Sunflower genetics: From ancestors to modern hybrids

By Aleksandra Radanović, Dragana Miladinović, Sandra Cvejić, Milan Jocković and Siniša Jocić, Institute of Field and Vegetable Crops, Novi Sad, Serbia

Domestication and the first steps of sunflower breeding date back more than 4 000 years. It has undergone significant changes in the past to finally find its place as one of the most significant oil crops today.

Recent advances in molecular techniques with improved experimental designs contributed to the understanding of the genetic and molecular basis underlying the architectural and phenotypic changes that occurred during domestication and improvements in sunflower breeding.

All modern domesticated sunflowers (*Helianthus annuus* L.) can be traced back to a single centre of domestication in the interior mid-latitudes of eastern North America. Today, sunflower is the fourth-most important oil crop in the world.

History and domestication

Archaeological findings show that the Native Americans started the domestication of sunflower in 4225 BC. Since the harvest of each sunflower was a special operation and any variation in the seed size was easy to see, it is logical that the plants with the largest seeds were left for planting in the following season. This was in essence a mass selection for the seed size. The cultivated sunflower as we know it today was most likely created by mass selection from the wild *Helianthus annuus (H. annuus)*, which has small seeds and a branched stem.

The history of sunflower as an oil crop can be divided into three basic periods. The first period is the use of varieties created by mass selection, while the second is the use of varieties created by the method of individual selection, and the third, which is still present, is the introduction of hybrids in the production of sunflower.

The phenotypic changes that sunflower has undergone during domestication largely follow the domestication syndrome. These adaptations to human cultivation include a dramatic increase in apical dominance, an increase in seed size, the loss of natural seed dispersal and seed dormancy, and the loss of self-incompatibility.

Genetic studies of sunflower domestication revealed that, contrary to findings in other plant species where it was found that the phenotypic differences caused by domestication are due to a smaller number of genes with a strong effect, in sunflower there is a larger number of genes involved in domestication, with the majority of genes showing a small or moderate phenotypic effect.

Another difference between wild and cultivated sunflowers is the copy number of long terminal repeats (LTR) retrotransposons and splicing divergence.

Characteristics and composition

Achene size and shattering

Wild *H. annuus* achene is of wide obovate shape, measuring 0,29 to 3,3mm in width, and 0,41 to 6,7mm in length. Its colour is somewhat brown, with two or three dark stripes that are variable in width. However, cultivated sunflower achene is significantly larger, measuring 7 to 25mm in length and 4 to 13mm in width (*Figure 1d*).

When sunflower was domesticated, its seeds increased in size and weight, while easy seed dispersal was disabled. Direct selection for increased seed oil in early oilseed sunflower breeding programmes indirectly selected for smaller achenes and shifted the phenotype toward the wild type. Burke *et al*. associated seven and five quantitative trait loci (QTL) to achene weight and width, respectively, while only two QTLs were associated with achene length.

Chromosomes 3, 6, 9, and 10 carry QTLs for more than one achene morphology parameter, confirming the polygenic nature of these traits. Burke *et al*. and Baack *et al*. discovered two QTLs that have been associated with the domestication QTLs on chromosomes 6 and 10, however, the position of the QTLs was different in the two studies.

Concerning shattering, Burke *et al*. found QTLs on chromosomes 11 and 17 that explained 6,6 and 5% of the phenotypic variation for this trait. Both QTLs expressed a dominant mode of action (MOA) of the cmsHA89 allele. In contrast to this this study, Wills and Burke mapped two QTLs that were associated with shattering on chromosomes 4 and 10.

Oil content

The populations of wild *H. annuus* usually contain below 30% of oil in seed. The first sunflower varieties with increased oil content that were created at sunflower breeding centres in the former Soviet Union had 40 to 45% oil content. Today, most of the sunflower hybrids have 45 to 50% of oil in seed.

Burke *et al*. mapped QTL controlling differences in seed oil content and composition between cultivated and wild sunflower and used the results to guide a genome-wide analysis of genetic variation for evidence of past selection. They identified a QTL for oil content on LG4 and determined the MOA of the cmsHA89 allele as partly recessive.

By analysing the sequence homology with *Arabidopsis* genes, these two putative domestication genes showed the highest homology with AT5G49460 and AT5G52840 genes, which encode ATP citrate lyase sub-unit B2, and have a function related to NADH ubiquinone

oxidoreductase, respectively. The third domestication gene showed homology with a gene that codes a sub-unit of pyruvate kinase, which is an enzyme that was involved in the conversion of carbohydrates to seed oil.

Oil composition

Standard sunflower oil is usually composed of polyunsaturated linoleic acid (18:2) and monosaturated oleic acid (18:1) in ratios of 70%:20%. Although the content of these two fatty acids could vary because of the environment, it is typical for sunflower oil that they jointly make about 90% of the total fatty acid content in the oil.

In cultivated sunflower, there are also traces of monosaturated myristoleic and mitoleic acids, as well as unsaturated myristic, arachidic and behenic acids. An analysis of the changes in the fatty acid content between wild and cultivated sunflower in the progeny of cultivated sunflower cmsHA89 and wild *H. annuus* (ANN1238) showed that only palmitic fatty acid content was similar between the examined cultivated and wild sunflower.

Chapman and Burke discovered that seven out of the eleven genes that underlie fatty acid biosynthesis and metabolism in wild and cultivated sunflower underwent selection (FAD2-1, FAD2-3, FAD3, FAD6, FAD7, FAB1, and FATB). The authors selected sequences that showed orthology with *Arabidopsis* for the study and analysed different desaturase and thioesterase enzymes that were involved in the fatty acid conversion pathway.

Plant architecture changes

Stem properties and height

Sunflower hybrids are typically non-branched annual plants, from 150 to 180cm in height, which are distinguished from other cultivated crops by large conspicuous inflorescence containing many large achenes.

Unlike cultivated sunflower, wild *H. annuus* is characterised by a plant height ranging from 63 to 171cm, highly branched growth form with numerous small flowering heads, and relatively small achenes that are released upon maturation (*Figure 1a*).

Stem diameter QTLs were found on chromosomes 1, 3, 6, 7, 11, and 17 in an F3 cross between cmsHA89 and wild *H. annuus* var. annuus individual that was collected in the United States, while Baack *et al*. found one stem diameter QTL on chromosome 3, and Dechaine *et al*. mapped a QTL associated with stem diameter on chromosome 13 by analysing the recombinant inbred lines (RILs) derived from the same parental material as in Burke *et al.* and testing it in two different environments.

Out of all the identified QTLs, only the QTL that was mapped by Corbi *et al*. on chromosome 13 fell in the similar region as the one mapped by Dechaine *et al*. The other QTLs were mapped in different positions compared to the already identified QTLs, and were located on chromosomes 1, 3, and 9.

Leaf properties

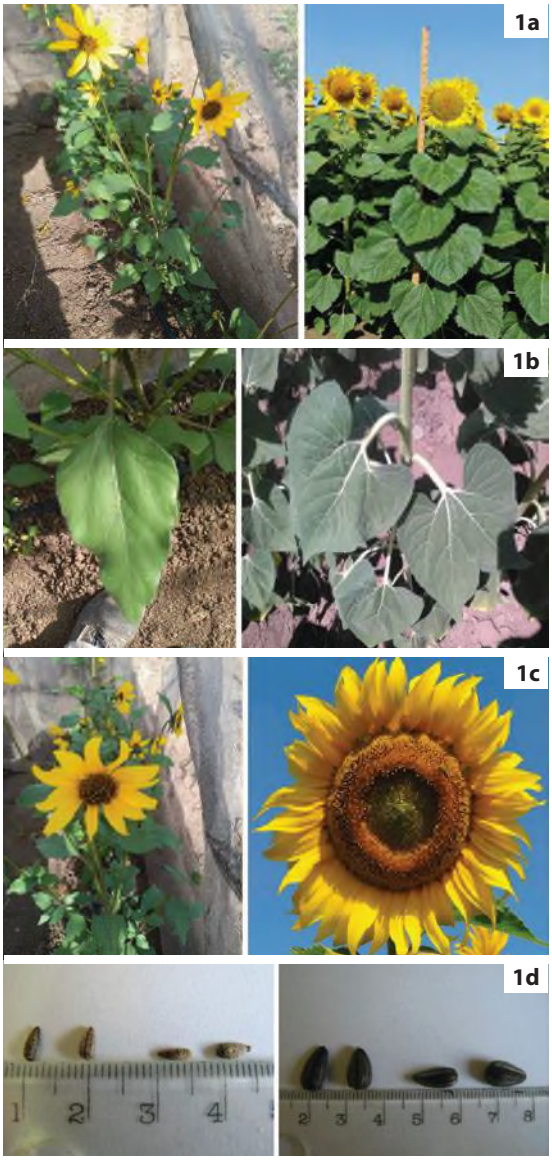
In sunflower, there is significant variability in all the leaf characteristics (*Figure 1b*). There is a considerable difference in the leaf number per plant in connection to the vegetation period, as it is described that early genotypes have a lower leaf number per plant, while genotypes with longer vegetation have a higher number of leaves per plant.

Burke *et al*. reported two QTLs on chromosomes 12 and 13 for leaf shape that expressed dominant and additive MOAs. The leaf size QTLs that were found on chromosomes 3 to 5, and 9 expressed recessive and partially recessive MOAs, while two QTLs expressed over-dominance. Three of the QTLs on chromosomes 1, 9, and 17 for the number of leaves on the main stem expressed a partially dominant MOA of the cultivated allele, while the remaining two on chromosomes 6 and 7 expressed additive and partially recessive MOAs.

Dechaine *et al*. mapped QTLs associated with leaf area while testing RILs in North Dakota and Nebraska in the United States, and found no mutual QTLs for the two locations. Among others, the authors mapped a QTL on chromosome 5 that was present in two locations, however, it was mapped in different positions on chromosome 5.

Unlike the studies mentioned above and in which a cross between cmsHA89 and wild *H. annuus* was used, Wills and Burke used a domesticated Hopi sunflower landrace to cross with wild *H. annuus* and mapped the QTLs for several main stem leaves on chromosomes 6, 7, 9, and 15, and for leaf size on chromosomes 5, 8, 10, and 14 to 16. The QTL for the number of main stem leaves on chromosome 15 explained 57% of the phenotypic variation and was the nearest to the simple sequence repeats (SSR) marker ORS687.

Figure 1: Differences in phenotypes of wild (left) and cultivated *H. annuus* (right) in terms of the (a) plant habitus, (b) leaf, (c) head and (d) seeds.



Head properties

The domestication of sunflower significantly changed head properties. Besides being monocephalic, cultivated sunflower has a significantly larger head diameter than wild sunflower (*Figure 1c*). The head diameter in wild sunflower ranges from 2,4 to 8cm, while the head diameter in cultivated sunflower falls between 20 to 30cm.

Furthermore, there is great variability in the head shape of cultivated sunflower as determined by breeder preference and head inclination. Depending on the region where cultivated sunflower is grown, there is a significant variability of head inclination which is connected to sun damage, damage caused by birds and head rot diseases.

Baack *et al.* reported several QTLs associated with head diameter on chromosomes 4, 6, 19, and 14, of which only the position of one QTL on chromosome 14 overlapped for the two tested environmental conditions (Nebraska and Indiana), which was mapped near marker HT319.

Changes in reproductive strategy

The domestication of sunflower was marked by a loss of self-incompatibility, favouring the pollination of one sunflower plant with the pollen of another and decreased seed dormancy. These traits have been completely or partly lost during domestication and breeding, thus cultivated sunflower is self-compatible and has a short-lived seed dormancy.

Ghandi *et al*. were the first to examine the QTLs for self-incompatibility and self-pollination in sunflower. The authors used a BC1 family obtained from a cross between an inbred line NMS373 (self-pollinated, non-dormant) and wild sunflower ANN1811 (self-incompatible, dormant). They mapped S locus (self-incompatibility locus) as an incomplete dominant allele on the lower end of chromosome 17.

Burke *et al*. found two major QTLs associated with the number of selfed seeds on the lower half of chromosome 17. These QTLs explained 12,7 and 68% of the phenotypic variation, and both were found to be partially recessive.

Life cycle shift

Flowering time is one of the most important domestication traits as it influences the success of the crop. Selection favoured consistent flowering time, however, a late flowering date was favoured in the early stages of domestication in primitive sunflower, while modern cultivated sunflower is characterised by relatively early flowering.

Burke *et al.* reported ten QTLs associated with days to flowering, five of which expressed an additive MOA. Three QTLs found on chromosomes 8 and 17 showed a dominant MOA, while the QTLs on chromosomes 4 and 7 expressed under-dominance and partial recessiveness. Lai *et al*. mapped a locus HT160 on chromosome 8. Based on homology, this locus was predicted to be the APETALA2-like protein and was previously reported as a QTL associated with flowering time and achene size.

Later, Chapman *et al.* mapped five candidate genes on chromosome 7 in the interval where the QTLs for flowering time and the number of main stem leaves were mapped previously. Two out of the five candidate genes, c1921 and c2588, that were mapped by Chapman *et al*. showed homology with the genes that code a DNA-binding with one finger (Dof)-like protein and a protein with the INDETERMINATE domain, respectively.

Dechaine *et al*. enriched a previously reported map by adding the domestication and/or improvement loci identified by Chapman *et al*. to the SSR markers that were used by Baack *et al*. and found QTLs associated with flowering time on certain chromosomes.

Blackman *et al*. conducted a comprehensive study of the different genes that have undergone changes during domestication and improvement. The authors used an integrated candidate approach by analysing the homology with genes of known function and the positions of QTLs associated with flowering times that have already been reported in the literature.

Additionally, they determined that the expression of duplicated homologs of the FLOWERING LOCUS T (FT) in sunflower has a role in sunflower domestication. Four FT-like paralogs have been isolated (HaFT1-4) in the sunflower genome. HaFT1 was under selection in domestication, while the other paralogs were selected during improvement. HaFT1-3 was mapped on chromosome 6 and HaFT1 underlies a major flowering time QTL.

With the availability of new technologies, Mandel *et al*. mapped the QTLs associated with days to flower by use of the Illumina Infinium 10 k SNP array for sunflower and found significant associations for this trait in ten genomic regions located on eight of 17 sunflower chromosomes.

Future prospects and implications

The QTLs reported in this review are associated with major domestication traits. These QTLs can be used as diagnostic markers in tracking introgression from wild into cultivated sunflower and eliminating unwanted sequences surrounding the gene of interest during introgression.

Another application would be to identify the types of crop-like traits that are favoured in the wild if they are subjected to manipulation. Linkage maps obtained from wild and cultivated sunflower crosses can differ from a cultivated sunflower cross due to suppressed recombination, as reported by Wills and Burke, making it difficult to compare QTLs obtained in different crosses.

As QTLs are also highly environmentally dependent, all these QTLs should be further validated in different crosses and by association analysis.

The real breakthrough in sunflower molecular biology was achieved with the publishing of the sunflower genome sequence. Further insight into the domestication process would be achieved by sequencing the wild *H. annuus* genome, preferably through choosing from the population used in most of the QTL analysis of sunflower domestication, such as Ann1238.

This sequence could be further used to define the location of QTLs associated in domestication that have been previously reported to gain more insight into the important metabolic pathways, as was done with the cultivated sunflower, and enable the replacement of SSR markers with more precise SNPs.

This article has been shortened for use in *Oilseeds Focus*. Visit *www.mdpi.com/2073-4425/9/11/528/htm* to access the full report.