

## Structure of wild annual sunflower (*Helianthus annuus* L.) accessions based on agro-morphological traits

Hamidreza Nooryazdan<sup>1</sup>, Hervé Serieys<sup>2</sup>, Roberto Bacilieri<sup>2</sup>, Jacques David<sup>2</sup>, André Bervillé<sup>2</sup>

<sup>1</sup>Persian Gulf University, Mahini Street, Bushehr, 7516913798, Iran, e-mail: [hrnooryazdan@pgu.ac.ir](mailto:hrnooryazdan@pgu.ac.ir)

<sup>2</sup>INRA. UMR Diversité et Génomes des Plantes Cultivées. 2 Place P Viala, 34060 Montpellier Cedex 1, France

### ABSTRACT

- The history of Sunflower domestication and especially the selection of high seed oil content explain several bottlenecks that reduced the genetic diversity of the crop compared to the wild relatives. The underlying genetic control of a crop's interaction with its environment is complex and often poorly understood. Studies of geographic and climatic variation can provide important information regarding patterns of adaptation to different habitats and ecological isolation. The objective of the present study was to determine the patterns of distribution of morphological variation for thirteen quantitative characters in 77 accessions of wild sunflower evaluated in the same site.
- The collection of genetic resources of wild *H. annuus*, as held at the French National Institute for Agricultural Research (INRA) in Montpellier, includes approximately 350 populations, which have been collected in the USA. Seventy – seven wild sunflower accessions from this collection were compared in France (Mauguio) for 13 quantitative characters using multivariate methods, including clustering, principal component and canonical discriminant analysis to assess the patterns of morphological and climatic variation. Geographic and climatic data from their sampled sites were analyzed and we sought for correlations of climatic variables and morphological traits. Populations from northern Mexico and southern Canada were not available at that time. Ward's hierarchical clustering (Ward, 1963) was used to identify groups of climatic sites and of morphology forms. Hierarchical clustering was applied by SAS software (SAS, 2000). Principal Component Analysis (PCA) (PRINCOMP procedure, SAS software) was performed on the correlation matrix of the morphologic and climatic variables based on the mean accession value. Canonical discriminant analysis (CDA) on the basis of climatic clustering was used to assess the differences between morphological variations of the 77 accessions for the thirteen characters. These were done using the PROC CANDISC of SAS software (SAS, 2000)
- Principal component analysis (PCA) was used to summarize the characteristic habits of annual wild sunflower sampled sites in terms of geography and climates, thus we characterized the ecogeographical profiles of wild sunflowers. Principal component 1 (PC 1) explains 47% of the dataset variation and shows that temperature and sunniness were closely related but contrasted with number of rain days. PC2 explains 29% of dataset variation and is dominated by effect of mean and total rainfall and contrasted with range of temperatures. PCA was used, also on morphological traits. The first three principal components accounted for 79% of the total variation. PC1 accounted for 35% of the total variance and had factors with high contribution as sowing flowering duration, plant height, leaf length Lamina and leaf petiol length. Hierarchical clustering using Ward' method (Ward, 1963) assembled accessions from the 18 States of USA into four clusters based the 13 morphological traits. The accessions from cluster 1 (North Dakota, South Dakota, Wyoming and Colorado) displayed weak plant height, sowing flowering duration and leaf length Lamina. The accessions from cluster 2 (California, Nevada, Arizona and Utah) displayed more intense apical and median branching than those from other clusters. The accession from cluster 3 (Iowa, Kansas, Oklahoma and Texas) displayed lower seed length (4.6 mm) and thousand seeds weight (8.1 g) than other clusters. The results displayed that the accessions from cluster 4 (Illinois, Missouri and Arkansas) had larger head, longer leaf size, thousand seeds weight and less median and apical branching. The first three canonical variables accounted for 99% of the total morphological variation. Sixty three percent of the variation was explained by canonical coefficient 1 (CAN1) which correlated with sowing flowering duration, plant height, max length of lateral branch and leaf size. CAN2 accounted for 32% of the variation and was correlated with branching trait, max length of lateral branch and seed length. Head diameter, leaf size and basal branching were correlated with CAN3 which explained 4% of the variation in the wild sunflowers. As we studied the traits in a common environment we speculated whether local adaptation patterns are sustained by our data. Indeed traits shared by accessions that belong to climatic clusters are suggested to be due to local adaptation.
- Principal component analysis of quantitative traits associated the components: (1) for precocity related traits in agreement with the geographical origin of populations and (2) to domestication traits. The climatic clustering displayed the local adaptation phenomenon.
- This study combines agro-morphological and climatically data aiming to determine patterns of distribution of morphological variation of 77 wild sunflowers. No such work has ever been reported for wild sunflower.

Key words: **Canonical discriminant analysis – Ecogeographical characterization – Local adaptation**

## INTRODUCTION

With regard to morphological variation, few studies have been done on the wild sunflowers (Schneiter 1998). Studies of geographic and climatic variation can provide important information regarding patterns of adaptation to different habitats and ecological isolation (Zink and Remsen 1987). The wild germplasm of the most cultivated species is likely to display a broad range of tolerance to abiotic stresses. Local adaptation has been defined as (1) genetic changes in a population in response to a geographically localized selection pressure or (2) if local adaptation exists genotypes from local populations will have higher relative fitness than those originating from other habitats, if local adaptation does exist in sunflower, then accessions from the same area should have similar fitness's. We decided therefore, since 1995, to gather in a unique site the genetic diversity of a large sample of wild sunflowers and test the presence of alleles and characters potentially interesting for breeding. This experiment aims to characterize the variation of some agro-morphological traits in wild sunflower populations and at the same time initiates a process of base broadening or neo-domestication that may help breeders to identify genetic factors favorable for the crop. Observation of phenotypes of wild sunflower growing in their natural sites does not enable to distinguish among environmental and genotype effects. The objective of the present study was to determine the patterns of distribution of morphological variation for thirteen quantitative characters in 77 accessions of wild sunflower evaluated in the same site, in Mauguio (France). These accessions were sampled from a wide geographical range and climates in USA. Coordinate of the prospecting sites were used to gather information from public climatic databases that were used to describe roughly the environmental conditions of the sites. An attempt was then made to study the joined variation of the 13 morphological, phenotypical and technological traits revealed in the common location site experiment with the climatic variation of the prospecting sites. We discuss the possibility that the value for some traits may be determined by local adaptation to specific climatic conditions.

## MATERIALS AND METHODS

Seventy-seven wild populations were chosen for further multiplication and testing, as representative of the natural geographic and climatic distribution, in order to represent the largest range of variation in wild sunflower (Serieys et al. 1997). All accessions were planted and observed in the same field near Montpellier at Mauguio (France). The experimental field coordinates are 43 360 360 0 N Lat. and 3 580 480 0 E Long. with mean annual temperature and total annual rainfalls of 13.83 C and 655 mm, respectively (Cramer 2005). For each accession, 15 seeds were sown in a greenhouse on April 1st, 1996. The plantlets were then transferred into the field on 3 m long and 0.75 m wide with plant-to-plant row spacing of 0.5 m. For every accession in a plot, five individuals were used for recording the data, except for sowing-flowering duration, which was recorded as the average of the plot. A total of thirteen characters were recorded at various growth stages using the standard evaluation system for sunflower by Schneiter et al. (1998), as follows: sowing- flowering duration (SOFD) in number of days; plant height (PHE) in cm; basal branching (BASBR); median branching (MEDBR) and apical branching intensities (APIBR); length of branches (LBr), leaf length lamina (LLL); leaf petiole length (LPL); head diameter (HDI); thousand seed weight (TSW); seed length (SL); seed width (SW); and seed oil content (SOC as %). BASBR, MEDBR and APIBR were quantified according to the scale developed by Serieys et al. (2008). The environmental data (latitude, longitude and altitude) were collected for each site of collection from the CLIMATE database version 2.1 (Cramer 2005). We aimed to seek for local adaptation by comparing correlation between morphological and climatic data (Jonas and Geber 1999). Under local adaptation the collect sites should have higher relative fitness than genotypes originating from other habitats. However, the fitness of these accessions is not known in their native location. When some morphological variables are correlated with climatic data whereas other morphological are not correlated we considered that local adaptation has occurred, and was responsible for a significant part of the phenotypic variability observed for such traits. Ward's hierarchical clustering (Ward 1963) was used to identify discrete groups of climatic sites and of morphology forms. Hierarchical clustering was applied by SAS Software (SAS 2000). There are no satisfactory methods to establish the number of clusters (Bock 1985). Traits differentiating climatic clusters were identified using standardized. Principal Component Analysis (PCA) (PRINCOMP procedure—SAS Software) was performed on the correlation matrix of the morphologic and climatic variables based on the mean accession value. The correlation between PCA morphologic and climatic axes was then computed. The results of PCA were confirmed by canonical discriminant analysis (CDA), but CDA was more efficient in identifying the traits common to all accessions for each cluster based on climatic data. From principal components analysis, linear combinations of the original variables that account for as much of the total variation as possible are constructed. From canonical correlation, a linear association between predictor variables and criterion measures are determined (Dillon and Goldstein 1984). CDA on the basis of climatic

clustering was used to assess the differences between morphological variations of the 77 accessions for the thirteen characters. These were done using the PROC CAN-DISC of SAS software (SAS 2000).

## RESULTS

Hierarchical clustering using Ward' method, (Ward 1963) assembled accessions from the 18 States of USA into four clusters based the 13 morphological traits (Table 1). The accessions from cluster 1 (North Dakota, South Dakota, Wyoming and Colorado) displayed weak PHE, SOFD, and LLL. The accessions from cluster 2 (California, Nevada, Arizona and Utah), displayed more intense apical and median branching (clusters 2 and 2.4, respectively) than those from other clusters. The accession from cluster 3 (Iowa, Kansas, Oklahoma and Texas) displayed lower seed length (4.6 mm) and TSW (8.1 g) than other clusters. The results displayed that the accessions from cluster 4 (Illinois, Missouri and Arkansas) had larger head, longer leaf size, TSW, and less median and apical branching. The hierarchical clustering procedure assembled the 77 *H. annuus* sites into four distinct clusters. Cluster 1 was made of 32 accession sites, while clusters 2, 3 and 4 contained 21, 21 and 3 accession sites, respectively. The canonical coefficients showing the contribution of each of the characters used for the canonical analysis to the total variation are presented in Table 2. The first three canonical variables accounted for 99% of the total morphological variation. Sixty- three percent of the variation was explained by Can 1, which correlated with sowing- flowering duration, plant height, Max length of lateral branch and leaf size. Can 2 accounted for 32% of the variation and was correlated with branching trait, Max length of lateral branch and seed length. Head diameter, leaf size and basal branching were correlated with Can 3, which explained 4% of the variation in the wild sunflowers. The canonical discriminant analysis performed on the wild *H. annuus* accessions from the four climatic clusters using the thirteen quantitative characters appeared unequally distributed along the three canonical axes (Fig. 1). The accessions from cluster 1 were negatively correlated with Can 1. Can2 was highly correlated with the accessions from the cluster 2. Can 3 appeared to be more positively correlated with the accessions from cluster 4.

## DISCUSSION

In this study we detected among 77 wild *H. annuus* accessions wide variation on the basis of CV for seed weight, petiole length, plant height, head diameter, max length of lateral branch and sowing- flowering duration, Differences for SOFD between accessions suggested that the development of sunflower maturity groups may be possible. Variation in flowering time has a geographic pattern (Aiken 2005) since significant positive relationships was revealed between Lat. and sowing-flowering duration. Thus, this suggests that SOFD in wild *H. annuus* has local adaptation. Adaptation can be studied by comparing natural accessions in a common environment (Jonas and Geber 1999). SOFD can be shaped by numerous biotic and abiotic factors such as temperature and drought. Variation in SOFD suggests adaptation to the length of the growing season. The important determinants of flowering time, temperature and day length have strong latitudinal gradients. We detected significant difference in SOFD between northern, central, southern and west southern populations (Table 1). Apparently, the wild populations of sunflower unlike cultivated sunflower have sensitivity to photoperiod, but genetic differences between the northern and other populations for genes of the photoperiodic pathway remain to be revealed. Heiser et al. (1969) have claimed that the common wild annual sunflower is the most geographically diverse species of sunflower in North America and exhibits the greatest morphological and habitat variability but there are a few studies about wild *H. annuus* diversity. Heiser (1978) has noted that archeological evidence strongly supports the origin of the domesticated sunflower from the central US. So hybridization between wild and cultivated sunflower in this area is probable. Heiser (1985) has suggested that prior to the colonization of North America by human's wild *H. annuus* was restricted to what is now the southwestern US. The climates in the Western and South-western US locations where these accessions originated tend to be desert-like environments. These regions have the more monthly sunniness and the less monthly rainfall. These accessions are candidate to exhibit drought tolerance. The drought-tolerant nature of sunflower has been attributed to its extensive root system, which can extract water and nutrients to soil in depth of 3 m (Jones and Johnson 1983). According to climatic clustering (Table 2), the Great Plains accessions showed the greater head diameter, le Poids de mille grains and leaf size and lower branching. Rieseberg and Seiler (1990) have analyzed 12 populations of wild *H. annuus* with isozyme markers and found those wild sunflowers with similar geographic provenances were not necessarily genetically allied. Arias and Rieseberg (1995) found the similar results. Cronn et al. (1997) have studied 146 accessions of wild and domesticated sunflower. They found evidence that genetic variation in wild sunflower from different regions of the United States is geographically structured. According to climatic clustering, it was found that the accessions from west and southwest of Unite States morphologically had different forms and exhibited the more branched than

other accessions' zone (Table 2). Branching trait variation could be an adaptation to climatic factors, where plants with higher branching are exposed to higher levels of sunniness and temperature range and therefore water stress, but much less so with plants with less branching in the East. It is evident that branching is an important trait in wild sun- flower (Schneiter 1997).

## REFERENCES

- Aiken, R.M. 2005. Applying thermal time scales to sunflower development. *Agron J* 97:746–754.
- Arias, D.M., and L.H. Rieseberg. 1995. Genetic relationships among domesticated and wild sunflowers. *Econ Bot* 49:239–248.
- Bock HH (1985) On some significance test in cluster analysis. *J Classif* 2:77–108.
- Cramer, W. 2005. Climate database version 2.1. Postdam Institute for Climate Impact Research
- Cronn, R., M. Brothers, K. Klier, P.K. Bretting and J.F. Wendel.1997. Allozyme variation in domesticated annual sunflower and its wild relatives. *Theor Appl Genet* 95:532–545
- Dillon, W.R., and M. Goldstein. 1984. Multivariate analysis: methods and applications. Wiley, New York
- Heiser, C.B. 1978. Taxonomy of *Helianthus* and origin of domesticated sunflower. P. 31–53. In: Carter, J.F. (ed) *Sunflower science and technology*. American Society of Agronomy Madison, Wisconsin, USA.
- Heiser, C.B. 1985. *Of plants and people*. University of Oklahoma Press, 237 p.
- Heiser, C.B., D. Smith, S. Clevenger, and W. Martin. 1969. The North American sunflower (*Helianthus*). *Mem Torrey Bot Club* 22 (3):218.
- Jonas, C.S., and M.A. Geber. 1999. Variation among populations of *Clarkia unguiculata* (Onagraceae) along altitudinal and altitudinal gradients. *Am J Bot* 86(3):333–343.
- Jones, O.R., and W.C. Johnson. 1983. Cropping practices: southern great plains. In: Dregne HE, Willis WO (eds) *Dryland agriculture*. *Agron Monogr* 23 ASACSSA and SSSA, Madison, pp 365–385.
- Rieseberg, L.H., and G. Seiler. 1990. Molecular evidence and the origin and development of the domesticated sunflower (*H. annuus* L.). *Econ Bot* 44:79–91.
- Rieseberg, L.H., A. Desrochers, and S.J. Youn. 1995. Interspecific pollen competition as a reproductive barrier between sympatric species of *Helianthus* (Asteraceae). *Am J Bot* 82:515–519.
- SAS Institute. 2000. *SAS/STAT user's guide*, Release 603 ed SAS Inst, Cary, NC.
- Schneiter, A.A. 1997. *Sunflower technology and production*. Madison, p 834.
- Schneiter, A.A., J.F. Miller, and D.R. Berglund. 1998. *Stages of sunflower development*. North Dakota, North Dakota State University.
- Serieys, H., Quillet M.C. Kaan F., and A. Bervillé.1997. Etude de la variabilité d'*Helianthus annuus* sauvages. Chapitre D4. P. 373–379. *Proceedings 6emes Jour scientifique du réseau biotechnologies végétales* Aupelf Uref Juillet Orsay France,
- Serieys, H., H. Nooryazdan, F. Kaan, R. Bacilieri, J. David, M.F. Ostrowski, M.H. Muller, and A. Bervillé. 2008. *H. annuus* natural populations to increase the whole genetic diversity of domesticated sunflower: the concept of neodomestication, vol 2 P. 736–741. In: *Proc. 17th Int. Sunfl. Conf.*, Cordoba, 10–15 June.
- Ward, J.H. 1963. Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 58:236–244
- Zink, R.M., and V.R. Remsen .1987. Evolutionary processes and patterns of geographic variation in birds. *Curr Ornithol* 4:1–69.

## Tables and Figure

**Table 1** Comparison using Ward hierarchic analysis of morphological traits of the four clusters based on climate data

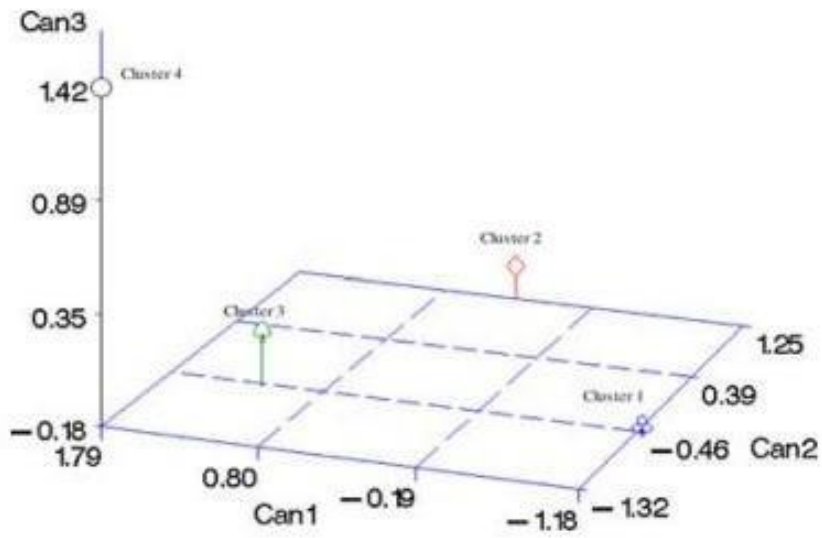
Trait	Unit	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Sowing-flowering duration	Days	60.1 <sup>d</sup>	71.1 <sup>c</sup>	79.8 <sup>b</sup>	86.6 <sup>a</sup>
Plant height	cm	143.6 <sup>b</sup>	208.1 <sup>a</sup>	224.4 <sup>a</sup>	216.4 <sup>a</sup>
Length of branch	cm	97.7 <sup>b</sup>	137.3 <sup>a</sup>	127.6 <sup>a</sup>	114.1 <sup>a,b</sup>
Basal branching	Number	2.5	2.7	2.6	2.6
Median branching	Number	2.3 <sup>a,b</sup>	2.4 <sup>a</sup>	2.1 <sup>b</sup>	1.8 <sup>c</sup>
Apical branching	Number	1.9 <sup>b</sup>	2 <sup>a</sup>	1.7 <sup>b</sup>	1.4 <sup>c</sup>
Leaf length lamina	cm	28.8 <sup>b</sup>	44.5 <sup>b</sup>	49.2 <sup>a,b</sup>	53.3 <sup>a</sup>
Leaf petiole length	cm	11.6 <sup>b</sup>	23 <sup>a</sup>	23.1 <sup>a</sup>	22.9 <sup>a</sup>
Thousand seed weight	g	10.1 <sup>a,b</sup>	10.7 <sup>a,b</sup>	8.1 <sup>c</sup>	13 <sup>a</sup>
Head diameter	mm	27.5 <sup>b</sup>	30 <sup>b</sup>	29.6 <sup>b</sup>	37.9 <sup>a</sup>
Seed length	mm	5.1 <sup>a,b</sup>	5.2 <sup>a</sup>	4.6 <sup>b</sup>	5.3 <sup>a</sup>
Seed width	mm	2.4 <sup>a,b</sup>	2.5 <sup>a,b</sup>	2.3 <sup>b</sup>	2.6 <sup>a</sup>

Means with the same letter in superscripts are not significant difference for each trait at the 5% level

**Table 5** Canonical coefficients (Can 1–3) showing the contribution of each quantitative character to total morphological variation

	Can 1	Can 2	Can 3
Proportion of total variation	0.63	0.32	0.04
Cumulative variation	0.63	0.95	0.99
Correlation coefficient with:			
Sowing-flowering duration	0.84	-0.11	-0.05
Plant height	0.83	0.29	0.15
Length of branch	0.51	0.70	0.07
Basal branching	-0.006	0.46	0.36
Median branching	-0.25	0.53	-0.01
Apical branching	-0.22	0.59	0.02
Leaf length lamina	0.80	0.17	0.27
Leaf petiole length	0.79	0.09	0.27
1,000 seed weight	-0.03	0.13	0.16
Head diameter	0.32	-0.04	0.40
Seed length	-0.15	0.30	0.06
Seed width	0.03	0.19	-0.09
Oil seed content	-0.24	0.11	0.26

Table 2



**Fig. 5** Diagram of the first three canonical variables (CAN) based on the mean values of the traits in climatic cluster of wild sunflowers