

Flowering phenology and hybridization rates between cultivated varieties and weedy populations of sunflower in France

Marie Roumet Marie, Camille Noilhan, Muriel Latreille, Marie-Hélène Muller
UMR AGAP, INRA, 2 place Pierre Viala, 34060 Montpellier, France
Corresponding author: Marie-Helene.Muller@supagro.inra.fr

ABSTRACT

- In France and Spain, weedy populations of sunflowers have been reported in the recent years in sunflower fields. They have been shown to result from the evolution of crop-wild hybrids unintentionally produced during F1-hybrid seed production and introduced through the seed lots. When they are present in high density in sunflower crop fields, they induce yield loss and interfere with the harvest. Weedy and cultivated sunflower belong to the same species, grow in close proximity and their flowering periods overlap partially. Gene flow from the crop to the weedy populations is thus highly probable. In the context of the recent development of herbicide tolerant varieties, it appears important to quantify gene flow, in order to assess and monitor the potential escape of the resistance trait into the wild. The aims of the present study were to:

- quantify the overlap in flowering between the weeds and the cultivated varieties
- estimate the rate of crosses between both forms and relate it to the flowering phenology of the weeds.

- In 2009, in the region of Lauragais (East of Toulouse, France), we surveyed six sunflower fields infested by weedy populations at contrasting levels. In each field, the flowering period of the cultivated variety was recorded. Between 100 and 300 weedy plants per population were followed for their flowering schedule (flowering date, number and area of flowering heads across time). From these data, we estimated the total area of heads that was in bloom over the season, and the proportion of that area that was in bloom during the flowering period of the crop. This proportion was denoted the flowering overlap of the weedy population with the crop. In one of these populations, all surveyed weedy plants, the cultivar, as well as up to 20 offspring per 94 weedy plants were genotyped with 13 microsatellite markers.

- The flowering overlap of the weedy population with the cultivated variety varied between 26 and 57 % among the fields, with highly infested fields showing a lower flowering overlap than the weakly infested ones. Between 13 and 54% of the weeds flowered completely outside the crop flowering period. The rate of pollination of the weeds by the crop was estimated on average to 11%. This rate varied widely across the season, with the early flowering heads being crossed at a rate reaching 35%.

- Our results showed that gene flow can be substantial between the crop and the sympatric weeds, although a number of factors can influence this rate such as the level of infestation and the flowering phenology of the weeds. The outcome of these crosses, namely how the resulting seeds will contribute to the evolution of the weedy population, will depend partly on their fitness relative to that of weed x weed offspring.

- This is the first quantification of crop-to-weed gene flow in sunflower at the population level, and integrating the effect of phenology. Our results can help to assess the durability of tolerant varieties, to design recommendations for their use, as well as to predict the fate of the weedy populations submitted to recurrent gene flow.

Keywords: gene flow - weedy sunflower - genetic diversity - flowering phenology - hybridization

INTRODUCTION

Despite the fact that *Helianthus annuus* originates from North America, weedy sunflowers showing wild phenotypic traits have been reported in and around cultivated sunflower fields in Europe and cause agricultural losses (Faure et al., 2002). For France and Spain, molecular data showed that they resulted from the evolution of crop-wild hybrids unintentionally produced during commercial seed production, and introduced into cultivated fields through the seed lots (Muller et al., 2009, 2011). Weedy and cultivated sunflower belong to the same species, are pollinated by bees, grow in the same place, and their flowering period overlap partially (Muller et al., 2009). Gene flow through pollen from the cultivated varieties into these weedy populations is then highly probable, and could deeply impact their evolution, especially when herbicide-tolerant varieties are cultivated (Massinga et al., 2003).

Many studies have addressed the issue of gene flow from cultivated to wild sunflower: range-wide studies of the co-occurrence of wild populations and cultivated fields (Burke et al., 2002), quantification of crop-wild hybridization rates in progenies (Arias and Rieseberg, 1994), analysis of the fitness of crop-wild hybrids in natural environments (Snow et al., 1998). All showed that gene flow is possible and can be substantial. In this study, we considered the case of gene flow from a cultivated field into a weedy population growing in that field. Our objectives were: (i) to quantify the potential of pollen flow through the description of the flowering overlap between the crop and the weedy population, and (ii) to estimate crop-weed hybridization rates at the population level, and as a function of the weeds' flowering time, using microsatellite markers.

MATERIALS AND METHODS

(i) Experimental design: Six fields sown with cultivated varieties of sunflower have been surveyed in 2009 in the Lauragais region (close to Toulouse, France). These fields were infested by weedy sunflower populations; the infestation level was classified as low ([L], a few weeds or clusters of weeds) or high ([H], large zones of the field with high densities of weeds). In May, when weeds and crop plants were at the 2-6 leaves stage, quadrats were defined, and all weeds growing within the quadrats were marked with a numerated ring. The number and size of the quadrats were adjusted as a function of the weeds' density (Table 1). In population Escalquens, leaf samples were taken for DNA analysis from all marked weeds and from the cultivated variety.

Table 1. Characteristics of the surveyed fields and results of the phenological survey.

Field [Infestation level]	Latitude Longitude	Sowing Harvest dates	Sample size (quadrats)	Crop flowering period ¹	Weeds' flowering period ²	Ov _F (%) ³	Isolated plants (%) ⁴
Odars [L]	01°34'20''E 43°32'14''N	5/12 9/8	92 (6)	[7/16, 8/3]	[-11,54]	57	16.3 ^{ab}
Verfeil [L]	01°40'14''E 43°38'52''N	5/8 9/2	109 (8)	[7/10, 8/1]	[-1,54]	42	26.6 ^{abc}
Villefranche [L]	01°42'28''E 43°24'40''N	5/7 9/2	136 (8)	[7/11, 8/2]	[-5,53]	51	13.2 ^a
Balma [H]	01°33'25''E 43°36'41''N	5/20 9/16	206 (14)	[7/25, 8/11]	[-16,53]	26	42.7 ^c
Escalquens [H]	01°34'16''E 43°31'50''N	5/10 9/8	258 (14)	[7/11, 8/3]	[-5,59]	39	36.4 ^{bc}
Fourquevaux [H]	01°37'28''E 43°31'00''N	5/6 9/9	223 (14)	[7/16, 8/2]	[-10,55]	34	54.2 ^d

1. [Tc₁, Tc₂], as defined in the text. 2. [Flowering date of the first flowering weed – End of flowering of the last flowering weed]. Dates are given relatively to the Tc₁ value of the field. 3. Flowering overlap of the weedy population with the cultivated variety (see text for details). 4. Frequency of isolated plants. Values with the same letter are not significantly different (P<0.05).

(ii) The fields were visited twice a week from the end of June to the first week of August, and once a week from the second week of August until the harvest. At each visit, the number of flowering heads (primary, secondary and tertiary) was recorded for each marked weed, as well as the flowering stage of the cultivated field. Additionally, for each weed, we recorded the flowering time (T_0), and the sexual morph (Hermaphrodite or male-sterile). We measured the diameter of the primary flowering head at the

end of its flowering. We estimated the average sizes of the secondary and tertiary heads by measuring a representative sample at the end of their flowering.

In population Escalquens, we bagged the primary heads to recover the seeds for DNA analysis. Additionally, for hermaphrodite branched plants, we bagged a secondary head before its flowering: this bag was harvested at the end of the season; the plant was classified as self-compatible if more than 4 seeds had been produced, and self-incompatible if no seeds had been produced. The self-compatibility was set to unknown otherwise.

From our observations, we estimated the flowering period of the cultivated field $[T_{c1}, T_{c2}]$, where T_{c1} was the date at which around 5% of the cultivated plants had started flowering, and T_{c2} the date at which more than 95% had finished flowering. T_{cult} was defined as the peak of flowering of the crop, when around 50% of the plants were in bloom. From the number of flowering heads and from their sizes, we estimated for each weed i the flowering area across time. We computed the individual flowering overlap (Ov_i) as the proportion of this area that flowered during the crop flowering period. The same computation was made at the field level (Ov_F): namely, for each day of visit, we summed the individual flowering areas to get the population flowering area; then we computed the proportion of the total population flowering area in bloom during the crop flowering period. We defined the frequency of isolated plants in field F as the frequency of weeds for which $Ov_i = 0$. The effect of the field infestation level on Ov_F , was tested using a one-way analysis of variance (ANOVA, performed in R-2.5.1 package lmtest). The significance of the differences of the frequencies of isolated plants between populations was tested using a chi-square test.

(iii) In population Escalquens, DNA was isolated from all surveyed weeds and from the cultivar. We randomly selected 94 among the weeds which produced a sufficient number of seeds: around 20 seeds produced by their primary head were randomly chosen and germinated. DNA was isolated from leaves of the seedlings. Molecular methods are described in Muller et al. (2011). The sample was genotyped with 13 microsatellite loci: ORS297, ORS337, ORS342, ORS344, ORS371, ORS380, ORS432, ORS610, ORS656, ORS674, ORS788, ORS887 and ORS925 (Tang et al., 2002).

For each offspring, given its genotype and the genotype of its mother, we used a procedure of exclusion (relying on the rules of Mendelian inheritance, Jones et al., 2010) to determine (i) if the variety cultivated in the field could be its father and (ii) if the offspring could result from self-fertilization. Then we computed familial weed-crop hybridization rates as the proportion of the outcrossed offspring of a mother plant that could have the cultivar as a father. The theoretic exclusion probability over these loci was estimated as 0.9995, using the program CERVUS (Kalinowski et al., 2007).

To take a better advantage of the information present in the data set, a likelihood method was adapted from the mating-model developed by Oddou-Muratorio et al. (2005): namely, we wrote the probability of the genotype g_o of each offspring o produced by a mother j with genotype g_j , given three possible crosses: selfing (which occurred with probability s_j), outcrossing by the cultivar (with probability $(1 - s_j) \pi_{j,cult}$), and outcrossing by another weed (with probability $(1 - s_j)(1 - \pi_{j,cult})$):

$$P(g_o | g_j) = s_j \times T(g_o | g_j, g_j) + (1 - s_j) \times \pi_{j,cult} \times T(g_o | g_j, g_{cult}) + (1 - s_j) \times (1 - \pi_{j,cult}) \times T(g_o | g_j, AF)$$

$T(g_o | g_j, X)$ is the Mendelian segregation probability of the offspring genotype given the genotype of the mother and given X , the genotype of the father. $X = g_j$ in case of selfing, $X = g_{cult}$, the genotype of the variety in case of outcrossing by the variety, and $X = AF$, the allele frequencies in the pollen cloud produced by the weeds, in case of outcrossing with another weed. These AF were computed from the genotypes of the whole data set of mother plants in population Escalquens (i.e. 246 plants).

The propensity of mother j to self, s_j , was written as a function of the sexual morph H_j , the self-compatibility status C_j of mother j , and s_{hc} the selfing rate of hermaphrodite, self-compatible plants: $s_j = H_j C_j s_{hc}$. Namely H_j was equal to 0 or 1, if mother j was respectively male sterile or hermaphrodite. C_j was equal to ε or $1 - \varepsilon$, if the mother plant had been respectively classified as self-incompatible or self-compatible. An error rate of $\varepsilon = 0.1$ was introduced to take into account the uncertainty surrounding our inference of self-compatibility. For unknown status, H_j and C_j were respectively equal to the frequencies of hermaphrodites (0.92) and of self-compatible hermaphrodites (0.166) in the whole population.

The expected proportion of crop-weed hybrids in the outcrossed progeny of mother j , $\pi_{j,cult}$, was assumed to be equal to the proportion in the pollen cloud of pollen produced by the cultivated variety (denoted as *cultivated pollen* in the following). It was written as a function of three components: $f_{w/c}$, the total production of pollen of the weedy population (denoted as *weedy pollen* in the following) relative to that of the cultivated field, $P_{cult}(T_{0j})$ and $P_{weed}(T_{0j})$, respectively the probability for a cultivated pollen and for a weedy pollen to be present in the pollen cloud at T_{0j} , the flowering date of mother j .

To compute $P_{cult}(T_{0j})$, we considered that all cultivated pollen was produced at date T_{cult} , and that mother j was at the temporal distance $|T_{0j} - T_{cult}|$ of the cultivated field. We then used a family of functions traditionally used to model spatial pollen dispersal (see Oddou-Muratorio et al., 2005), to express the probability for a pollen grain produced at date T_{cult} to be present in the pollen cloud at date T_{0j} :

$$P_{cult}(T_{0j}) = \frac{\frac{b}{2\pi a^2 \Gamma(2/b)} \exp\left[-\left(\frac{|T_{cult} - T_{0j}|}{a}\right)^b\right]}{\int_{T_{0min}}^{T_{0max}} \frac{b}{2\pi a^2 \Gamma(2/b)} \exp\left[-\left(\frac{|T_{cult} - t|}{a}\right)^b\right] dt}$$

Where Γ is the gamma function; a is a scale parameter and b is a shape parameter affecting the tail of the distribution. This function reduces to the bivariate normal distribution for $b=2$, and to the bivariate exponential distribution for $b=1$. The denominator allows defining this function in the time interval containing the flowering dates of the weeds (from $T_{0min} = 10^{\text{th}}$ of July to $T_{0max} = 22^{\text{th}}$ of August).

To compute $P_{weed}(T_{0j})$, we assumed that the production of weedy pollen was homogenous across the season. Then, $P_{weed}(T_{0j}) = 1/(T_{0min} - T_{0max})$, a uniform distribution defined in the interval $[T_{0min} - T_{0max}]$.

Then $\pi_{j,cult}$ can be written as

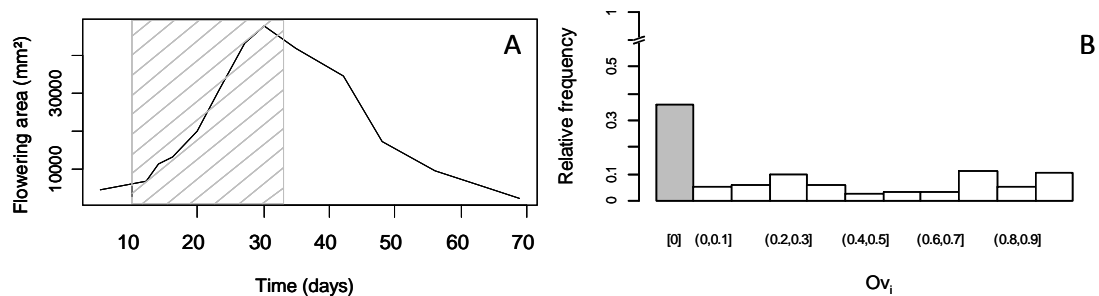
$$\pi_{j,cult}(f_{w/c}, T_{0j}) = \frac{P_{cult}(T_{0j})}{P_{cult}(T_{0j}) + f_{w/c} P_{weed}(T_{0j})}$$

The likelihood of the full data set was computed as the product of the probabilities of each offspring. The parameter values were estimated by maximum-likelihood using a quasi Newton algorithm. Confidence intervals were computed using 1000 bootstrap of mothers. Different functions were compared via likelihood ratio tests. Computations were performed using R (R Development Core Team 2008).

RESULTS

In the six studied fields, the weedy population flowered longer than the cultivated variety (Table 1): on average, the flowering duration of crop and weedy populations extended respectively over 19.8 and 62.6 days. In all fields, the first weeds started flowering before the crop, and some weeds were still in bloom at harvest. The flowering periods of the weedy populations and of the cultivated fields were largely overlapping: overall, around 40% of the total weed flowering area was blooming during the crop flowering period (Table 1 and Figure 1A). As shown in Figure 1B, the overlap in flowering periods varied between individuals of the same population: each weedy population was composed of plants which entirely or partially flowered during $[T_{c1}, T_{c2}]$ and of plants completely isolated from the crop. The vast majority (99.7 %) of the isolated plants flowered after T_{c1} .

Figure 1 A. Flowering periods of the cultivated variety and of the weedy sunflower population in Escalquens. The curve represents the variation of the weedy population flowering area with time (date 0 corresponds to the 1st of July 2009); the dashed zone corresponds to the crop flowering period. **B.** Distribution of the individual flowering overlaps in Escalquens; the shaded bar corresponds to individuals for which no flowering occurred during the flowering period of the crop (i.e. isolated individuals).

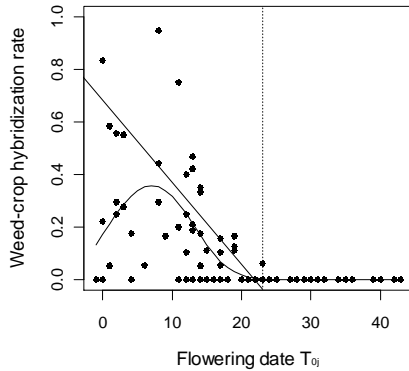


The overlap in flowering period of the weedy populations with the crop varied with the field infestation level: it was around 50% in weakly infested fields and only 33% in highly infested ones ($P =$

0.042, Table 1). This result reflected differences for the frequency of isolated plants, which was significantly higher in highly infested fields (43.9% vs. 18.4%, $\chi^2 = 62.18$, $P = 3.1 \cdot 10^{-15}$).

The familial weed-crop hybridization rates estimated with the exclusion procedure were highly variable during the crop flowering period and then decreased to zero (Figure 2). On average, the proportion of crop-weeds hybrids among the offspring reached 11.2%.

Figure 2. Weed-crop hybridization rates as a function of the flowering date of the mother plant. Filled circles: Familial weed-crop hybridization rates estimated with the exclusion procedure. Continuous line: variation of $\pi_{j,cult}$, the expected proportion of weed-crop hybrids among outcrossed offspring, as estimated by maximum-likelihood. The dotted vertical line is the end of the crop flowering period.



Likelihood-ratio tests identified that the function that best accounts for temporal dispersion of the pollen produced by the cultivated field was the normal function (Figure 2). Estimated parameters values and their confidence intervals were $a=7.56$ [7.36-8.05], $f_{w/c} = 6.92$ [6.92-9.99] and $s_{hc}=0.0996$ [0.0679-0.104].

The results of the mating-model were in accordance with the simple exclusion procedure, but allowed predicting the contribution of flowering time to hybridization rate variation. Hybridization rates were predicted to reach a value of 35% at the maximum of the crop flowering period.

DISCUSSION

Overall, our results showed that crop-to-weed gene flow is possible: flowering overlap was substantial even if variable across populations. Additionally, we showed that plants were not similarly exposed to gene flow from the crop: some plants can ‘escape’ from cultivated pollen by flowering outside of the crop flowering period.

Using two different procedures, we estimated weed-crop hybridization rates as high as 95% for individual progenies, and on average of 35% at the peak of crop flowering. Arias and Rieseberg (1994) and Massinga et al. (2003) estimated values of the same order of magnitude between cultivated plants and sympatric wild plants, but did not take into account the flowering period. During the flowering of the crop, around one weed was flowering per 10 cultivated plants: the rate of 35% for weed pollination by the crop then appears relatively low. The pollinators’ behaviour could induce assortative mating between weeds instead of random pollen dispersal between crop and weeds (Pham-Delegue et al., 1989). Indeed, the high variance among familial weed-crop hybridization rates (Figure 2) probably reflected the correlation of dispersal events by bees and the effect of other factors such as the local density of weeds and plant height.

The outcome of these crosses will depend on the relative fitness of the resulting hybrids. However, for the release of tolerant varieties, this hybridization rate can anyway be considered as rather high, especially since no strong fitness effects of the IMI-resistance gene has been shown (Massinga et al., 2005). An herbicide resistance transgene, present in cultivated variety of *Brassica napus* has been shown to persist in populations of the weedy relative *Brassica rapa*, even in absence of herbicide pressure (Warwick et al., 2008). Moreover even with low crossing rates, gene flow is expected to occur when high selective pressures are favouring a given gene (Haygood et al., 2003). If no strong measures are taken to eliminate all flowering weeds in the vicinity of herbicide tolerant cultivated fields, the tolerance trait could readily move into the weedy populations.

ACKNOWLEDGMENTS

Many thanks are due to Muriel Tavaud and the CETIOM experimental station of EnCrambade for assistance during the field season, to Adeline Cayre and Aline Rocher for their help during the DNA isolation stage. This work was funded by the CETIOM.

REFERENCES

- Arias, D.M. & Rieseberg, L.H. 1994. Gene flow between cultivated and wild sunflowers. *Theor Appl Genet* 89: 655-660.
- Burke, J.M., Gardner, K. & Rieseberg, L.H. 2002. The potential for gene flow between cultivated and wild sunflower (*Helianthus annuus*) in the United States. *Am J Bot* 89: 1550-1552.
- Faure, N., Serieys, H. & Bervillé, A. 2002. Potential gene flow from cultivated sunflower to volunteer, wild *Helianthus* species in Europe. *Agric Ecosyst Environ* 89: 183-190.
- Haygood, R., Ives, A.R. & Andow, D.A. 2003. Consequences of recurrent gene flow from crops to wild relatives. *Proc Roy Soc Lond* 270: 1879-1886.
- Jones, A.G., Small, C.M., Paczolt, K.A. & Ratterman, N.L. 2010. A practical guide to methods of parentage analysis. *Mol Ecol Resour* 10: 6-30.
- Kalinowski, S.T., Taper, M.L. & Marshall, T.C. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol* 13: 1099-1106.
- Massinga, R.A., Al-Khatib, K., St. Amand, P. & Miller, J.F. 2003. Gene flow from imidazolinone-resistant domesticated sunflower to wild relatives. *Weed Sci* 51: 854-862.
- Massinga, R.A., Al-Khatib, K., St. Amand, P. & Miller, J.F. 2005. Relative fitness of imazamox-resistant common sunflower and prairie sunflower. *Weed Sci* 53: 166-174.
- Muller, M.-H., Delieux, F., Fernandez-Martinez, J.M., Garric, B., Lecomte, V., Anglade, G., Leflon, M., Motard, C. & Segura, R. 2009. Occurrence, distribution and distinctive morphological traits of weedy *Helianthus annuus* L. populations in Spain and France. *Genet Resour Crop Evol* 56: 869-877.
- Muller, M.-H., Latreille, M. & Tollon, C. 2011. The origin and evolution of a recent agricultural weed: population genetic diversity of weedy populations of sunflower (*Helianthus annuus* L.) in Spain and France. *Evol Appl* 4: 499-514.
- Oddou-Muratorio, S., Klein, E.K. & Austerlitz, F. 2005. Pollen flow in the wildservice tree, *Sorbus torminalis* (L.) Crantz. II. Pollen dispersal and heterogeneity in mating success inferred from parent-offspring analysis. *Mol Ecol* 14: 4441-4452.
- Pham-Delegue, M.H., Etievant, P., Guichard, E. & Masson, C. 1989. Sunflower volatiles involved in honeybee discrimination among genotypes and flowering stages. *J Chem Ecol* 15: 329-343.
- Snow, A.A., Moran-Palma, P., Rieseberg, L.H., Wszelaki, A. & Seiler, G.J. 1998. Fecundity, phenology, and seed dormancy of F₁ wild-crop hybrids in sunflower (*Helianthus annuus*, Asteraceae). *Am J Bot* 85: 794-801.
- Tang, S., Yu, J.-K., Slabaugh, M.B., Shintani, D.K. & Knapp, S.J. 2002. Simple sequence repeat map of the sunflower genome. *Theor Appl Genet* 105: 1124-1136.
- Warwick, S.I., Légère, A., Simard, M.-J. & James, T. 2008. Do escaped transgenes persist in nature? The case of an herbicide resistance transgene in a weedy *Brassica rapa* population. *Mol Ecol* 17: 1387-1395.
- Whitton, J., Wolf, D.E., Arias, D.M., Snow, A.A. & Rieseberg, L.H. 1997. The persistence of cultivar alleles in wild populations of sunflowers five generations after hybridization. *Theor Appl Genet* 95: 33-40.