

GENE FLOW AND INTROGRESSION BETWEEN DOMESTICATED CROPS AND THEIR WILD RELATIVES

Roberto Papa

Dipartimento di Scienze degli Alimenti, Università Politecnica delle Marche, Ancona, Italy.
r.papa@univpm.it - Fax: +39-071-2204858 - Phone: +39-071-2204984

Summary

Gene flow between crops and their wild relatives is an important process that has strong implications both for conservation of genetic diversity and for plant breeding. Introgression between wild and domesticated populations has been shown to be a widespread phenomenon in most species, both those allogamous and autogamous. The combined effects of gene flow and selection on the structure of genetic diversity of wild and domesticated populations is discussed.

Keywords

Biodiversity, domestication, landraces, selection, wild progenitor

Along with other evolutionary forces that include genetic drift, mutation and selection, the process of migration is driving the changes in gene frequencies in natural populations. Divergence between populations can be caused by genetic drift because of complete or partial isolation, or by heterogeneous selection among different demes. Migration of individuals (e.g. seeds) or gametes (e.g. pollen) between subpopulations counteracts this divergence between populations, causing a reduction in the genetic diversity between subpopulations, and increasing the level of genetic diversity within subpopulations.

As with genetic drift, migration is a neutral process and it influences gene frequencies of natural populations similarly across the whole genome, while because of recombination, selection induces genetic changes only on target loci (Cavalli Sforza, 1966, Lewontin and Krakauer, 1975). Migration models can be grouped into two categories: 1) symmetric models, which usually assume the same size in all subpopulations and symmetric migration rates (e.g. island model, Wright, 1931). More generally, these migration models assume that for every subpopulation, the number of individuals migrating into the subpopulation is equal to the number of individuals migrating out of the subpopulation. 2) Asymmetric models (e.g. mainland-island model, Harrison 1991), where both subpopulation size and migration rate can vary.

Symmetric migration will affect the structure of genetic diversity by the homogenization of gene frequencies among subpopulations, but it will not affect the total diversity because the gene frequencies of the whole population will remain constant. Differently asymmetric models can drastically affect both the total genetic diversity and the gene frequencies of the whole population, and may lead to the extinction of some subpopulations. Indeed, if migration is present in just one direction, or if one direction is largely prevalent, because of the different sizes (N) between subpopulations or because the migration rates (m) are asymmetric (e.g. unilateral incompatibility, prevalent wind direction, different pollen-tube growth rates), the gene frequencies of the recipient subpopulation will gradually tend to the gene frequencies of the donor subpopulation, leading, in extreme cases, to the extinction of the sink subpopulation (genetic assimilation or demographic swamping; Wolfe et al., 2001). According to the neutral

island model of migration, the amount of gene flow needed to prevent independent evolution between subpopulations is relatively low, about 1 migrant per generation (Nm), suggesting that even if the rate of dispersal and outcrossing are very variable among, and even within, species, in most cases gene flow would be expected to have a prominent role in shaping the structure of genetic diversity of natural populations.

Gene flow between wild and domesticated populations is likely to have been important following domestication, and would have partially restored the genetic diversity of domesticated populations (Badr et al., 2000). In the areas of distribution of the wild relatives, this would also act as a source of the new alleles that have been selected by farmers. In the same areas, landraces are still cultivated and gene flow from wild to domesticated crops remains an important source of genetic variation for their evolution (Jarvis and Hodgkin, 1999). Nowadays, modern plant breeding has also introduced several genes from wild relatives, which have been mainly resistance genes, but have also included QTLs for yield and quality (Tanskley and McCouch, 1997). Gene flow from a domesticated population may cause a reduction in the genetic diversity of the wild population, leading in some cases to the genetic extinction of the wild population (Ellstrand et al., 1999, Wolfe et al., 2001, Gepts and Papa, 2003). Finally, the advent of transgenic varieties in areas of sympatry with wild populations increases the potential risks of transfer of novel genes that may be undesirable in wild populations (i.e. herbicide tolerance), favouring the development of more aggressive weeds. (Jenczewski et al., 2003).

Crops and their wild progenitors show marked phenotypic differences that have been referred to as the “domestication syndrome” (Hammer 1984). These differences result from natural or human selection pressures that are imposed at various stages of the development of the plant under cultivation (Harlan 1992). During the domestication process, similar sets of traits were selected in the domesticated form: a lack of spatial and temporal dispersal systems (shattering, dormancy); the size of the organs utilized by humans (e.g. seeds); a compact plant architecture (determinacy); and a reduced growth cycle (early flowering and maturation, photoperiod insensitiveness). Most of the traits of the domestication syndrome are due to the action of one, or a few, genes. Moreover, although with some exceptions (Burke et al 2002), most of the “domesticated alleles” originate from mutations with a loss of function, and most of the domesticated alleles are recessive (Gepts, 2002).

The process of domestication has consisted of a selection process for adaptation to the cultivated environments and to human needs. Domestication has acted in two ways that have reduced the genetic diversity of domesticated crops in comparison to their wild progenitors (domestication bottleneck): a random genetic drift at the whole genome level, because of the small population size of the original domesticated population; and selection at target and linked loci, because of recombination (hitchhiking). The effects of hitchhiking are expected to be substantial in autogamous species, while much lower in allogamous species. Indeed, in the allogamous species *Zea mays*, hitchhiking appears to have affected restricted genomic regions (a few hundreds base pairs) around selected sites of domestication loci (Wang et al., 1999, 2001; Tenaillon et al., 2001; Clark et al., 2004).

Despite phenotypic differences that have originated from the domestication process, in general, crops and their wild progenitors belong to the same biological species (Harlan and de Wet, 1971). Thus, while wild and domesticated populations are sympatric (i.e. grow in close proximity) and share a similar phenology (i.e. their flowering periods overlap), hybridization is expected to occur and result in fertile progeny. With the possible exception of *Arachis hypogea*, this is indeed what has been observed in both allogamous and autogamous species in various studies that have used different methods and approaches (Ellstrand et al. 1999).

In some cases hybridization between wild and domesticated populations can promote the development of weedy populations that can be found in areas of cultivation, in abandoned fields or in other disturbed areas (e.g. roadsides), and these will present intermediate characteristics between the wild and domesticated forms. However, in some cases a weedy population can originate from genotypes that have escaped from cultivation. In both cases, the weedy populations may act as a “genetic bridge”, favouring the introgression between the wild and domesticated types. However, wild and domesticated types often maintain their phenotypic distinctiveness when they are sympatric. For cross-pollinating species such as maize, the bi-directional selection for alternate states of traits of the domestication syndrome is a factor that is likely to explain the continued existence of crops and their wild relatives in sympatry. For self-pollinating species, the autogamous reproductive system provides a plausible alternative explanation.

As mentioned above, several examples of introgression between wild and domesticated populations have been shown that indicate that gene flow acts in both directions: from domesticated to wild, and from wild to domesticated (Ellstrand et al., 1999; Jarvis and Hodgkin, 1999). Without considering introgression due to modern plant breeding, gene flow from wild to domesticated populations can occur only when farmers use part of their crop production as seed for the next generation of planting, without replacing the seed with that of commercial varieties. This situation generally occurs in traditional agriculture by the use of landrace populations. Indeed, in the context of modern agriculture, gene flow can act only in one direction: from domesticated to wild populations. Recently, Papa and Gepts (2003) used admixture proportions to show that gene flow between wild and domesticated landrace populations of common beans leads to asymmetric introgression, with a higher rate from domesticated to wild populations than in the reverse direction. Similarly, other studies have suggested that introgression is higher from domesticated to wild populations (Jenczewski et al., 1999, Matsuoka et al., 2000, Wolfe et al., 2001), probably because of the much higher population sizes of the domesticated populations, as compared to the wild populations, and because of the different types of selection in the cultivated and wild environments (Papa and Gepts 2004). First, in the cultivated environment, in addition to the selection pressures imposed by the agroecosystem, a direct selection operated by the farmers may also be acting. Secondly, because the “domesticated alleles” of the genes responsible for the genetic control of the domestication syndrome are generally recessive, the first generation hybrids will show a phenotype similar to the wild parent, and they will thus be much better adapted to the wild than to the cultivated environment. For this reason, selection in the cultivated environment will act mainly through the first hybrid generation, while in the wild environment it will mainly act in the following generations, allowing recombination and thus introgression of alleles from the domesticated population.

Thus, both differences in population sizes and selection regimes promote asymmetric introgression, and will lead to the displacement of the wild alleles, which will be replaced with those of the domesticated population. Because of the lower diversity present in the domesticated as compared to the wild forms (domestication bottleneck), the effect of asymmetric introgression will be a reduction in the genetic diversity of the wild population. However, because of recombination and selection at domestication loci, this process will not affect the whole genome evenly. Indeed, because of linkage around the domestication loci, maintenance of the original diversity of the wild populations would be expected. Results obtained in the strict selfer *Phaseolus vulgaris* (Papa and Gepts, unpublished results) confirm this hypothesis, and they show that the areas of the genome where introgression from the domesticated forms is limited by selection can extend over a few centimorgans. Indeed, the size of these “protected windows of the genome” will depend on the effective recombination

rate. High recombination, such as that observed in allogamous species, will restrict these areas to very small portions of the genome, while in autogamous species the size of these regions may be much larger. Considering the results related to the structure of genetic diversity along the genomes of the allogamous *Zea mays* (Wang et al., 1999) and the autogamous *Arabidopsis thaliana* (Nordborg et al., 2002), we can expect that the size of these “protected windows” can vary, depending on the breeding systems of a species, from a few hundreds of base pairs to hundreds of Kb.

Overall, it has been shown that gene flow between wild and domesticated populations is a relevant process in shaping the structure of genetic diversity and in the evolution of crops and their wild relatives. The combined effects of gene flow and selection appear to promote higher introgression from domesticated populations, rather than in the reverse direction. This asymmetric introgression from domesticated to wild populations will be likely to reduce the diversity of the wild forms, which in extreme cases could lead to the genetic extinction of the wild population. Moreover, introgression appears to be unevenly distributed along the genome, particularly in autogamous species. To disentangle the effects of migration and selection, both putative neutral markers and molecular markers that are likely to be affected by selection should be used, possibly in combination with molecular maps.

THE ROLE OF BIOTECHNOLOGY
Villa Gualino, Turin, Italy – 5-7 March, 2005

REFERENCE LIST

1. Badr A, Muller K, Schafer-Pregl R, El Rabey H, Effgen S, Ibrahim HH, Pozzi C, Rohde W, Salamini F. (2000). On the Origin and Domestication History of Barley (*Hordeum vulgare*). *Molecular Biology and Evolution*, 17: 499-510.
2. Cavalli-Sforza, L. L., 1966 Population structure and human evolution. *Proc. R. Soc. Lond. Ser. B* 164:362-379
3. Clark, R. M., E. Linton, J. Messing and J. F. Doebley, 2004 Pattern of diversity in the genomic region near the maize domestication gene *tb1*. *Proc. Natl. Acad. Sci. USA* 101: 700–707
4. Ellstrand, N., H. Prentice, and J. Hancock. 1999. Gene flow and introgression from domesticated plants into their wild relatives. *Annu. Rev. Ecol. Syst.* 30: 539-563.
5. Hammer, K. 1984. Das Domestikations syndrom. *Kulturpflanze* 32: 11-34.
6. Harlan, J. R. 1992. *Crops and man*. American Society of Agronomy, Madison, WI.
7. Harlan, J. R., and J.M.J. de Wet 1971. Towards a rational classification of cultivated plants. *Taxon* 20: 509-517.
8. Lewontin, R. C. and J. Krakauer, 1975 Testing the heterogeneity of F values. *Genetics* 80:397-398
9. Harrison, S., 1991. Local extinction in a metapopulation context: an empirical evaluation. *Biological Journal of the Linnean Society* 42, 73-88
10. Gepts P. 2002. A comparison between crop domestication, classical plant breeding, and genetic engineering. *Crop Science* 42: 1780-1790
11. Gepts, P., and Papa., R. (2003). Possible effects of (trans) gene flow from crops on the genetic diversity from landraces and wild relatives. *Environmental Biosafety Research* 2: 89-103
12. Jarvis, D.I. and Hodgkin, T. (1999) Wild relatives and crop cultivars: detecting natural introgression and farmer selection of new genetic combinations in agroecosystems. *Molecular Ecology* 8, S159–S173.
13. Jenczewski, E., Prosperi, J.M. and Ronfort, J. (1999). Differentiation between natural and cultivated populations of *Medicago sativa* (Leguminosae) from Spain: analysis with random amplified polymorphic DNA (RAPD) markers and comparison to allozymes. *Molecular Ecology* 8, 1317-1330.
14. Jenczewski, E., Ronfort, J., and A.M. Chèvre, (2003). Crop-to-wild gene flow, introgression and possible fitness effects of transgenes. *Environmental Biosafety Research* 1, 9-24
15. Papa R, and Gepts P. 2003. Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris* L.) from Mesoamerica. *Theor. Appl. Genet.*
16. Papa R and Gepts P. 2004 Asymmetric gene flow and introgression between domesticated and wild populations. In: H.C.M. den Nijs, D. Bartsch, J. Sweet (eds.), *Introgression from Genetically Modified Plants into Wild Relatives*. CABI, Wallingford, Oxfordshire, UK
17. Matsuoka, Y., Vigouroux, Y., Goodman, M.M., Jesus Sanchez, G., Buckler, E., and Doebley, J., (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Science USA* 30, 6080-6084.
18. Nordborg, M.; Borevitz, J.; Bergelson, J.; Berry, C.; Chory, J.; Hagenblad, J.; Kreitman, M.; Maloof, J.; Noyes, T.; Oefner, P.J.; Stahl, E.A.; Weigel, D. The extent of linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genetics* 2002, 30, 190-193.
19. Tanksley, S.D. and McCouch, S.R. 1997. Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 1997, 277, 1063-1066
20. Tenaillon, M.I., Sawkins, M.C., Long, A.D., Gaut, R.L., Doebley, J.F., and Gaut, B.S. (2001). Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). *Proc. Natl. Acad. Sci. USA* 98, 9161–9166
21. Wang, R.L., A. Stec, J. Hey, L. Lukens, and J. Doebley 1999. The limits of selection during maize domestication. *Nature* 398: 236-239.
22. Wang R-L, Stec A, Hey J, Lukens L, Doebley J (2001) correction: The limits of selection during maize domestication. *Nature* 410: 718.
23. Wright, S., 1931 Evolution in Mendelian populations. *Genetics* 16: 97-159.
24. Wolfe, D.E., Takebayashi N., and L. H. Risenberg, 2001. Predicting the risk of extinction through hybridization. *Conservation Biology* 15:1039-1053

THE ROLE OF BIOTECHNOLOGY
Villa Gualino, Turin, Italy – 5-7 March, 2005