

Detecting multiple origins of domesticated crops

Kenneth M. Olsen* and Briana L. Gross

Department of Biology, Washington University, St. Louis, MO 63130-4899

Crop domestication has long been studied both as a model for understanding the process of evolution (1) and for gaining insights into the history of human civilization (2). In recent decades, a wealth of neutral molecular markers (e.g., SNPs, microsatellites, amplified fragment length polymorphisms) has become available for many crop species, permitting genomewide examinations of genetic diversity in crops and their wild relatives. Quite commonly, these surveys reveal that the present-day representatives of a crop all show close genetic similarity to a specific geographical subset of a wild species (e.g., ref. 3; Fig. 1A); this pattern is typically interpreted as evidence that the crop was domesticated a single time and in a single geographical region. Interestingly, this picture of a single, geographically localized crop origin does not always match the picture inferred from archaeological data.

In particular, for Old World cereal crops, archaeobotanical data have indicated that domestication was likely a geographically diffuse and protracted process, involving long-term predomestication plant use across wide areas of the Near East (4). As such, the present-day diversity of a crop might well be expected to reflect multiple, geographically disparate origins of domestication (Fig. 1B). Why then, do multilocus genetic studies reveal monophyletic origins of Old World cereal crops? A study by Allaby *et al.* (5) in this issue of PNAS suggests that, given certain population parameters, the genetic signature of a single origin can arise even if multiple domestication events have occurred.

Allaby *et al.* (5) have conducted a series of simulations in which they ask the following question: given a true domestication history that involves two independent events (Fig. 1B), how often do multilocus, neutral markers lead to the inference of a single domestication origin (Fig. 1A)? Their results suggest that with a sufficiently protracted domestication period the genetic evidence of multiple domestications events may be lost. The key to this “protracted model” lies in the power of genetic drift: if domestication occurs over a long period, with effective population sizes of just a few hundred individuals for many generations, then drift may be strong enough to eliminate the genetic traces of multiple domestications, leaving only a single

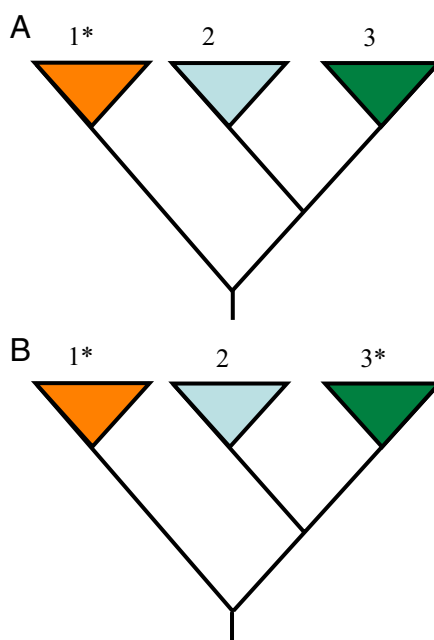


Fig. 1. Potential relationships between a crop and wild progenitor populations. Crops carry a subset of the neutral genetic variation found in their wild relatives. (A) Single origin of domestication. (B) Multiple domestication origins. Each numbered clade corresponds to a genetically and geographically definable group of wild populations; asterisks indicate clades sharing neutral variation with the crop.

event detected in a crop’s present-day genetic diversity.

The Genetic Consequences of Domestication

Crop domestication, which began $\approx 12,000$ years ago, represents a recent event on an evolutionary time scale. Because crops have not had time to diverge from their wild ancestors through mutation at neutral genes, the neutral genetic diversity in a crop is expected to be a subset of that found in wild populations. Two key processes govern the reduction in genetic diversity that occurs in a crop lineage during domestication: selection by humans for desirable “domestication traits” (e.g., loss of seed dispersal mechanisms) and genetic drift in the form of “domestication bottlenecks” (6), which occur as plants are taken from wild populations and brought into cultivation. Whereas selection only affects genetic diversity at the genes that underlie traits being selected upon (or loci in linkage disequilibrium with those genes), bottlenecks reduce neutral genetic

diversity across the entire genome. Because of domestication bottlenecks, crops often contain $\approx 70\%$ of the neutral genetic diversity present in their wild ancestors (7, 8).

The classic model of a domestication bottleneck envisions genetic drift as being restricted to the initial stages of domestication (ref. 6 and Fig. 2A). The strength of genetic drift during this bottleneck is determined by two interacting factors: the size of the bottlenecked population (N_b) and the bottleneck’s duration (d generations). The severity of the bottleneck (k) can thus be quantified as the ratio of these values: $k = N_b/d$ (9); a smaller value of k predicts a more severe reduction in neutral genetic diversity. Under the classic model, it is assumed that once a crop has passed through this initial domestication bottleneck and becomes widely cultivated, the population size (N_p) is large enough that any subsequent drift effects are minimal (Fig. 2A).

The protracted model of Allaby *et al.* (5) differs from this classic model in that genetic drift operates not only during an initial domestication bottleneck, but also for many subsequent generations. In the simulations used for their study, an initial population bottleneck ($N_b = 20$, $d = 10$) is followed by tens or hundreds of additional generations with 50–150 individuals, either with a single domestication event (Fig. 2B) or two independent events followed by admixture (Fig. 2C). As is illustrated in Fig. 2B, the bottleneck severity during this protracted period can easily surpass the severity of the initial domestication bottleneck. The compounded effects of these bottlenecks would thus be very effective at eliminating genetic diversity in the crop lineage, including, potentially, evidence of multiple domestication events. This prediction appears to be borne out in the simulations conducted by Allaby *et al.* For example, in a model with two independent domestication events and a population of $n = 100$ maintained for 200 generations after admixture (see Fig. 2C), $\geq 90\%$ of simulations yield patterns of genetic diversity

Author contributions: K.M.O. and B.L.G. wrote the paper.

The authors declare no conflict of interest.

See companion article on page 13982.

*To whom correspondence should be addressed. E-mail: kolsen@wustl.edu.

© 2008 by The National Academy of Sciences of the USA

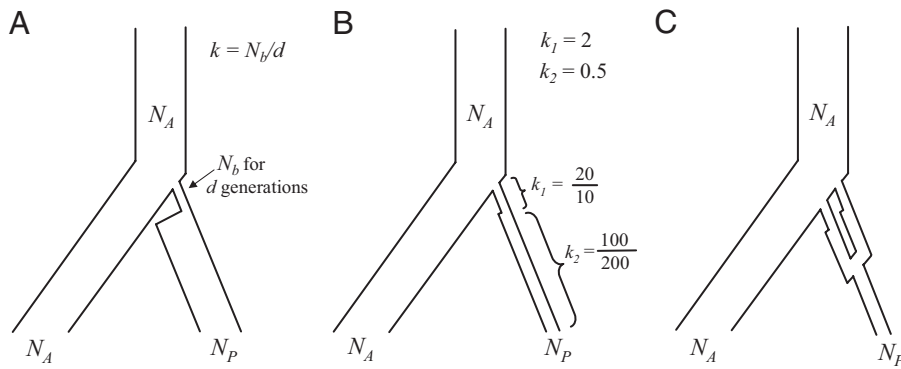


Fig. 2. Models of domestication bottlenecks. (A) The classic model (6). (B) The protracted model (5). Population sizes remain small for tens of hundreds of generations after the initial bottleneck, creating protracted and severe bottleneck effects. Values of k are shown for parameters used by Allaby *et al.* (5) in their simulations. (C) The protracted model shown with two independent domestication events, followed by admixture. N_A , population size of the wild ancestor population; N_b , population size during the domestication bottleneck; d , duration (in generations) of the domestication bottleneck; k , bottleneck severity, calculated as N_b/d (9); N_P , present-day crop population size.

indicating a single origin of domestication; this result holds true for a range of recombination parameters (5).

Is the Protracted Model a Realistic Model for Crop Evolution?

The key to this question lies in genetic diversity and whether biologically realistic levels of neutral variation could actually persist in a crop lineage through the compounded bottlenecks modeled by Allaby *et al.* (5) (e.g., $\approx 70\%$ of the ancestor's genetic diversity; refs. 7 and 8). If so, their model may be a plausible representation of the domestication process. This hypothesis could be tested by

quantifying genetic diversity at multiple time points during simulations of their model. Other important insights could be gained by explicitly modeling population structure/gene flow among wild populations and/or crop lineages; these biologically important forces are likely to strongly influence patterns of monophyly in the simulations.

Assuming the protracted model does prove to be a realistic model of domestication, what does the Allaby *et al.* study (5) tell us about the validity of multilocus studies and “monophyletic” clusters for inferring crop origins? Insight into this question is provided by the specia-

tion literature, where recent research has examined the potential for multiple independent origins of a single taxon. A caveat revealed in these studies is that multilocus cluster analyses can generate a monophyletic grouping for a hybrid species even if that species has evolved multiple times independently (10); this scenario potentially parallels that of independent domestication events followed by admixture. Thus, in assessing whether a crop's monophyletic grouping is real, it is critical to examine whether the alleles in the crop are truly a subset of those in the wild populations with which the crop is most closely clustered, as would be expected in a progenitor/derivative relationship.

Finally, it is important to draw a distinction between what genetic data can tell us about the origin of extant crop varieties, and what they can (and cannot) tell us about the overall history of a crop's cultivation. Domestication bottlenecks eliminate genetic variation, and given protracted bottlenecks (e.g., Fig. 2 B and C), the alleles representing multiple domestications may be lost. This does not mean that multiple, independent domestications did not occur, only that their descendants are not represented in the crop today. For understanding the complete picture of crop domestication, we will always benefit by combining the insights gained from both genetic and archaeological data.

ACKNOWLEDGMENTS. Research on the population genetics of crop domestication in K.M.O.'s laboratory is supported by National Science Foundation Grant DBI-0638820.

- Darwin C (1859) *Origin of Species* (Murray, London).
- Diamond JM (1997) *Guns, Germs, and Steel: The Fates of Human Societies* (Norton, New York).
- Spooner DM, McLean K, Ramsay G, Waugh R, Bryan GJ (2005) A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *Proc Natl Acad Sci USA* 102:14694–14699.
- Weiss E, Kislev ME, Hartmann A (2006) Autonomous cultivation before domestication. *Science* 312:1608–1610.
- Allaby RG, Fuller DQ, Brown TA (2008) The genetic expectations of a protracted model for the origins of domesticated crops. *Proc Natl Acad Sci USA* 105:13982–13986.
- Eyre-Walker A, Gaut RL, Hilton H, Feldman DL, Gaut BS (1998) Investigation of the bottleneck leading to the domestication of maize. *Proc Natl Acad Sci USA* 95:4441–4446.
- Buckler ES, Thornsberry JM, Kresovich S (2001) Molecular diversity, structure, and domestication of grasses. *Genet Res* 77:213–218.
- Gepts P (2004) Crop domestication as a long-term selection experiment. *Plant Breeding Rev* 24: 1–44.
- Wright SI, *et al.* (2005) The effects of artificial selection of the maize genome. *Science* 308:1310–1314.
- Gross BL, Schwarzback AE, Rieseberg LH (2003) Origin(s) of the diploid hybrid species *Helianthus deserticola* (Asteraceae). *Am J Bot* 90:1708–1719.