants ranged from 60 to 160 cm. To balance the impact of each trait, we split the available variability for each trait on the corresponding scale, which has the expression of the same nine points. The resulting collection description matrix after the recalculation had the same numerical expression: from 1 to nine.

Evaluating the appearance of each of the methods, we found the most appropriate method of full communication, the results of which are presented in Figure 1.

The imaginary line of reliable groups is performed according to the indicator nine – which is half of the detected variability.

According to the dendrogram, the first, largest cluster consisting of nine lines appeared to be related. Similarly, in our molecular analysis study (Vedmed'va et al. 2010), these lines appeared to be related. However, the use of molecular markers has been highlighted as an excellent M10 line. On the dendrogram, this line was in the first cluster.

The second cluster consisted of three lines that were identified as sister lines by crossing each other: APS56, InK103 and InK235.

The next cluster included three lines DN47-2, ZL169 and MV3. DN47-2 by known lineages should not have fallen into this cluster.

The fourth cluster included seven lines, of which one subgroup contained ZL169/431 and MV1 – related to the previous cluster, namely two lines from ZL169 and MV3, and the second subcluster contained the third group lines from Table 1.

The fifth cluster is represented by two lines that, according to our data, had no genetic closeness. The sixth cluster fully corresponded to the third group of related lines. The seventh and eighth are represented by separate lines. The ninth cluster almost completely corresponded to the fifth group.

Summarizing the result, a genetic affinity for morphological traits was confirmed by the cluster hierarchical analysis method for 28 lines out of 39,

Table 2: Deciphering the qualitative characteristics of morphological features and codes of their expression by DHS.

No.	Traits	Codes of expressions	Differences
4.	Leaf: the shape of the distal part	1–9	4
5.	Leaf: Auricules	1, 3, 5, 7, 9	7
6.	Leaf: green colour (light, medium, dark)	3, 5, 7	4
7.	Leaf: leaf blister (absent, moderate, strong, very strong)	1, 3, 5, 7, 9	10
8.	Leaf: Serrature	1, 3, 5, 7, 9	10
9.	Leaf: shape at the transversal cross	1, 3, 4, 5	5
10.	Leaf: wings of the blade	1, 2, 3	7
11.	Leaf: the angle between the lowest lateral veins	1, 2, 3	6
12.	Leaf: height of distal part to the place of the base of the blade (on the level of 2/3 plant height)	3, 5, 7	9
13.	Stem: pubescence of the up-part (upper 5 cm) (absent, light, moderate, strong, very strong)	1, 3, 5, 7, 9	12
15.	Ray flowers: density (non-density, moderate, strong)	3, 5, 7	8
16.	Ray flowers: shape	1, 2, 3, 4	2
17.	Ray flowers: position (flattened, bent in length, wavy, bent in the direction of the head)	1, 2, 3, 4	3
18.	Ray flowers: length (short, medium, long)	3, 5, 7	8
19.	Ray flowers: colour (1 lemon, 2 light yellow, 3 yellow, 4 orange-yellow, 5 orange, 6 purple, 7 red-brown, 8 striped, 9 apricots)	1–8	9
20.	Floret flowers: colour (yellow, orange, purple)	1, 2, 3	1
21.	Floret flowers: Anthocyanin colouration of the stigma (absent, exist)	1, 9	0
22.	Floret flowers: the intensity of the anthocyanin colouration of the stigma (light, moderate, strong)	3, 5, 7	0
23.	Floret flowers: pollen-producing capacity (absent, exist)	1, 9	1
24.	External involucrum: shape	1, 2, 3	3
25.	External involucrum: length of the tip	1, 3, 5, 7, 9	8
26.	External involucrum: Green colour of the external part (light, moderate, dark	3, 5, 7	6
29.	Plant: branching (exclude branching due to the conditions) (absent, exist)	1,9	1
30.	Plant: type of branching (like for the 29)	1,2,3,4,5	4
32.	Head: position	1–9	7
34.	Head: shape from seed side	1–6	5

accounting for 72% of the lines. Five large clusters were identified, four of which coincide with clusters 1, 2, 5, 6, and the third and fourth clusters practically merged into a larger cluster. Separate clusters were established for LD72/1, InK1589, KG16 and InDN47 lines. The origin of the VA1B and VA2B lines is unknown, and a joint

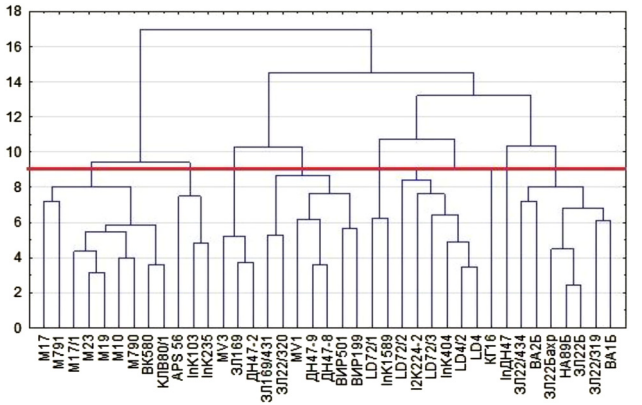

Figure 1: Dendrogram for the morphological characteristics of 39 sunflower lines, expressed on a nine-point scale. Complete connection method. Euclidean state.

movement into a similar cluster is possible. In general, only three lineages in the cluster affinity information available did not correspond to the affinity of cluster placement: InDN47, ZL169/431 and ZL22/320. Two of the offspring and their discrepancy may indicate an insufficient level of backcrossing (only five generations of crosses were conducted in this case, unlike other backcrosses where 6–7 generations were created).

Other displacements were related to the third and fourth groups, which in general had many features in common and were practically separated at the margin of difference.

As to exclude the 'problematic' three lines InDN47, ZL169/431 and ZL22/320 from the experiment, we obtain practically corresponding to the available genetic information the clustering of the lines presented in Figure 2.

Five clusters and the sixth of two lines, InK1589 and KG16, are identified, identifying them with great similarity in characteristics.

The 'problem' three lines InDN47, ZL169/431 and ZL22/320 require verification of affinity with established clusters by direct crossing and molecular identification. But the correspondence of the clusters formed by morphological traits to the genetic affinity of the lines at 72% is a pretty good result. The tested method allows using the developed descriptive material for establishing genetic affinity in sunflower lines.

Many times the use of molecular dendrogram construction is found in scientific research. Only a few studies use plant morphology to establish affinity for sunflowers (Gutierrez et al. 2010; Schwarzbach et al. 2001; Sujatha et al. 2008). But they use wild species and their natural hybrids. The transfer of the method of morphological analysis of affinity with the use of dendrograms to the original breeding lines

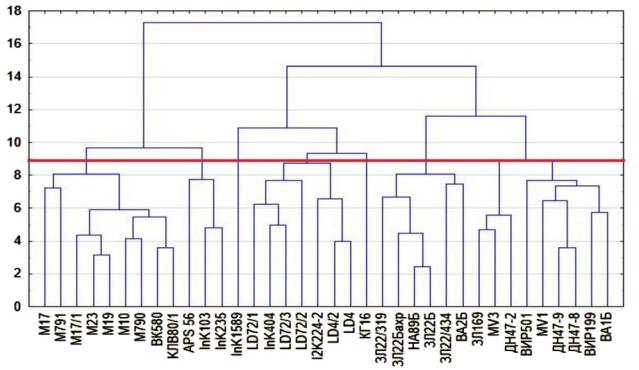

Figure 2: Dendrogram for the morphological characteristics of 36 sunflower lines, expressed on a nine-point scale. Complete connection method. Euclidean state.

is not currently used in scientific studies, but all modern breeding has followed the same path to morphology selection. But before it was the experience of the breeder, his intuition and the current level of statistical calculations allows you to draw appropriate conclusions without having a great experience of the breeder.

Conclusions

A study of the morphological lines of sunflower belonging to the same species divided the lines into groups. The groups corresponded to established lineages.

The possibility of using cluster analysis of morphological traits to establish the genetic affinity of sunflower lines has been proved. In the studying material from 39 lines, the cluster coincidence according to morphological features by 72–92% corresponds to genealogies and genetic identification.

Author contribution: The author has accepted responsibility for the entire content of this submitted manuscript and approved submission.
Research funding: Funding was provided by the National Academy of Agrarian Sciences of Ukraine and the Institute of Oilseeds of NAAS.
Conflict of interest statement: The author declares that they have no conflict of interest.

References

Darvishzadeh, R., Azizi, M., Hatami-Maleki, H., Bernousi, I., Abdollahi Mandoulakani, B, Jafari, M., and Sarrafi, A. (2010). Molecular characterization and similarity relationships among

sunflower (*Helianthus annuus* L.) inbred lines using some mapped simple sequence repeats. Afr. J. Biotechnol. 9: 7280–7288.

Gutierrez, A., Carrera, A., Basualdo, J., Rodriguez, R., Cantamutto, M., and Poverene, M. (2010). Gene flow between cultivated sunflower and *Helianthus petiolaris* (Asteraceae). Euphytica 172: 67–76.

Jan, C.C., Vick, B.A., Miller, J.F., Kahler, L., and Butler, E.T. III. (1998). Construction of a construction of an RFLP linkage map for cultivated sunflower. Theor. Appl. Genet. 96: 15–22.

Kholghi, M., Darvishzadeh, R., Bernousi, I., Pirzad, A, and Laurentin, H. (2012). Assessment of genomic diversity among and within Iranian confectionery sunflower (*Helianthus annuus* L.) populations by using simple sequence repeat markers. Acta Agric. Scand. Sect. B Soil Plant Sci. 62: 488–498.

Liu, J., Mo, J.S., Liu, G.S., Qi, D.M., and Li, F.F. (2001). RAPD analysis of the germplasm resources of sunflower. Acta Bot. Sin. 43: 151–157.

Schwarzbach, A.E., Donovan, L.A., and Rieseberg, L.H. (2001). Transgressive character expression in a hybrid sunflower species. Am. J. Bot. 88: 270–277.

Sivolap, Y.M., and Solodenko, A.E. (1998). Inter- and intraspecies differentiation in the genus Helianthus by RAPD analysis. Helia 21: 19–26.

Sujatha, M., Prabakaran, A.J., Dwivedi, S.L., and Chandra, S. (2008). Cytomorphological and molecular diversity in backcross-derived inbred lines of sunflower (*Helianthus annuus* L.). Genome 51: 282–293.

UPOV TG/81/5. (2000). Guidelines for the conduct of tests for distinctness uniformity and stability original: sunflower (Helianthus annuus L.). English date: 2000-02-18, Available at: https://www.upov.int/edocs/mdocs/upov/en/tg/tg_81_5_proj.pdf.

Vedmed'va, K.V., Solodenko, A.Y., and Tolmachov, V.V. (2010). Genetichna vyyezdnaya missiya v analogii sonyashnika. Vísnik Zaporíz'kogo natsíonal'nogo unívesitetu, pp. 11–17. Available at: http://web.znu.edu.ua/herald/issues/2010/bio_2010_1/011-17.pdf.

Vorob'yev, N.I., Sviridova, O.V., and Kutuzova, R.S. (2006). *Metodicheskiye rekomendatsii po ispol'zovaniyu graf-analiza v issledovaniyakh sistem, sostoyashchikh iz bioticheskikh i abioticheskikh komponentov. SPb.* GNU VNIISKHM, p. 59s, Available at: http://www.arriam.spb.ru/.shared/GIMM/catalogue/Graph2006.pdf.

Whankaew, S., Hasthanasombut, S., Thaikert, R., Tappiban, P., Smith, D.R., and Triwitayakorn, K. (2014). Molecular characterization and genetic relationship of marigolds (Tagetes spp.) based on simple sequence repeat markers. Plant Genet. Resour. Charact. Util. 12: 317–322.

Wood, T.E. and Nakazato, T. (2009). Investigating species boundaries in the Giliopsis group of *Ipomopsis* (Polemoniaceae): strong discordance among molecular and morphological markers. Am. J. Bot. 94: 853–861.

Yang, S.P., Zhong, O.W., Tian, J., Wang, L., Zhao, M., Li, L., and Sun, X. (2018). Characterization and development of EST-SSR markers to study the genetic diversity and populations analysis of Jerusalem artichoke (*Helianthus tuberosus* L.). Genes Genom. 40: 1023–1032.

Zeinalzadeh-Tabrizi, H., Haliloglu, K., Ghaffari, M., and Hosseinpour, A. (2018). Assessment of genetic diversity among sunflower genotypes using microsatellite markers. Mol. Biol. Res. Commun. 7: 143–152.