

# Does introgression of crop alleles into wild and weedy living populations create cryptic in situ germplasm banks?

Norman C. Ellstrand 

Department of Botany and Plant Sciences,  
University of California, Riverside, CA, USA

## Correspondence

Norman C. Ellstrand, Department of Botany and Plant Sciences, University of California, Riverside, CA, USA.  
Email: norman.ellstrand@ucr.edu

## Funding information

USDA NIFA Hatch, Grant/Award Number: 81361

Natural gene flow often delivers alleles from cultivated species into unmanaged populations of weedy or wild relatives. Unmanaged populations with introgressed crop alleles may become the unintended repositories of crop alleles that are no longer available to breeders. Descriptive studies have confirmed the introgression of alleles from 17 crop species into free-living populations. Multigeneration experimental studies allowing hybrid-derived individuals to evolve under natural conditions are informative, but remain few. A study in this issue of *Molecular Ecology* (*Molecular Ecology*, 26, 2017) documents genomewide crop allele evolution in experimental populations of crop-wild hybrid-derived sunflowers. Approximately, three-quarters of the wild alleles increased in frequency;—however, 5% of the crop alleles in each of the two sites increased (*Molecular Ecology*, 26, 2017; Figure 1). The remainder behaved as if they were neutral. This study's results plus those of prior studies demonstrate that introgressed populations of crop-wild relatives contain a mixture of wild and crop alleles and thereby can be valued as a distinct kind of in situ germplasm resource relative to nonintrogressed populations.

## KEYWORDS

crop-wild relatives, gene flow, in situ germplasm conservation, introgression

Spontaneous introgression from cultivated species into unmanaged populations of weedy or wild relatives can have interesting consequences (Ellstrand, 2003a). Certain hybrid-derived weeds have become more difficult to control, in one case disrupting a regional economy (Merotto et al., 2016). Crop gene flow has played a role in the extinction of at least one wild taxon (Kiang, Antonovics, & Wu, 1979). Transgene flow from cultivated species into free-living populations of *Agrostis*, *Brassica* and *Gossypium* has generated attention from the popular media (Ellstrand, 2012).

Crop-wild trysts may have beneficial effects. One liaison led to bread wheat's evolution (Feldman, Lupton, & Miller, 1995). Also unmanaged populations that have received crop alleles via recurrent gene flow for decades (e.g., Linder, Taha, Seiler, Snow, & Rieseberg, 1998) may become unintended repositories of crop alleles that are unavailable in ex situ germplasm collections. Those populations could be mined for useful historic alleles ("heirloom alleles") for crop improvement (Ellstrand, 2003b). Their utility depends on which crop alleles persist after immigration, how many persist and whether they are useful.

Population genetic theory has straightforward general predictions for the evolutionary impact of recurrent, unilateral gene flow from a source to a sink (Table 4.2, Ellstrand, 2003a). If immigrant crop alleles are neutral in unmanaged populations, their frequency in the sink should evolve to match that of the source. If detrimental, they are maintained in selection–migration equilibrium. Advantageous alleles should eventually come to fixation as gene flow and selection work together. Close linkage with alleles at other genes under strong positive or negative selection influences the fate of neutral or weakly selected immigrant alleles accordingly.

Descriptive studies have confirmed the introgression of alleles from at least 17 crop species into free-living populations (Ellstrand et al., 2013). Most of these studies are based on a handful of marker loci; thus, whether those genes are representative of the genome is unclear (but see Hufford et al., 2013). Likewise, without considerable genomic information and analysis, a descriptive study cannot determine the processes by which introgressed alleles have persisted.



**FIGURE 1** A lone wild sunflower in the foreground. A field of thousands of cultivated sunflowers in the background

Multigeneration experiments that allow hybrid-derived individuals to evolve under natural conditions can measure changes in crop-specific allele frequencies over time. Surprisingly, such studies remain exceedingly few. Snow et al. (2010) monitored crop allele evolution in four natural, self-seeding, annual populations initiated with hybrids between cultivated radish, *Raphanus sativus*, and the wild species *R. raphanistrum*. Despite variation between the three loci monitored and locus-by-population variation, two of the three crop-specific alleles showed a more-or-less consistent decline. Nonetheless, all three alleles remained present after 10 years at all four experimental sites.

Hooftman et al. (2011) were the first to take a genomewide experimental approach. They started by hybridizing the highly selfing annual weed *Lactuca serriola* with a lettuce (*L. sativa*) cultivar. After a bout of self-fertilization, the resulting seedlings were backcrossed to the weed. The resulting lineage was grown at three locations simulating the typical habitat of *L. serriola*. Plants were allowed to self through two generations in the field. The frequencies of more than 200 molecular markers were monitored. The most striking result was the evolution of increased linkage disequilibrium involving blocks of wild alleles and blocks of cultivar-specific alleles in all locations. Neither crop nor wild alleles were significantly favoured.

Corbi, Baack, Dechaine, Seiler, & Burke (2017) take the next step, studying crop allele evolution in a highly outcrossing species. They started with 169 recombinant inbred lines (RILs) derived from a single hybrid individual from a cross between a wild sunflower and an oilseed cultivar (both *Helianthus annuus*). The development of the RILs by 6–8 generations of self-pollination after the  $F_3$  generation reduced initial linkage disequilibrium present in the original hybrid. The experiment involved two generations of open pollination in two locations that were ecologically and geographically disparate but simulated wild sunflower habitat. They measured allele frequency changes for a genomewide collection of hundreds of SNPs.

The results are approximately the same for both sites and striking for so few generations (Corbi et al., 2017; Figure 1). Only about 8% of the crop alleles significantly increased in frequency in at least one location (about 5% for each site). About 20% of the SNP loci exhibited changes consistent with drift. The wild allele increased

frequency for about three-fourths of the loci. An accompanying common garden study showed the evolution of an increasingly wild phenotype in the experimental plants.

What do Corbi et al. (2017) and previous studies tell us about crop allele introgression into wild genomes under natural conditions, especially with regard to enhancing their germplasm value? Overall, alleles associated with domestication traits are strongly disadvantageous. Descriptive studies have largely shown the same trend (e.g., Hufford et al., 2013), and, in the few cases where such alleles are known to be maintained in free-living populations, their detrimental phenotypes are ameliorated by the evolution of de-domestication phenotypes at other loci (e.g., Xia, Xia, Ellstrand, Yang, & Lu, 2011). Generally, we expect deleterious alleles to go extinct in the wild or be maintained at low frequency in selection–migration equilibrium. The absence of domestication alleles in an introgressed wild population is hardly a germplasm problem because they are traits that breeders would never need for crop improvement.

Crop-specific alleles not associated with domestication traits or enhancing domestication traits are less likely to be strongly disadvantageous under natural conditions. These remaining alleles could range from disadvantageous to neutral to highly advantageous. But crop alleles that significantly increase in frequency in experimental sites mimicking the typical wild/weedy habitat are paradoxical. Why are they present in a crop cultivar but absent from the modern version of the crop's wild progenitor?

Corbi et al. (2017) venture a few guesses. The first is that the contemporary cultivated sunflower has an improvement history that involves introgression from *Helianthus* species other than its progenitor. They might hold alleles not present in the progenitor. The second is that the hybrids' wild parent was native to Nebraska and distant from the experimental sites (North Dakota and Iowa). Certain Nebraska alleles might be maladaptive in the other locations. Their final hypothesis is that “some increases in the crop allele are due to beneficial mutations that occurred following sunflower domestication” (Corbi et al., 2017). Note, “following sunflower domestication” could be as recent as a decade ago or as long as thousands of years ago. All the suggestions are reasonable; each could apply separately or in combination with the many favoured crop alleles discovered in the study. The first and last identify avenues for the independent evolution of alleles in a crop that would do well in, but are absent from, wild populations.

The study at hand reveals that a single individual of a single sunflower cultivar has alleles at hundreds of loci that are advantageous in the wild and, following spontaneous hybridization, should increase in frequency. Three times as many crop-specific alleles appear to be more-or-less neutral. If neutral alleles enter the wild population via a single bout of gene flow, their chance of fixation is equal to their initial frequency in the recipient population, but under long-term recurrent unilateral gene flow, they should evolve to match their frequency in the source population. For wild sunflower in North America, gene flow from the domesticated to wild populations is recurrent, but intermittent (“occasional” . . . or “frequent . . . but not constant”, J. M. Burke, *personal communication*, July 16, 2017). Gene flow from multiple individuals of a single cultivar should provide more advantageous alleles and a succession of cultivars over decades even more so.

If Corbi et al. (2017) and other genomewide studies are representative, given some genetic differentiation between a crop and a wild relative, the occasional receipt of crop alleles via gene flow should create introgressed populations that are repositories of a sizeable number and variety of crop alleles. The North American industrial oilseed sunflower production began to bloom almost 100 years ago (J. M. Burke, *personal communication*, July 16, 2017). If some of the gene-flow events occurred 100 or even 50 years ago, introgressed wild populations should hold alleles absent in contemporary cultivars, lost via selection and/or drift as those cultivars evolved. Furthermore, because the cultivated range of sunflower has shifted over time, some introgressed wild populations that hold older crop alleles do so without the threat of dilution of “modern” alleles emanating out of contemporary plantations (J. M. Burke, *personal communication*).

Wild and weedy populations adjacent to cultivated relatives typically have higher levels of genetic diversity than those that are distant (Table 9.1, Ellstrand, 2003a). Gene flow from the crop accounts for that difference. While current immigration might act to dilute the contributions of past immigration, long-term weed soil seed banks buffer the influences of contemporary evolutionary events in the standing population, particularly for annuals (e.g., Templeton & Levin, 1979).

Wild populations introgressed with historic, “lost” crop alleles are a different and perhaps more valuable *in situ* germplasm source for crop improvement compared to those without introgressed alleles. Here are a few possible reasons: the crop alleles present are robust, having persisted in a more rugged environment than under cultivation (That may not hold for weeds that must grow under cultivated conditions, for example, certain weedy rice populations). Crop-introgressed populations should be easier to cross. Finally, progress during backcrossing should be easier because the introgressed genome already has a substantial fraction of the crop genome (for Corbi et al. (2017) about 20%). At the very least, wild or weedy populations introgressed with crop alleles offer a different kind of *in situ* germplasm resource compared to those that are not introgressed.

Testing such speculation requires further research. Sunflower is one of few ideal systems. Interestingly, a high-quality reference genome for *Helianthus annuus* was published while I was writing this article (Badouin et al., 2017). The availability of that resource makes the prospects for continued research especially sunny.

## AUTHOR CONTRIBUTION

N.C.E. had fun writing this perspective.

## ORCID

Norman C. Ellstrand  <http://orcid.org/0000-0002-2687-8845>

## REFERENCES

Badouin, H., Gouzy, J., Grassa, C. J., Murat, F., Stanton, S. E., ... Langlade, N. B. (2017). The sunflower genome provides insights into oil

metabolism, flowering and Asterid evolution. *Nature*, 546, 148–152. <https://doi.org/10.1038/nature22380>

- Corbi, J., Baack, E. J., Dechaine, J. M., Seiler, G., & Burke, J. M. (2017). Genome-wide analysis of allele frequency change in sunflower wild-crop hybrid populations evolving under natural conditions. *Molecular Ecology*, 27(1), 233–247. <https://doi.org/10.1111/mec.14202>
- Ellstrand, N. C. (2003a). Current knowledge of gene flow in plants: Implications for transgene flow. *Philosophical Transactions of the Royal Society of London B*, 358, 1163–1170. <https://doi.org/10.1098/rstb.2003.1299>
- Ellstrand, N. C. (2003b). *Dangerous liaisons? When cultivated plants mate with their wild relatives*. Baltimore, MD: JHU Press.
- Ellstrand, N. C. (2012). Over a decade of crop transgenes out-of-place. In C. A. Wozniak, & A. McHughen (Eds.), *Regulation of agricultural biotechnology: The United States and Canada* (pp. 123–135). Dordrecht: Springer. <https://doi.org/10.1007/978-94-007-2156-2>
- Ellstrand, N. C., Meirmans, P., Rong, J., Bartsch, D., Ghosh, A., de Jong, T., ... Hooftman, D. (2013). Introgression of crop alleles into wild or weedy populations. *Annual Review of Ecology, Evolution, and Systematics*, 44, 325–346. <https://doi.org/10.1146/annurev-ecolsys-110512-135840>
- Feldman, M., Lupton, F. G. H., & Miller, T. E. (1995). Wheats. In J. Smartt, & N. W. Simmonds (Eds.), *Evolution of crop plants*, 2nd ed. (pp. 184–192). Harlow: Longman.
- Hooftman, D. A. P., Flavell, A. J., Jansen, H., den Nijs, J. C. M., Syed, N. H., Sørensen, A. P., ... van de Wiel, C. C. M. (2011). Locus-dependent selection in crop-wild hybrids of lettuce under field conditions and its implication for GM crop development. *Evolutionary Applications*, 4, 648–659. <https://doi.org/10.1111/j.1752-4571.2011.00188.x>
- Hufford, M. B., Lubinsky, P., Pyhäjärvi, T., Devengeno, M. T., Ellstrand, N. C., & Ross-Ibarra, J. (2013). The genomic signature of crop-wild introgression in maize. *PLoS Genetics*, 9(5), e1003477. <https://doi.org/10.1371/journal.pgen.1003477>
- Kiang, Y. T., Antonovics, J., & Wu, L. (1979). The extinction of wild rice (*Oryza perennis formosana*) in Taiwan. *Journal of Asian Ecology*, 1, 1–9.
- Linder, C. R., Taha, I., Seiler, G. J., Snow, A. A., & Rieseberg, L. H. (1998). Long-term introgression of crop genes into wild sunflower populations. *Theoretical and Applied Genetics*, 96, 339–347. <https://doi.org/10.1007/s001220050746>
- Merotto, A. Jr., Goulart, I. C. G. R., Nunes, A. L., Kalsing, A., Markus, C., Menezes, V. G., & Wander, A. E. (2016). Evolutionary and social consequences of introgression of NON-transgenic herbicide resistance from rice to weedy rice in Brazil. *Evolutionary Applications*, 9, 837–846. <https://doi.org/10.1111/eva.12387>
- Snow, A. A., Culley, T. M., Campbell, L. G., Sweeney, P. M., Hegde, S. G., & Ellstrand, N. C. (2010). Long-term persistence of crop alleles in weedy populations of wild radish (*Raphanus raphanistrum*). *New Phytologist*, 186, 537–548. <https://doi.org/10.1111/j.1469-8137.2009.03172.x>
- Templeton, A. R., & Levin, D. A. (1979). Evolutionary consequences of seed pools. *American Naturalist*, 114, 232–249. <https://doi.org/10.1086/283471>
- Xia, H. B., Xia, H., Ellstrand, N. C., Yang, C., & Lu, B. R. (2011). Rapid evolutionary divergence and ecotypic diversification of germination behavior in weedy rice populations. *New Phytologist*, 191, 1119–1127. <https://doi.org/10.1111/j.1469-8137.2011.03766.x>

**How to cite this article:** Ellstrand NC. Does introgression of crop alleles into wild and weedy living populations create cryptic *in situ* germplasm banks? *Mol Ecol*. 2018;27:38–40. <https://doi.org/10.1111/mec.14390>