The contribution of genetic and genomic approaches to plant domestication studies

Paul Gepts

The application of genomic approaches to the phenomenon of plant domestication promises a better understanding of the origins of agriculture, but also of the way plant genomes in general are organized and expressed. Building on earlier genetic research, more detailed information has become available on the organization of genetic diversity at the genome level and the effects of gene flow on diversity in different regions of the genome. In addition, putative domestication genes have been identified through population genomics approaches (selective sweeps or divergence scanning). Further information has been obtained on the origin of domestication syndrome mutations and the dispersal and adaptation of crops after domestication. For the future, increasingly multidisciplinary approaches using combinations of genomics and other approaches will prevail.

Addresses
University of California, Department of Plant Sciences/MS 1, Section of Crop and Ecosystem Sciences, 1 Shields Avenue, Davis, CA 95616, United States of America

Corresponding author: Gepts, Paul (plgepts@ucdavis.edu)

Introduction

The importance of the Neolithic revolution, which took place some 12,000–10,000 years ago, cannot be overstated. At that time, the human species embarked on a radical change in the way it had procured its food and other products for millions of years. In addition to the transition from hunting-gathering to plant (and animal) agriculture, other major changes were taking place more or less simultaneously including sedentism, the formation of villages, and the introduction of ceramics. Eventually, these developments led to the formation of more hierarchical societies and states and the development of civilizations [1,2]. The Neolithic revolution had a tremendous impact on the biosphere and human societies, consequences that still affect us today. This transition took place in several areas of the Earth and often gave rise to a complementary group of crops that satisfied human needs, not only nutritionally but for other uses as well (Figure 1).

One of the key technological elements of the transition to agriculture is domestication. The cultivation of plants represents an important change in their environment. These changes set in motion several evolutionary processes that led to major morphological and physiological changes, such that often crop plants became quite different from their wild progenitors and, in the most pronounced cases of domestication, could not survive without human intervention [3].

Domestication draws attention from several viewpoints. In addition to its role in human cultural evolution, it is also an experimental system for the study of biological evolution. In addition, many of the differences between wild and domesticated types are important traits or processes in basic biology such as plant development, reproduction, and adaptation [4]. Furthermore, the efficient utilization of crop genetic resources for varietal breeding depends on detailed information about genetic relationships in these collections, which were affected in large part by the domestication process [5,6].

Elucidating the transition from hunting-gathering to agriculture has always been a multidisciplinary research endeavor (e.g. [7–13,14•,15]; for a recent e.g.: [16•]). As science has progressed, new approaches have been incorporated (molecular biology: [17,18•]; bioinformatics: e.g. [19]).

Genomic approaches have become especially important since the introduction of next-generation DNA sequencing [20,21] because these technologies allow for the sequencing of multiple genotypes in a low-cost, high-throughput way, and open new avenues for population, landscape, and ecological genomics [22–25].

Origins of domestication and structure of wild and domesticated gene pools

One of the most generalizable observations in crop domestication studies is the reduction in genetic diversity during and subsequent to domestication, for example, [26–29,30•,31,32,33••,34,35,36•]. This reduction is due to the two-fold action of genetic drift and selection that operate at each stage of the domestication process (Figure 2). Additional factors affect reduction in genetic diversity, including the specific stage of the crop evolution process, with the initial domestication phase usually
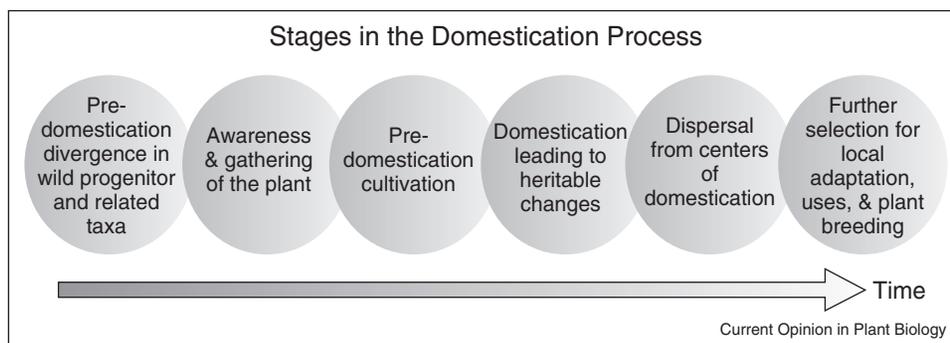
Figure 1

Current Opinion in Plant Biology
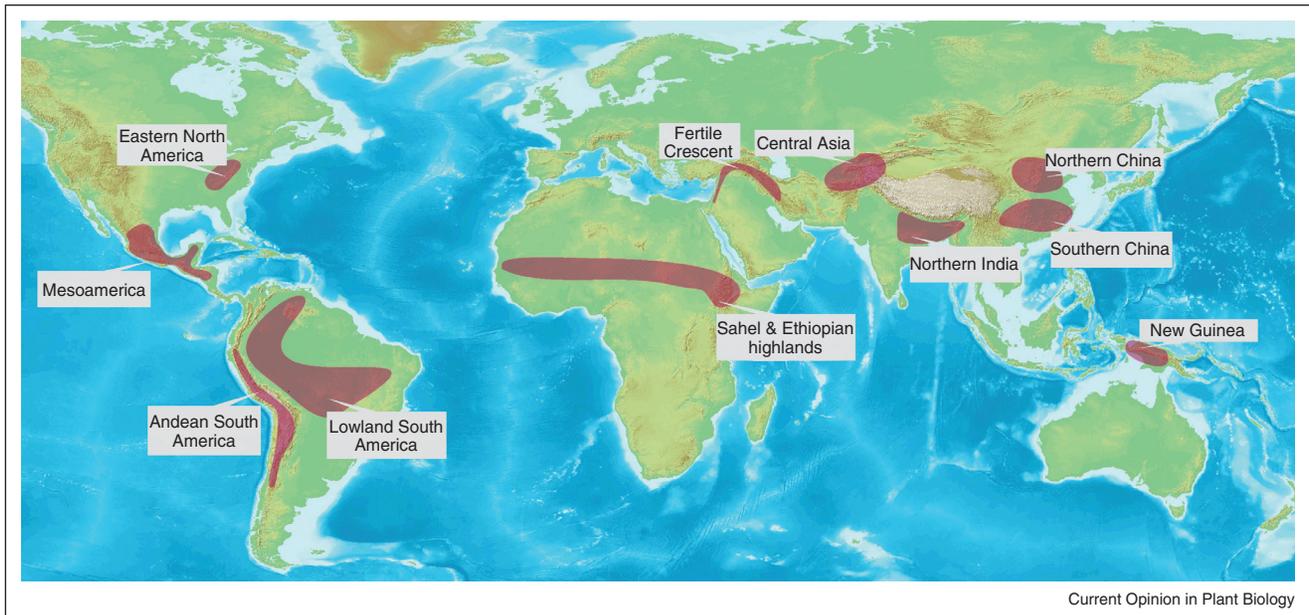
Centers of crop origin and domestication. Shaded regions indicate approximate location of centers of origin and domestication of selected crops. Regions are approximate due to uncertainties in the distribution of the wild progenitors, limited numbers of sites with archaeobotanical remains, and the area of actual domestication. Examples of crops domesticated in their respective centers are: Lowland South America: cassava, sweet potato pineapple, groundnut, cashew, guaraná, peach palm, *Capsicum* peppers; Andean South America: potato, other Andean root crops (e.g. oca, arracacha), common & lima bean, lupin, *Capsicum* peppers, Pima cotton, coca, quinoa, squash; Mesoamerica: Maize, common & lima bean, *Capsicum annuum* pepper, Upland cotton, sisal, papaya, avocado, prickly pear, squash, tomato, vanilla, cacao, amaranth; Eastern North America: sunflower, sumpweed, goosefoot; Sahel & Ethiopian Highlands: African rice, pearl millet, sorghum, cowpea, Bambara groundnut, hyacinth bean, tef, fonio, yam, watermelon, melon, okra, kenaf, coffee, ensete, noog; Fertile Crescent: wheat, barley, oat, pea, chickpea (garbanzo), lentil, faba bean, flax, olive, date palm, grape, onion, lettuce, saffron, poppy; Central Asia: apple, carrot; Northern India: Asian rice (indica), mung bean, pigeon pea; Northern China: broomcorn millet, foxtail millet, Chinese cabbage, soybean; Southern China: Asian rice (japonica), quince, persimmon, litchi, peach, tea; New Guinea: banana, sugar cane.

causing the biggest drop in genetic diversity (e.g. [36•,37]) compared to later steps, including modern plant breeding efforts [32,35,37] (Fig. 2).

Recent genomic studies have improved our knowledge of the fate of genetic diversity in two ways. First, they have expanded our horizons to other crops than the major cereal crops and grain legumes and provided contrasting views of changes in genetic diversity. Next-generation sequencing, various forms of reduced representation libraries, and synteny combined with considerable bioinformatic analyses make it now easier to directly sequence

Figure 2

Stages in the Domestication Process

Pre-domestication divergence in wild progenitor and related taxa

Awareness & gathering of the plant

Pre-domestication cultivation

Domestication leading to heritable changes

Dispersal from centers of domestication

Further selection for local adaptation, uses, & plant breeding

Time

Current Opinion in Plant Biology

Successive stages in the transition from wild-gathered to domesticated plants. The actual duration of the transition may depend on the crop plant and center of domestication, but is thought to have lasted from several 100 to several 1000 years.

additional taxa (even non-model ones) or additional accessions within taxa, with or without the availability of a whole-genome reference sequence (e.g. [27,33[••],38,39[•],40[•]]). The reduction in genetic diversity can be much more modest in fruit and nut trees or vines [39[•]], which have been propagated mostly vegetatively with a limited number of sexual generations. By contrast, certain vegetables (e.g. [33[••]]) exhibit a stronger reduction in genetic diversity, probably reflecting a more severe bottleneck as may be expected from crops planted mainly in gardens or smaller field surfaces and maintained for specific horticultural phenotypes. An exception to the pattern of reduced genetic diversity upon domestication, is carrot [40[•]] in which no reduction was observed, due to post-domestication gene flow between wild and domesticated types and a strong inbreeding depression.

Second, compared to genetic information, genomic approaches provide additional information related to the general reduction in genetic diversity. Linkage disequilibrium (LD) and the size of haplotype blocks are increased in domesticates compared to their wild ancestor [26,28,37,41[••]]. Further increases in LD can be seen within the domesticated gene pool as a consequence of selection for particular domesticated types such as in yard-long beans, an Asian variety of cowpea (*Vigna unguiculata* L. Walp.), the latter having been domesticated in Africa [42]. Individual genome areas may show different patterns. For example, in barley (*Hordeum vulgare*) in the *Rrs2* region controlling resistance to barley leaf scald (*Rhynchosporium secale* (Oudem.) J.J. Davis), an increase in the number of haplotypes and recombinations was observed in the domesticated gene pool [31], potentially attributable to recent selection by the pathogen, heterogeneous selection in time and space, or redundancy. Thus, genomics provides both a broader view across taxa and across individual genomes on the effects of domestication on genetic diversity.

Effects of gene flow
The widespread occurrence of gene flow should not be a surprise given that in most cases crop plants and their wild progenitors belong to the same biological species. Gene flow between wild and domesticated populations has multiple potential consequences, from the positive (increase in diversity and adaptation [43,44[•]]) to the negative (unwanted escape of genes leading to crop weeds or spread of transgenes [45,46]). The magnitude of gene flow between wild and domesticated types can be studied in sympatric pairs of wild and domesticated populations [47–50,44[•]]. These studies share several results. First, the wild and domesticated types remain overall phenotypically distinct in spite of the detectable presence of gene flow. Second, introgression is not uniform across the genome. Regions lacking introgression showed increased differentiation, decreased diversity, and a higher number of fixed or private differences. Gene flow may also be asymmetric although the direction is inconsistent: mainly from domesticated to wild in common bean [47] and conversely in maize [44[•]]. Genetic diversity is decreased and differentiation is increased around putative domestication genes, suggesting that introgressed traits may be selected against. In maize, introgression from *Zea mays* ssp. *mexicana* into domesticated maize at higher altitude in central Mexico explains why highland Mexico maize is closer to *Z. mays* ssp. *parviglumis*, the actual progenitor of domesticated maize, even though the latter is distributed at lower and mid-altitudes in Mexico [51].

Gene flow can lead to incongruence between genome-wide and local genomic patterns of genetic diversity, LD and genealogical history, as shown by rice [29,52]. Although the former article [29] follows a 'bottom up' approach based on population genetic argument (including the identification of low-diversity regions presumably resulting from selection), the later article [52] reviews top-down experiments leading to the cloning of specific genes underlying important rice domestication traits. Both, however, end up positing selection of domestication alleles in *japonica* (or proto-*japonica*) types followed by introgression into *indica* types. Clearly, a better understanding of genetic diversity in the rice ancestor, *Oryza rufipogon* is needed, based on more comprehensive field collections to elucidate fully the rice domestication history. Overall, the abundance of markers provided by genomic approaches allows us to identify specific regions of the genome that have been affected by gene flow and the relationship of these regions with those harboring genes controlling the domestication syndrome.

Discovery of domestication syndrome genes through population genomic approaches — selective sweeps and association studies
The domestication syndrome is a set of traits that were or are still selected for because they confer adaptation to a cultivated environment such as a reduction in seed dormancy, a more compact, less branched growth habit, a larger number or greater size of fruit and seed, and reduced seed shattering. Additional traits include those that please the consumer through novelty (e.g. pigmentation), better taste, reduced toxicity, or additional uses. Multiple criteria can now be used to determine whether a gene is the underlying molecular basis of a phenotypic trait involved in the domestication syndrome. From a bottom-up standpoint, population genetics argument can be used such as increases in LD and F_{ST} or decreases in genetic diversity in domesticated types compared to the wild progenitor. From a top-down perspective, biparental or multi-parental mapping and synteny with candidate genes can be used. An advantage of the bottom-up approach is that it does not assume prior knowledge of relevant domestication traits [4].

Traditionally, the inheritance of the domestication syndromes has been determined by a top-down approach,

namely QTL analyses in biparental crosses domestica-ted × wild. Results showed that many traits were controlled at least in part by a few major genes, that genetic effects predominated over environmental ones, and that some genes were linked [6]. In addition, candidate genes and synteny can be used to identify causal genes. An example of the latter is that determinacy, a growth habit and flowering time trait, is controlled by *PvTFL1y*, a homologue of the *Arabidopsis TFL1* gene [53]. Given the importance of earliness for food security of subsistence farmers, it is perhaps not surprising that farmers have selected the determinacy trait multiple times in the common bean domesticated gene pool, as revealed by the diversity of mutations identified in *PvTFL1y* [54••]. These results also underscore the need to understand selection by farmers to better interpret genetic diversity studies in crops such as common bean (e.g. [55,56]) and pearl millet (see later section).

These top-down studies are now complemented with bottom-up, population genetic screenings of genome-wide diversity, including sequence diversity departures from neutrality, scans for selective sweeps or highly divergent regions (F_{ST}), or genome-wide association studies (GWAS). Further research is then needed to identify or confirm the causal genes involved. Comparison of sequence diversity combined with neutrality tests in wild and domesticated types can identify regions that appear to be affected by domestication selection. Furthermore, wild-landrace comparisons focus specifically on the effect of initial domestication, whereas landrace-improved cultivar comparisons measure the effect of selection subsequent to domestication (including modern breeding). A few studies [26,36•,32] now suggest that some 3–4% of the genome is affected by domestication and improvement. Comparison of map locations of regions identified by QTL and re-sequencing showed overlap for both domestication and post-domestication selections [36•]. Some of the regions identified tended to be linked as has been reported before in several crops [6].

In maize, modern breeding has had a negligible effect on genome-wide diversity or mean haplotype length; selection coefficients are lower than those for domestication [32]. It is noteworthy that ~25% of domestication features were also breeding features, indicating broad overlap between the two trait concepts. An intriguing observation is the reduction in expression variation suggesting removal of *cis*-located variation and a selection focus on highly expressed genes [32], perhaps to achieve genotypically and environmentally more stable trait expression necessary for broader adaptation. In an analysis of modern US maize germplasm, an increase in differentiation and linkage disequilibrium was observed concurrently with the development of heterotic groups that constitute the basis of current hybrid maize improvement. Overall, 5% of SNPs showed some evidence of directional selection,

involving shade and stress response, lignin biosynthesis, and auxin response and synthesis [57•]. An additional recent example of selective sweeps include several regions in sorghum encompassing genes for starch synthesis enzymes, seed shattering, plant height and maturity loci [41••]. The latter study stands out by its large plant ($n \sim 1000$ accessions) and genome ($\sim$265,000 SNPs) samples.

Origin of the domestication syndrome: pre-existing or *de novo* variation?

In the study of the inheritance of adaptedness, a central question is the origin of mutations upon which selection acts. Either these mutations are part of the standing variation in the population under selection, or they arise *de novo*. Compared with *de novo* mutations, existing mutations will allow populations to evolve faster because there will be no lag period and they may occur at higher frequency. In addition, selection of the two types of mutations will have different selection signatures [58–60]. Selection for adaptation in humans has acted both on standing variation and *de novo* mutations [59,60].

A selection experiment for flowering date in two maize inbred lines [61] showed generally a rapid response to selection for both earlier and later flowering in spite of the presumably uniform genetic background. There was also evidence for the segregation of a major gene in the late flowering selection of one of the inbreds. Thus, the response was based not only on standing variation but also on the appearance of novel, large-effect mutations.

The single-stem phenotype of domesticated maize is controlled by the dominant *Tb-1* allele. This allele arose before domestication through the insertion of a *Hopscotch* transposon some 60 kb upstream of the actual gene leading to an overexpression at the locus, consistent with the dominant nature of the allele selected during domestication. The insertion was dated at some 28,000–23,000 years BP, thus, predating domestication [62•]. An unanswered question is the dynamics of standing variation mutations, such as *Tb1*, in the original population before domestication selection. What mechanisms or forces account for their survival, given that their single-branched phenotype may reduce fitness in natural populations?

A more systematic application of genomic approaches would allow us to assess how often the two types of mutations have been selected during domestication. In turn, this information could contribute to the debate on the speed of domestication. The existence of standing variation, the rapid response to selection observed in empirical selection experiments and computer models, and the simple genetic architecture of the domestication syndrome suggest that, from a genetic perspective, domestication could have proceeded rapidly. However, archaeological data appears to suggest otherwise, that

is, that domestication may have taken place during a long pre-domestication cultivation period of several 100 to several 1000 years based on the presence of arable species and large quantities of seeds in some archaeological sites [63]. Further research is needed on the potential speed *per se* of selection, with regard to a better definition and evidence for the pre-domestication cultivation period.

Post-domestication dispersal, adaptation, and the role of contemporary farmers

In its narrow definition, domestication refers to the initial process of selection that took a wild plant to a domesticated one adapted to cultivation and consumers' needs in its center of domestication. The evolution of crops, however, does not stop there. Significant changes to the crop's genetic constitution can still take place in subsequent evolution through migration and dissemination, setting in motion a succession of genetic drift, selection, and recombination episodes that can have a profound effect on the organization of genetic diversity of the crop. Numerous studies have traced the dispersal of crops from their centers of domestication. However, far fewer have yet to apply genomic approaches to this issue.

Genomic approaches could avail themselves of the possibility of using ancient DNA in crop remains. The use of DNA from plant remnants of varying age to explore issues of crop evolution and domestication has been limited although some outstanding examples exist (e.g. [64]). For example, in a study on the spread of sweet potato [65•], strong evidence was provided of a pre-Columbian spread of sweet potato from tropical America to the Pacific islands based on an analysis of DNA from contemporary populations and herbarium samples, including three herbaria from the Cook expedition in 1769.

Pearl millet is a major cereal in Western Africa and plays an important role as a drought-tolerant crop. Flowering time is a crucial adaptation trait given the existence of a north-south gradient in rainfall across the Sahel. Contemporary pearl millet farmers adapt to this situation in part through the choice of early and late planting material. Hence, there has been much interest in identifying potential targets of selection among different flowering time genes by a combination of approaches, while disentangling selective from demographic and gene flow effects through the development of alternative demographic models by Approximate Bayesian Computation [30•,66–69].

Transcriptomics, proteomics and metabolomics of domestication

More recently, a few studies have addressed the functional genomics aspects of domestication to assess how many and which genes show differences in expression in comparisons between wild and domesticated types. In maize and tomato [32,70,71]. In addition, placing these analyses in a phylogenetic context including wild and domesticated species helps in identifying distinct expression patterns that correlate with adaptive phenotypes, such as in *S. pennellii*, a desert-distributed species, and infer domestication effects [71].

In developing cotton fibers, some 15% of 1300 proteins were significantly up-regulated or down-regulated. Most of the changes took place in the early development stages consistent with human selection for earlier activation of fiber elongation in domesticated compared to wild types. Nevertheless, there was only limited overlap between transcript-level and protein-level inferences, as may be expected because protein abundances depend on translation, localization, post-translational modifications and degradation processes [72].

Existing studies offer a glimpse of the complexity of gene networks and developmental and biosynthetic pathways and promise a better understanding of the far-reaching consequences domestication can have on these networks and pathways [Xin *et al.*, 2013] or how domestication has preferentially recruited specific categories of genes, such as regulatory factor families [73]. A better understanding may also help reconcile the sometimes contradictory results between QTL and population and structural genomic studies, with regard to the number, magnitude and genome location of domestication genes. For example, some genes may be more prone to being selected because they occupy specific key positions in gene networks with downstream control.

Vavilov's Law of Homologous Variation and the molecular basis of phenotypic convergence

The Law of Homologous Variation [74] is based on the observation that similar phenotypes were selected during domestication of related species, raising the question whether convergent phenotypic evolution is based on convergent molecular evolution (i.e. involvement of orthologous genes). Results are contradictory. For cereal seed dispersal, the same or different genes may be involved [75–77]. By contrast, in a broad range of dicots and monocots, homologues of the *Arabidopsis TFL1* gene control flowering time and/or growth habit [54••,78–80]. Several factors appear to favor convergent evolution such as a nodal position in gene regulatory networks, simple (biosynthetic) pathways, minimal pleiotropic effect, homologous plant structures, and overall closer phylogenetic relationships [81••]. A better understanding of the molecular convergence issue may help the development of new crops.

Conclusions and a view towards the future

Genomics allows a substantial, in-depth sampling of species involved and their respective genomes to study the effect of domestication, but also provides opportunities

to address new questions. New factors affecting gene expression should be integrated into domestication studies (epigenetics [82], small RNAs [83]) as a complement to more systematic transcriptomic, proteomic, and metabolomic research. A better understanding is needed also of organism × organism relationship and how these have affected crop evolution, including host–pathogen, host–pest, crop relationships in associated cropping, and perhaps most importantly for crops, the role of farmers in molding crop biodiversity for their own specific goals [84–86]. Given the importance of wild progenitors to analyze changes observed in domestication, more attention should be paid to the organization of genetic diversity in these wild relatives and their adaptation (including phenotypic plasticity, to biotic and abiotic conditions, including Pleistocene and early Holocene conditions [87], combining biological approaches with demographic and geospatial modeling [32,88,89,90•,91]. Because of the increasing importance of wild progenitors to address climate change issues in plant breeding [92], this approach will also facilitate the development of climate-tolerant cultivars. Genomic approaches, therefore, provide a brand new way to look at 'old' problems and a platform for multidisciplinary research, combining biological and social sciences.

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- •• of outstanding interest

1. Gepts P, Famula TR, Bettinger RL, Brush SB, Damania AB, McGuire PE, Qualset CO (Eds): *Biodiversity in Agriculture: Domestication, Evolution, and Sustainability*. Cambridge, UK: Cambridge University Press; 2012.

2. Gepts P: **231. Domestication: plants**. In *Encyclopedia of Agriculture and Food Systems*. Edited by Van Alfen N. Amsterdam: Elsevier; 2014. in press.

3. Darwin C: *The Variation of Plants and Animals Under Domestication*. London: J. Murray; 1868, .

4. Ross-Ibarra J, Morrell PL, Gaut BS: **Plant domestication, a unique opportunity to identify the genetic basis of adaptation**. *Proc Natl Acad Sci* 2007, **104**:8641-8648.

5. Kelly JD, Kolkman JM, Schneider K: **Breeding for yield in dry bean (*Phaseolus vulgaris* L.)**. *Euphytica* 1998, **102**:343-356.

6. Gepts P: **Domestication as a long-term selection experiment**. *Plant Breed Rev* 2004, **24(Part 2)**:1-44.

7. Harlan JR, de Wet JMJ: **On the quality of evidence for origin and dispersal of cultivated plants**. *Curr Anthropol* 1973, **14**:51-62.

8. Zeder MA, Emshwiller E, Smith BD, Bradley DG: **Documenting domestication: the intersection of genetics and archaeology**. *Trends Genet* 2006, **22**:139-155.

9. Dempewolf H, Rieseberg L, Cronk Q: **Crop domestication in the Compositae: a family-wide trait assessment**. *Genet Resour Crop Evol* 2008, **55**:1141-1157.

10. Glémin S, Bataillon T: **A comparative view of the evolution of grasses under domestication**. *New Phytol* 2009, **183**:273-290.

11. Purugganan MD, Fuller DQ: **The nature of selection during plant domestication**. *Nature* 2009, **457**:843-848.

12. Tang H, Sezen U, Paterson AH: **Domestication and plant genomes**. *Curr Opin Plant Biol* 2010, **13**:160-166.

13. Miller AJ, Gross BL: **From forest to field: perennial fruit crop domestication**. *Am J Bot* 2011, **98**:1389-1414.

14. Meyer RS, DuVal AE, Jensen HR: **Patterns and processes in**
• **crop domestication: an historical review and quantitative analysis of 203 global food crops**. *New Phytol* 2012, **196**:29-48.
The most up-to-date general review on crop domestication available, covering a broad range of crop types, including less familiar ones. Discusses trends characterizing domestication such as life cycle, ploidy level, reproductive strategies, domestication syndrome, and spatial and temporal trends.

15. Meyer RS, Purugganan MD: **Evolution of crop species: genetics of domestication and diversification**. *Nat Rev Genet* 2013, **14**:840-852.

16. Kraft HK, Brown CH, Nabhan GP, Luedeling E, Luna Ruiz JJ,
• Coppens d'Eeckenbrugge G, Hijmans RJ, Gepts P: **Multiple lines of evidence for the origin of domesticated chili pepper, *Capsicum annuum*, in Mexico**. *Proc Natl Acad Sci* 2014. in press.
The approach presented here, based on genetic and geographic distances, allows the combination of disparate types of data to identify a putative center of origin of a crop.

17. Doebley JF, Gaut BS, Smith BD: **The molecular genetics of crop domestication**. *Cell* 2006, **127**:1309-1321.

18. Olsen KM, Wendel JF: **Crop plants as models for understanding**
• **plant adaptation and diversification**. *Front Plant Sci* 2013:4.
An excellent review discussing the molecular basis of domestication traits, including genetic targets of selection, mutational mechanisms, the relationship between genotype and phenotype through the application of other 'omics' technologies, and parallel evolution.

19. Csilléry K, Blum MGB, Gaggiotti OE, François O: **Approximate Bayesian computation (ABC) in practice**. *Trends Ecol Evol* 2010, **25**:410-418.

20. Mardis ER: **Next-generation sequencing platforms**. *Annu Rev Anal Chem* 2013, **6**:287-303.

21. Visendi P, Batley J, Edwards D: **Next generation sequencing and germplasm resources**. In *Genomics of Plant Genetic Resources Vol. 1 Managing, Sequencing and Mining Genetic Resources*. Edited by Tuberosa R, Graner A, Frison E. Springer; 2014:369-390.

22. Ekblom R, Galindo J: **Applications of next generation sequencing in molecular ecology of non-model organisms**. *Heredity* 2011, **107**:1-15.

23. Schoville SD, Bonin A, François O, Lobreaux S, Melodelima C, Manel S: **Adaptive genetic variation on the landscape: methods and cases**. *Annu Rev Ecol Evol System* 2012, **43**:23-43.

24. Manel S, Holderegger R: **Ten years of landscape genetics**. *Trends Ecol Evol* 2013, **28**:614-621.

25. Savolainen O, Lascoux M, Merila J: **Ecological genomics of local adaptation**. *Nat Rev Genet* 2013, **14**:807-820.

26. Lam H-M, Xu X, Liu X, Chen W, Yang G, Wong F-L, Li M-W, He W, Qin N, Wang B *et al.*: **Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection**. *Nat Genet* 2010, **42**:1053-1059.

27. Alo F, Furman BJ, Akhunov E, Dvorak J, Gepts P: **Leveraging genomic resources of model species for the assessment of phylogeny in wild and domesticated lentil**. *J Hered* 2011, **102**:325-329.

28. Mamidi S, Rossi M, Annam D, Moghaddam S, Lee R, Papa R, McClean P: **Investigation of the domestication of common bean (*Phaseolus vulgaris*) using multilocus sequence data**. *Funct Plant Biol* 2011, **38**:953-967.

29. He Z, Zhai W, Wen H, Tang T, Wang Y, Lu X, Greenberg AJ, Hudson RR, Wu C-I, Shi S: **Two evolutionary histories in the genome of rice: the roles of domestication genes**. *PLoS Genet* 2011, **7**:e1002100.

30. Clotault J, Thuillet AC, Buiron M, De Mita S, Couderc M,
• Haussmann BIG, Mariac C, Vigouroux Y: **Evolutionary history of pearl millet (*Pennisetum glaucum* L. R. Br.) and selection on flowering genes since its domestication**. *Mol Biol Evol* 2012, **29**:1199-1212.

This work first used Approximate Bayesian Computing to establish a most likely demographic model, which involved exponential population growth and wild-domesticated gene flow with domestication some 4800 years ago. The authors then identified 6 of 15 flowering time genes (especially those associated with circadian clock) with significant signatures of selection associated with domestication or improvement. This work is part of a long-running series to characterize the interaction between genetic diversity, the environment, and farmer management of crop diversity in the Sahel (see text).

31. Fu YB: **Population-based resequencing analysis of wild and cultivated barley revealed weak domestication signal of selection and bottleneck in the *Rrs2* scald resistance gene region**. *Genome* 2012, **55**:93-104.

32. Hufford MB, Xu X, van Heerwaarden J, Pyhajarvi T, Chia JM, Cartwright RA, Elshire RJ, Glaubitz JC, Guill KE, Kaeppler SM *et al.*: **Comparative population genomics of maize domestication and improvement**. *Nat Genet* 2012, **44** 808–U118.

33. Qi J, Liu X, Shen D, Miao H, Xie B, Li X, Zeng P, Wang S, Shang Y,
•• Gu X *et al.*: **A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity**. *Nat Genet* 2013, **45**:1510-1515.
Based on an earlier paper reporting on a reference whole-genome sequence, the authors then re-sequenced a representative sample of 115 wild and domesticated lines to conduct the full array of analyses related to domestication, including identification of major groups by STRUCTURE, documentation of the bottleneck, identification of selective sweeps and LD blocks, and mapping of domestication genes by QTL analysis and sequence polymorphisms. This paper illustrates how genomics can rapidly provide information about domestication even in 'minor' crops.

34. Bitocchi E, Bellucci E, Giardini A, Rau D, Rodriguez M, Biagetti E, Santilocchi R, Spagnoletti Zeuli P, Gioia T, Logozzo G *et al.*: **Molecular analysis of the parallel domestication of the common bean (*Phaseolus vulgaris*) in Mesoamerica and the Andes**. *New Phytol* 2013, **197**:300-313.

35. Morrell PL, Gonzales AM, Meyer KKT, Clegg MT: **Resequencing data indicate a modest effect of domestication on diversity in barley: a cultigen with multiple origins**. *J Hered* 2013, **105**:253-264.

36. Li Y-h, Zhao S-c, Ma J-x, Li D, Yan L, Li J, Qi X-t, Guo X-s, Zhang L,
• He W-m *et al.*: **Molecular footprints of domestication and improvement in soybean revealed by whole genome re-sequencing**. *BMC Genomics* 2013, **14**:579.
Documents how the footprint of domestication is much stronger for the domestication phase than the subsequent breeding phase. Identify or confirm candidate genes selected during domestication and breeding.

37. Cavanagh CR, Chao S, Wang S, Huang BE, Stephen S, Kiani S, Forrest K, Saintenac C, Brown-Guedira GL, Akhunova A *et al.*: **Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars**. *Proc Natl Acad Sci* 2013, **110**:8057-8062.

38. Soto-Cerda B, Diederichsen A, Ragupathy R, Cloutier S: **Genetic characterization of a core collection of flax (*Linum usitatissimum* L.) suitable for association mapping studies and evidence of divergent selection between fiber and linseed types**. *BMC Plant Biol* 2013, **13**:78.

39. Myles S, Boyko AR, Owens CL, Brown PJ, Grassi F, Aradhya MK,
• Prins B, Reynolds A, Chia J-M, Ware D *et al.*: **Genetic structure and domestication history of the grape**. *Proc Natl Acad Sci* 2011, **108**:3530-3535.
This article on grapes provides a contrasting view of the effect of domestication on genetic organization of the domesticated gene pool and genome. It shows limited reduction of genetic diversity upon domestication, rapid LD decay, and mostly clonal propagation or close familial relationships among varietals.

40. Iorizzo M, Senalik DA, Ellison SL, Grzebelus D, Cavagnaro PF,
• Allender C, Brunet J, Spooner DM, Van Deynze A, Simon PW: **Genetic structure and domestication of carrot (*Daucus carota* subsp. *sativus*) (Apiaceae)**. *Am J Bot* 2013, **100**:930-938.
Carrot is one of the only crops for which no reduction can be detected upon domestication. This article also identifies a likely center of domestication in Central Asia and proposes that the orange carrots are a fairly recent selection, derived from yellow carrots.

41. Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T,
•• Upadhyaya HD, Riera-Lizarazu O, Brown PJ, Acharya CB, Mitchell SE *et al.*: **Population genomic and genome-wide association studies of agroclimatic traits in sorghum**. *Proc Natl Acad Sci* 2013, **110**:453-458.
This paper illustrates the successful application of genome-wide sequence diversity scan and association mapping based on genotyping-by-sequencing of a large sample. Simultaneous SNP discovery and genotyping limits any ascertainment bias. Corroborates map position of previously mapped genes for starch enzymes, seed dispersal, plant height and maturity.

42. Xu P, Wu X, Wang B, Luo J, Liu Y, Ehlers JD, Close TJ, Roberts PA, Lu Z, Wang S *et al.*: **Genome wide linkage disequilibrium in Chinese asparagus bean (*Vigna. unguiculata* ssp. *sesquipedalis*) germplasm: implications for domestication history and genome wide association studies**. *Heredity* 2012, **109**:34-40.

43. Zizumbo-Villarreal D, Colunga-GarcíaMarín P, Payró de la Cruz E, Delgado-Valerio P, Gepts P: **Population structure and evolutionary dynamics of wild-weedy-domesticated complexes of common bean in a Mesoamerican region**. *Crop Sci* 2005, **35**:1073-1083.

44. Hufford MB, Lubinksy P, Pyhäjärvi T, Devengenzo MT,
• Ellstrand NC, Ross-Ibarra J: **The genomic signature of crop-wild introgression in maize**. *PLoS Genet* 2013, **9**:e1003477.
This article uses the established approach of analyzing paired sympatric wild and domesticated populations (e.g. Papa *et al.* [48]) to characterize gene flow between the two population types. Like Papa *et al.*, they observe that non-introgressed regions are enriched with domestication genes. The article also shows that gene flow from the wild type to domesticated maize was more ancient than in the opposite direction and may have contributed to the local adaptation of maize.

45. Olsen KM, Caicedo AL, Jia Y: **Evolutionary genomics of weedy rice in the USA**. *J Integr Plant Biol* 2007, **49**:811-816.

46. Piñeyro-Nelson A, Van Heerwaarden J, Perales HR, Serratos-Hernández JA, Rangel A, Hufford MB, Gepts P, Garay-Arroyo A, Rivera-Bustamante R, Álvarez-Buylla ER: **Transgenes in Mexican maize: molecular evidence and methodological considerations for GMO detection in landrace populations**. *Mol Ecol* 2009, **18**:750-761.

47. Papa R, Gepts P: **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* 2003, **106**:239-250.

48. Papa R, Acosta J, Delgado-Salinas A, Gepts P: **A genome-wide analysis of differentiation between wild and domesticated *Phaseolus vulgaris* from Mesoamerica**. *Theor Appl Genet* 2005, **111**:1147-1158.

49. Papa R, Bellucci E, Rossi M, Leonardi S, Rau D, Gepts P, Nanni L, Attene G: **Tagging the signatures of domestication in common bean (*Phaseolus vulgaris*) by means of pooled DNA samples**. *Ann Bot* 2007, **100**:1039-1051.

50. Russell J, Dawson IK, Flavell AJ, Steffenson B, Weltzien E, Booth A, Ceccarelli S, Grando S, Waugh R: **Analysis of >1000 single nucleotide polymorphisms in geographically matched samples of landrace and wild barley indicates secondary contact and chromosome-level differences in diversity around domestication genes**. *New Phytol* 2011, **191**:564-578.

51. van Heerwaarden J, Doebley J, Briggs WH, Glaubitz JC, Goodman MM, Sánchez González JdJ, Ross-Ibarra J: **Genetic signals of origin, spread, and introgression in a large sample of maize landraces**. *Proc Natl Acad Sci* 2011, **108**:1088-1092.

52. McCouch SR, Kovach MJ, Sweeney M, Jiang H, Semon M: **The dynamics of rice domestication: a balance between gene flow and genetic isolation**. In *Biodiversity in Agriculture: Domestication, Evolution and Sustainability*. Edited by Gepts P, Famula TR, Bettinger RL, Brush SB, Damania AB, McGuire PE, Qualset CO. Cambridge University Press; 2012:311-329.

53. Repinski SL, Kwak M, Gepts P: **The common bean growth habit gene PvTFL1γ is a functional homolog of Arabidopsis TFL1**. *Theor Appl Genet* 2012, **124**:1539-1547.

54. Kwak M, Toro O, Debouck D, Gepts P: **Multiple origins of the
•• determinate growth habit in domesticated common bean (*Phaseolus vulgaris* L.)**. *Ann Bot* 2012, **110**:1573-1580.

This paper illustrates the high level of diversity at a single locus, *PvTFL1y*, a homologue of the *Arabidopsis TFL1* locus, responsible for the determinate growth habit. At least five different mutation types are documented leading to determinacy, as a result of farmer selection in different areas of the center of domestication, consistent with the importance of earliness in subsistence agriculture.

55. Asfaw A, Almekinders CJM, Blair MW, Struik PC: **Participatory approach in common bean (*Phaseolus vulgaris* L.) breeding for drought tolerance for southern Ethiopia**. *Plant Breed* 2012, **131**:125-134.

56. Worthington M, Soleri D, Aragón-Cuevas F, Gepts P: **Genetic composition and spatial distribution of farmer-managed bean plantings: an example from a village in Oaxaca, Mexico**. *Crop Sci* 2012, **52**:1721-1735.

57. van Heerwaarden J, Hufford MB, Ross-Ibarra J: **Historical**
• **genomics of North American maize**. *Proc Natl Acad Sci* 2012, **109**:12420-12425.
This article documents the effect of modern breeding on the organization of the maize genome (e.g. LD) and on changes in genetic diversity (e.g. appearance of heterotic groups).

58. Barrett RDH, Schluter D: **Adaptation from standing genetic variation**. *Trends Ecol Evol* 2008, **23**:38-44.

59. Peter BM, Huerta-Sanchez E, Nielsen R: **Distinguishing between selective sweeps from standing variation and from a de novo mutation**. *PLoS Genet* 2012, **8**:e1003011.

60. Fu W, Akey JM: **Selection and adaptation in the human genome**. *Annu Rev Genomics Human Genet* 2013, **14**:467-489.

61. Durand E, Tenaillon M, Ridel C, Coubriche D, Jamin P, Jouanne S, Ressayre A, Charcosset A, Dillmann C: **Standing variation and new mutations both contribute to a fast response to selection for flowering time in maize inbreds**. *BMC Evol Biol* 2010, **10**:2.

62. Studer A, Zhao Q, Ross-Ibarra J, Doebley J: **Identification of a**
• **functional transposon insertion in the maize domestication gene tb1**. *Nat Genet* 2011, **43**:1160-1163.
This article provides a clear example of the contribution of standing variation in wild ancestral populations to the mutation pool selected during domestication, based on the identification of a causal mutation and the timing of this mutation well before domestication.

63. Purugganan MD, Fuller DQ: **Archaeological data reveal slow rates of evolution during plant domestication**. *Evolution* 2011, **65**:171-183.

64. Jaenicke-Després V, Buckler ES, Smith BD, Gilbert MTP, Cooper A, Doebley J, Pääbo S: **Early allelic selection in maize as revealed by ancient DNA**. *Science* 2003, **302**:1206-1208.

65. Roullier C, Benoit L, McKey DB, Lebot V: **Historical collections**
• **reveal patterns of diffusion of sweet potato in Oceania obscured by modern plant movements and recombination**. *Proc Natl Acad Sci* 2013, **110**:2205-2210.
The research illustrates how historic records, in this case, herbarium samples from the 18th century coincident with European discovery voyages, can document pre-Columbian patterns of dispersal of sweet potato providing evidence of pre-European contacts between South America and the Pacific islands.

66. Mariac C, Jehin L, Saidou AA, Thuillet AC, Couderc M, Sire P, Jugde H, Adam H, Bezancon G, Pham JL *et al.*: **Genetic basis of pearl millet adaptation along an environmental gradient investigated by a combination of genome scan and association mapping**. *Mol Ecol* 2011, **20**:80-91.

67. Saïdou AA, Mariac C, Luong V, Pham JL, Bezancon G, Vigouroux Y: **Association studies identify natural variation at PHYC linked to flowering time and morphological variation in pearl millet**. *Genetics* 2009, **182**:899-910.

68. Lakis G, Ousmane AM, Sanoussi D, Habibou A, Badamassi M, Lamy F, Jika N, Sidikou R, Adam T, Sarr A *et al.*: **Evolutionary dynamics of cycle length in pearl millet: the role of farmer's practices and gene flow**. *Genetica* 2011, **139**:1367-1380.

69. Lakis G, Navascués M, Rekima S, Simon M, Remigereau M-S, Leveugle M, Takvorian N, Lamy F, Depaulis F, Robert T: **Evolution of neutral and flowering genes along pearl millet (*Pennisetum glaucum*) domestication**. *PLoS One* 2012, **7**:e36642.

70. Swanson-Wagner R, Briskine R, Schaefer R, Hufford MB, Ross-Ibarra J, Myers CL, Tiffin P, Springer NM: **Reshaping of the maize transcriptome by domestication**. *Proc Natl Acad Sci* 2012, **109**:11878-11883.

71. Koenig D, Jiménez-Gómez JM, Kimura S, Fulop D, Chitwood DH, Headland LR, Kumar R, Covington MF, Devisetty UK, Tat AV *et al.*: **Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato**. *Proc Natl Acad Sci* 2013, **110**:E2655-E2662.

72. Hu G, Koh J, Yoo M-J, Grupp K, Chen S, Wendel JF: **Proteomic profiling of developing cotton fibers from wild and domesticated Gossypium barbadense**. *New Phytol* 2013, **200**:570-582;
Xin H, Zhang J, Zhu W, Wang N, Fang P, Han Y, Ming R, Li S: **The effects of artificial selection on sugar metabolism and transporter genes in grape**. *Tree Genet Genom* 2013, **9**:1343-1349.

73. Huang Z, Houten J, Gonzalez G, Xiao H, Knaap E: **Genome-wide identification, phylogeny and expression analysis of SUN, OFP and YABBY gene family in tomato**. *Mol Genet Genomics* 2013, **288**:111-129.

74. Vavilov NI: **The law of homologous series in variation**. *J Genet* 1922, **12**:47-89.

75. Li W, Gill BS: **Multiple genetic pathways for seed shattering in the grasses**. *Funct Integrative Genomics* 2006, **6**:300-309.

76. Tang H, Cuevas HE, Das S, Sezen UU, Zhou C, Guo H, Goff VH, Ge Z, Clemente TE, Paterson AH: **Seed shattering in a wild sorghum is conferred by a locus unrelated to domestication**. *Proc Natl Acad Sci* 2013, **110**:15824-15829.

77. Lin Z, Li X, Shannon LM, Yeh C-T, Wang ML, Bai G, Peng Z, Li J, Trick HN, Clemente TE *et al.*: **Parallel domestication of the *Shattering1* genes in cereals**. *Nat Genet* 2012, **44**:720-724.

78. Pnueli L, Carmel-Goren L, Hareven D, Gutfinger T, Alvarez J, Ganal M, Zamir D, Lifschitz E: **The *SELF-PRUNING* gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of *CEN* and *TFL1***. *Development* 1998, **125**:1979-1989.

79. Boss PK, Sreekantan L, Thomas MR: **A grapevine TFL1 homologue can delay flowering and alter floral development when overexpressed in heterologous species**. *Funct Plant Biol* 2006, **33**:31-41.

80. Comadran J, Kilian B, Russell J, Ramsay L, Stein N, Ganal M, Shaw P, Bayer M, Thomas W, Marshall D *et al.*: **Natural variation in a homolog of Antirrhinum CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley**. *Nat Genet* 2012, **44**:1388-1392.

81. Lenser T, Theißen G: **Molecular mechanisms involved in**
•• **convergent crop domestication**. *Trends Plant Sci* 2013, **18**:704-714.
A thoughtful discussion on the factors that influence the appearance or not of convergence in crop domestication, like pleiotropy, standing variation, and morphological relationships. It points to the need to understand the position of domestication genes in overall gene networks in plants. This article may have an impact on ongoing efforts to domesticate new crops.

82. Diez CM, Roessler K, Gaut BS: **Epigenetics and plant genome evolution**. *Curr Opin Plant Biol* 2014, **18**:1-8.

83. Wang Y, Bai XF, Yan CH, Gui YJ, Wei XH, Zhu QH, Guo LBA, Fan LJ: **Genomic dissection of small RNAs in wild rice (*Oryza rufipogon*): lessons for rice domestication**. *New Phytol* 2012, **196**:914-925.

84. Raffaele S, Farrer RA, Cano LM, Studholme DJ, MacLean D, Thines M, Jiang RHY, Zody MC, Kunjeti SG, Donofrio NM *et al.*: **Genome evolution following host jumps in the Irish Potato Famine pathogen lineage**. *Science* 2010, **330**:1540-1543.

85. Heil M: **Indirect defence via tritrophic interactions**. *New Phytol* 2008, **178**:41-61.

86. Zimmerer KS: **The compatibility of agricultural intensification in a global hotspot of smallholder agrobiodiversity (Bolivia)**. *Proc Natl Acad Sci* 2013, **110**:2769-2774.

87. Piperno DR, Holst I, Winter K, McMillan O: **Teosinte before domestication: experimental study of growth and phenotypic variability in late Pleistocene and early Holocene environments**. *Quatern Int* 2014. in press, http://dx.doi.org/10.1016/j.quaint.2013.12.049.

88. Pyhäjärvi T, Hufford MB, Mezmouk S, Ross-Ibarra J: **Complex patterns of local adaptation in teosinte**. *Genome Biol Evol* 2013, **5**:1594-1609.

89. Besnard G, Khadari B, Navascués M, Fernández-Mazuecos M, El Bakkali A, Arrigo N, Baali-Cherif D, Brunini-Bronzini de Caraffa V, Santoni S, Vargas P et al.: **The complex history of the olive tree: from Late Quaternary diversification of Mediterranean lineages to primary domestication in the northern Levant**. *Proc R Soc B: Biol Sci* 2013:280.

90. Mamidi S, Rossi M, Moghaddam SM, Annam D, Lee R, Papa R,
•   McClean PE: **Demographic factors shaped diversity in the two gene pools of wild common bean *Phaseolus vulgaris* L.**. *Heredity* 2013, **110**:267-276.
This analysis of demographic parameters in a wild progenitor before domestication helps understand the subsequent divergence through two independent domestications of geographically distinct gene pools and their respective levels of diversity. It also provides a date for the divergence between the two gene pools.

91. van Etten J, Hijmans RJ: **A geospatial modelling approach integrating archaeobotany and genetics to trace the origin and dispersal of domesticated plants**. *PLoS ONE* 2010, **5**:e12060.

92. McCouch S, Baute GJ, Bradeen J, Bramel P, Bretting PK, Buckler E, Burke JM, Charest D, Cloutier S, Cole G et al.: **Agriculture: feeding the future**. *Nature* 2013, **499**:23-24.