ELSEVIER

Domestication and plant genomes

Haibao Tang, Uzay Sezen and Andrew H Paterson

The techniques of plant improvement have been evolving with the advancement of technology, progressing from crop domestication by Neolithic humans to scientific plant breeding, and now including DNA-based genotyping and genetic engineering. Archeological findings have shown that early human ancestors often unintentionally selected for and finally fixed a few major domestication traits over time. Recent advancement of molecular and genomic tools has enabled scientists to pinpoint changes to specific chromosomal regions and genetic loci that are responsible for dramatic morphological and other transitions that distinguish crops from their wild progenitors. Extensive studies in a multitude of additional crop species, facilitated by rapid progress in sequencing and resequencing(s) of crop genomes, will further our understanding of the genomic impact from both the unusual population history of cultivated plants and millennia of human selection.

Address
Plant Genome Mapping Laboratory, University of Georgia, Athens, GA 30602, USA

Corresponding author: Paterson, Andrew H (paterson@uga.edu)

Introduction

Prehistoric people were able to transform a few wild plant species into the crops that now provide humanity with a wide range of ecosystem services. Compared to their wild relatives, domesticated crops typically show synchronization of flowering time, enlargement of reproductive organs (fruit or seed), loss of seed dispersal, increased apical dominance, and other characteristics collectively known as the 'domestication syndrome' [1]. In this review, we first discuss new findings on the history of agriculture and then describe population and genomic changes underlying the domestication process. We then recapitulate some recurring strategies for dissecting and quantifying these changes with a focus on recent developments.

History of agriculture and domestication

Extensive literature regarding the fascinating history of agriculture and subsequent domestication of wild organisms has recently reached toward a new level of integration, through the synthesis of archeological and genetic findings [2,3••]. According to this new synthesis, the transition into agriculture took multiple steps and domestication appears to have been a series of nondeliberate attempts spread over a few thousand years rather than a few human generations [4•].

Clues as to what prompted the early hunter-gatherers to alter their lifestyle may reside in climate data. Globally distributed oxygen isotope measurements from ocean sediments allow reconstruction of the earth's climate back to 5 million years ago [5]. Ice core data from Greenland and Antarctica have a resolution of a decade and goes back to 800 000 years ago, covering 20 glaciations [6]. Modern humans evolved in the past 100 000 years through these severe climatic perturbations. Our omnivorous ancestors appear to have switched to a plant-based diet during the cold spells. Compelling evidence comes from a 23 000-year-old Ohalo II cave in Israel, which contained ground stone tools and a wide collection of plant and animal remains left by hunter-gatherers that survived the last glacial maximum [7]. Humans must have had an excellent understanding of plant management including seasonality and the life cycles of animals and plants, and the cognitive prerequisites required for agriculture existed perhaps as early as 50 000 years ago.

Agriculture was not practiced until the climatically stable Holocene warming. During the most recent glaciation there was a warm period sandwiched between the Oldest Dryas (18 000–14 600 bp) and Younger Dryas (12 900–11 500 bp) cold periods. This warm period may have temporarily rescued the profitable hunter-gatherer lifestyle, delaying the emergence of agriculture. Human populations during the Younger Dryas reduced their prey size, became more sedentary, and once again increased the plant content of their diet [7]. Cold arid conditions promoted the use of grasses and cereal seeds by hunter-gatherers. Granaries discovered at the site of Dhra' in the Jordan Valley are a striking demonstration that preagricultural plant management was already present 11 000 years ago [8•].

Surprisingly, the characteristics known as the domestication syndrome were slow to develop. One hallmark of domestication is the reduction of seed shattering which shows a gradual increase in frequency over a few thousand years [3••]. People had many ways to maximize harvests, despite shattering. Ethnographic evidence suggests that

Figure 1
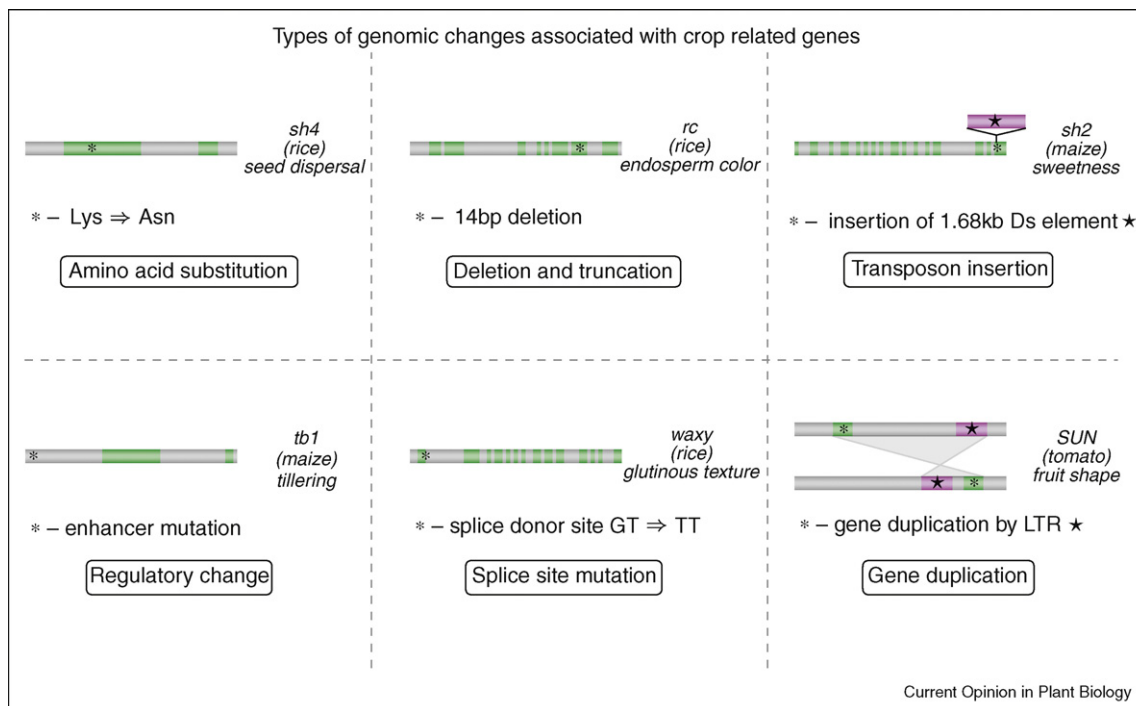
Types of changes associated with crop-related genes. One specific example is given for each type of genomic change (sh4 [16], rc [18], sh2 [65], tb1 [17], waxy [66], and SUN [15•]). More comprehensive gene lists and associated changes can be found in [13,19].

hand gathering from the ground, unripe harvesting, beating, or hand stripping of ripe seeds into a basket, and twisting the plant stems to form a cone all provided sufficient yields and could have delayed the rise of non-shattering forms [9,10]. Such practices may have minimized selection on the shattering trait in some human cultures, even though simulations suggest that selection could still have allowed fixation of nonshattering within a period of 100 years [11•]. Recent archeobotanical findings at Tianluoshan, China provide an excellent example for the slow rate of increase in nonshattering rice frequency [12••].

Genomic changes associated with the domestication

Several genes that were targets of domestication or crop improvement have been identified [13], influencing fruit size [14] and shape [15•], seed dispersal [16], tillering [17], seed color [18], and many other traits (reviewed in [13,19]). The mutations in these genes that were incorporated into crops most likely occurred naturally in the progenitor plant population. For example, the nonfunctional (domesticated) allele of the rice shattering gene sh4 is also found in some individuals of the progenitor species O. rufipogon [20]. It is unclear whether the same mutation is somehow maintained in low frequency in the wild population, or instead went extinct in the wild but later introgressed from cultivated individuals.

Such mutated alleles conferring domestication-related traits are presumed to reduce reproductive fitness of wild individuals bearing the alleles, and therefore to occur in wild populations only at relatively low frequencies. The persistence of 'domestication alleles' in natural populations is easily explained for those that are recessive and may represent loss-of-function alleles, but more difficult to explain for the appreciable number that are additive or dominant, and which have modified function [13,21].

The form and nature of the genetic mutations associated with transitions from wild to domesticated plants is highly variable (Figure 1). Some mutations causing frameshifts or premature termination of the protein product are considered nonfunctional 'knock-outs' in the domesticated species. By contrast, some mutations are in the regulatory elements and modify spatio-temporal expression patterns and/or levels [13]. In the case of sh4 [16], amino acid substitutions appear to disrupt the interaction of the protein with downstream targets.

Population footprints of crop domestication and improvement

Study of both mutations in crop-related genes and neutral DNA markers is changing our understanding of the tempo and duration of domestication. The spread of 'domestication alleles' through populations is typically

studied by collecting genetic information from a diverse sampling of both domesticated and wild plant varieties. One common feature of the domesticated genomes is the reduction of genetic diversity in crops relative to the wild progenitors [16]. This reduction has two causes. First, domestication is typically thought to have involved initial populations of small size, constraining *genome-wide* genetic diversity by the familiar concept of a 'genetic bottleneck'. The second factor is the 'selective sweep', or directional selection for *local* genomic regions that distinguish crops from their ancestors. Both forces can be tested for deviations from the neutral Wright–Fisher model, which assumes constant population size and no selection [22].

Demographic changes and tempo of domestication

The development of neutral DNA markers has made it possible to study the demographic changes associated with domestication independently of the effects of selection on specific genomic regions [23]. Although conventional wisdom was that crops lost a significant portion of ancestral diversity, recent sequence data suggest otherwise. For example, domesticated maize has about 60–80% of the diversity in its progenitor teosinte [24,25]. Surprisingly, this estimate is typical of several crops, including einkorn wheat (70–100%) [26], sorghum (~80%) [27] and chile peppers (~90%) [28]. However, such estimates are likely overestimates, given the potential selection bias and possible genetic erosion in the wild population. It is also possible that some crops have partially restored diversity through recent gene flow from wild population after the initial domestication [29].

Population bottleneck is usually quantified by two factors — the bottleneck population size (Nb) and duration of the bottleneck (d). The severity of the bottleneck is given by coefficient $k = Nb/d$ [25]. Results from earlier analysis [30] suggested that most domesticates form a monophyletic group consistent with a single, rapid localized domestication event. Recent archeobotanical evidence and coalescent simulations instead favor a 'protracted' model of domestication [4•,31•]. For example, the classical population model in maize suggested only one bottleneck with a single value of k [25], whereas the 'protracted' model fits several values of k, representing multiple bottlenecks of different strengths. A prolonged series of bottlenecks associated with domestication might reflect the dispersal of key genotypes, and further improvements by plant breeding [32]. Frequency spectra of allele variants also reveal unique demographic history of particular domesticated species. For example, an excess of high-frequency SNP alleles suggests a rather complicated breeding history of rice [33].

Effect of artificial selection and natural selection

Detailed analyses of domestication genes reveal remarkable reduction of diversity that drove only a few hap-lotypes into fixation. Selection (both artificial and natural) is expected to reduce diversity at domestication-related genes and tightly linked loci (selective sweep) as favorable alleles are driven to high frequency. Such reduction of diversity is expected to be more striking than the effects of genetic bottleneck alone. The size and shape of the selective sweeps depend on the time and strength of selection as well as local recombination rates in the genome. Several studies in maize reported particularly large sweep blocks [34•,35]. However, in some domesticated species like sorghum, lower levels of variations at neutral loci make it more difficult to detect selection, perhaps because of much more recent domestication [36].

Artificial selection can unintentionally impose genetic load on a crop genome by harboring subtle but deleterious mutations that are closely linked to major domestication genes and preserved by 'genetic hitchhiking', rather than being purged as they would in natural populations. Recent genome comparisons of two rice cultivars (*japonica* and *indica*) show a high level of deleterious mutations, suggesting a genome-wide relaxation of selective constraints because of domestication [37], consistent with findings in domesticated animals [38]. One could envision that the resulting genetic load might be especially heavy in heterochromatic regions of a genome that are recalcitrant to recombination [39,40].

Recombination of favorable alleles through hybridization and breeding

The history of domestication and breeding can also be revealed by tracing the distribution of major domestication genes in chronologically and geographically stratified landraces and cultivars. For example, among the six rice domestication-related genes identified, spread of the mutations in *Rc* and *qSW5* was probably the most ancient since these mutations can be readily found in most heritage landraces, while *qSH1* was relatively recent, only found in a few modern temperate *japonica* cultivars [19].

Assessing domestication-related genes and genomic regions

Quantitative trait locus (QTL) mapping is a powerful way to study the domestication-related genes and chromosomal regions, with only the requirement of the presence of both domesticated and nondomesticated alleles in the mapping population [41]. QTLs with simple genetics and large phenotypic effect can be identified easily, and indeed most of the domestication genes identified so far are from this category. However, QTL studies are dependent on the environment and the parental lines, and are time and labor intensive to conduct.

A different approach — 'selection scan' takes advantage of the unusual polymorphism patterns of domestication-

related genes [25,42•]. In this method, one looks for loci that show significant reduction of sequence diversity in domesticated genotypes compared to wild samples. The selection scan is a relatively high-throughput method, often identifying a large set of 'candidate genes' compared to conventional QTL mapping which interrogates only a few loci at a time. It was previously thought that only a few key genes of 'large effect' were responsible to transform wild teosinte into maize [43]. However, a study based on 774 genes extrapolates that 2–4% of the maize genes are under domestication-related selection [25]. Another study found 36 of 492 (7%) sunflower genes to show evidence of selection [42•]. The number of genes under selection may vary among different domesticated species, because of differences in the domestication history and extent of recombination.

Tests for convergent evolution of domestication genes

Early analysis of a few domestication-related QTLs suggested that they occur in corresponding map locations across different cereal species, more often than explained by chance [44]. It was also postulated that convergent phenotypic evolution of major cereal crops might be explained by independent selection of mutations in orthologous gene loci [44]. It now appears that different domestication traits have different patterns of genetic architecture. For example, maize flowering time QTLs show synteny conservation with rice [44,45]. In contrast, the numbers and locations of many loci controlling seed shattering differ in barley, maize, rice, and sorghum, indicating multiple genetic pathways [46]. Even when the genetic control of a particular trait is well conserved at the positional level, it remains to be proven whether corresponding genes are responsible [47,48], although one can point to many examples of corresponding genetic control of discrete phenotypes, in some cases even across monocots and dicots [49]. For other traits, related species might have different morphological or phylogenetic constraints and therefore the major genetic determinants vary. For example, *ramosa1* controls the floral branching system in the panicoids (maize, *Miscanthus*, and sorghum) but not in rice [50].

To date, a conspicuously high frequency of the domestication genes identified is transcriptional regulators [13], yet their downstream targets are still unknown. These can be studied through a variety of experimental and computational methods, including ChIP-sequencing [51], yeast two hybrid screens [52], and genome-wide expression QTLs (eQTLs), which query the potential interactions among many genes [48]. Additionally, we can study the set of genes that are differentially expressed in the domesticated versus the wild individuals and see how these genes are related to one another in the context

of a large regulatory network [48]. Loss-of-function mutations (as opposed to only regulatory changes) can also be examined by comparing the genomes of different related crop species. For example, pseudogenes in the corresponding chromosomal locations that have simultaneously experienced loss-of-function mutations in rice and sorghum might reveal potential targets of domestication that perhaps resulted from recent convergent changes [53•].

New avenues for studying domestication
High-throughput genotyping technology
Sequencing technology is becoming increasingly parallel and high-throughput while cost per base continues to plummet. High density tiling arrays and next generation sequencing provide for efficient sampling of genetic diversity. Using resequencing microarrays to map genome-wide SNP variations, a recent study revealed phylogenetic relationships, population structure and introgression history among 20 rice cultivars and landraces [54•]. Another proof-of-concept study used short-read sequencing technology to map the 'Green Revolution' gene *sd1* [49] in a 160-individual recombination inbred population [55••]. Resequencing is becoming a cost-effective genotyping method so long as data processing can be handled efficiently. Although the current read length of next generation sequencing is still not ideal for *de novo* sequencing of large and repetitive crop genomes, steady progress is being made [56].

Diversity-based mapping populations
Typical plant populations for genetic mapping of quantitative traits are of two types — naturally occurring lines and synthetic lines [41]. The advantage of synthetic populations, such as F2 or backcross, is that their strong linkage disequilibrium affords high power to detect QTL. However, because of limited allelic variation and number of recombinations, the map resolution of synthetic populations is low. This is complemented by using naturally occurring populations (or a diverse collection of germplasm), in which there are many historical recombinations. However, interpretations of correlations between genotype and traits using naturally occurring populations are often complicated by population structure and unknown genetic relatedness among individuals, producing false associations, and comprehensive scanning of the genome requires far greater numbers of genotypes to be assayed than QTL mapping.

Efforts that combine the relative strengths of the two types of populations are made feasible by advances in genotyping efficiency. For example, Buckler *et al.* recently established a nested association mapping (NAM) population in maize, where 25 different maize lines were all crossed with the same parent B73, and 200

recombinant inbred lines were generated from each cross [57●●]. This composite population captures a significant fraction of maize diversity and has been used to identify numerous QTLs that are shared among different families [57●●]. This is in contrast to the classical family of only two parents, in which only subsets of QTLs are detected. Development of 'multiparent advanced generation inter-cross' (MAGIC) offers a similar experimental platform for analyzing gene–trait correlations [58]. Such mapping strategies will provide a clearer picture of the genetic architecture underlying many domestication-related traits.

The next wave of domestications

Although often thought of as a Neolithic activity, dom-estication is an ongoing process driven by changing human needs and agricultural conditions. Expansion of agriculture to provide plant biomass for production of fuels and/or feedstocks will require additions to our pre-sent repertoire of crops. Scientific breeding of leading candidates such as *Miscanthus* [59] is only beginning, and many early priorities are 'domestication traits' about which there exists much information in close relatives such as sorghum and sugarcane [44,60–62]. The accumu-lating knowledge of QTLs and genes associated with domestication of other grasses may be useful in biofuel crops, albeit in different ways — for example, to suppress flowering rather than accelerate it, and to increase height rather than reduce it.

Likewise, growing attention to the needs of Africa is reawakening the merits of further improving native plants already reasonably well suited to local cultivation, and which have been subjected to degrees of domestication ranging from none (Moringa) to advanced (okra) [63]. The conserved synteny and collinearity within most clades of flowering plants [64], together with growing information about genes and their functions in botanical models and major crops, provide means by which to accelerate pro-gress in improvement of such plants to better suit human needs.

Conclusions and future prospects

Plant domestication, breeding, and biotechnology have modified plant genomes to tailor crops to the needs of humanity with increasing efficiency and precision. Un-derstanding such processes, crop domestication in particular, is crucial today because of the rising demand for improving yield and quality of grain crops, as well as a new wave of domestications to obtain additional ecosys-tem services from plants. Crops also form a particularly good system for the study of accelerated evolution. The study of domestication intersects both archeology and genetics, and informs us about the nature of both gene functions and selective constraints. Further knowledge of the genomics underlying crop domestication facilitates advancement of evolutionary theory while offering a solid foundation for full-fledged crop engineering in the near future.

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. Hammer K: **Das Domestikationssyndrom**. *Kulturpflanze* 1984, **32**:11-34.

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

3. Purugganan MD, Fuller DQ: **The nature of selection during plant**
●● **domestication**. *Nature* 2009, **457**:843-848.
In this review, genes associated with domestication and their gradual increase in frequency are superimposed on archeobotanical data. The authors utilize rice as an example for postdomestication cultural selective pressures that play a role in crop diversification.

4. Brown TA, Jones MK, Powell W, Allaby RG: **The complex origins**
● **of domesticated crops in the fertile crescent**. *Trends Ecol Evol* 2009, **24**:103-109.
Genetic evidence until recently had supported a 'rapid transition' model of domestication. The synthesis of archeological and genetic data leading to a new model of multiregional agriculture and protracted domestication is reviewed.

5. Lisiecki LE, Raymo ME: **A Pliocene–Pleistocene stack of 57 globally distributed benthic** δ' **^{18}O records**. *Paleoceanography* 2005, **20**: PA1003, 1–17.

6. Luthi D, Le Floch M, Bereiter B, Blunier T, Barnola J-M, Siegenthaler U, Raynaud D, Jouzel J, Fischer H, Kawamura K *et al.*: **High-resolution carbon dioxide concentration record 650,000–800,000 years before present**. *Nature* 2008, **453**:379-382.

7. Weiss E, Wetterstrom W, Nadel D, Bar-Yosef O: **The broad spectrum revisited: evidence from plant remains**. *Proc Natl Acad Sci U S A* 2004, **101**:9551-9555.

8. Kuijt I, Finlayson B: **Evidence for food storage and**
● **predomestication granaries 11,000 years ago in the Jordan Valley**. *Proc Natl Acad Sci U S A* 2009, **106**:10966-10970.
The discovery of highly sophisticated granaries built with suspended floors provides compelling evidence for intensive plant use even before domestication.

9. Kislev ME, Weiss E, Hartmann A: **Impetus for sowing and the beginning of agriculture: ground collecting of wild cereals**. *Proc Natl Acad Sci U S A* 2004, **101**:2692-2695.

10. Vaughan DA, Balazs E, Heslop-Harrison JS: **From crop domestication to super-domestication**. *Ann Bot (Lond)* 2007, **100**:893-901.

11. Zhang LB, Zhu Q, Wu ZQ, Ross-Ibarra J, Gaut BS, Ge S, Sang T:
● **Selection on grain shattering genes and rates of rice domestication**. *New Phytol* 2009, **184**:708-720.
This work studied the level of DNA variation at two genetic loci *sh4* and *qSH1* that are responsible for grain shattering, in both cultivated rice and wild progenitors. The *sh4* allele was fixed in all rice cultivars and neutrality tests and coalescent simulations suggest that selection on *sh4* could be strong enough to allow rapid fixation.

12. Fuller DQ, Qin L, Zheng Y, Zhao Z, Chen X, Hosoya LA, Sun GP:
●● **The domestication process and domestication rate in rice: spikelet bases from the Lower Yangtze**. *Science* 2009, **323**:1607-1610.

Using archeological data from a Neolithic site in China, the gradual increase in domesticated nonshattering rice is demonstrated. Previously, genetic models supported the hypothesis that under strong selection nonshattering domesticates could arise within a hundred years (Ref. [11●]).

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

14. Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD: **fw2.2: a quantitative trait locus key to the evolution of tomato fruit size**. *Science* 2000, **289**:85-88.

15. Xiao H, Jiang N, Schaffner E, Stockinger EJ, van der Knaap E: **A**
●   **retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit**. *Science* 2008, **319**:1527-1530.
This work identified a major gene *SUN* that controls elongated fruit shape of tomato, resulting from LTR retrotransposon-mediated gene duplication. This study illustrated how gene duplication, in addition to base pair or amino acid changes, can play an important role in creating morphological diversity.

16. Li C, Zhou A, Sang T: **Rice domestication by reducing shattering**. *Science* 2006, **311**:1936-1939.

17. Wang RL, Stec A, Hey J, Lukens L, Doebley J: **The limits of selection during maize domestication**. *Nature* 1999, **398**:236-239.

18. Sweeney MT, Thomson MJ, Pfeil BE, McCouch S: **Caught red-handed: Rc encodes a basic helix–loop–helix protein conditioning red pericarp in rice**. *Plant Cell* 2006, **18**:283-294.

19. Izawa T, Konishi S, Shomura A, Yano M: **DNA changes tell us about rice domestication**. *Curr Opin Plant Biol* 2009, **12**:185-192.

20. Lin Z, Griffith ME, Li X, Zhu Z, Tan L, Fu Y, Zhang W, Wang X, Xie D, Sun C: **Origin of seed shattering in rice (*Oryza sativa* L.)**. *Planta* 2007, **226**:11-20.

21. Lin Y-R, Zhu L, Ren S, Yang J, Schertz KF, Paterson AH: **A *Sorghum propinquum* BAC library, suitable for cloning genes associated with loss-of-function mutations during crop domestication**. *Mol Breed* 1999, **5**:511-520.

22. Yamasaki M, Wright SI, McMullen MD: **Genomic screening for artificial selection during domestication and improvement in maize**. *Ann Bot (Lond)* 2007, **100**:967-973.

23. Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA: **Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa***. *Proc Natl Acad Sci U S A* 2006, **103**:9578-9583.

24. Tenaillon MI, U'Ren J, Tenaillon O, Gaut BS: **Selection versus demography: a multilocus investigation of the domestication process in maize**. *Mol Biol Evol* 2004, **21**:1214-1225.

25. Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD, Gaut BS: **The effects of artificial selection on the maize genome**. *Science* 2005, **308**:1310-1314.

26. Kilian B, Ozkan H, Walther A, Kohl J, Dagan T, Salamini F, Martin W: **Molecular diversity at 18 loci in 321 wild and 92 domesticate lines reveal no reduction of nucleotide diversity during *Triticum monococcum* (Einkorn) domestication: implications for the origin of agriculture**. *Mol Biol Evol* 2007, **24**:2657-2668.

27. Casa AM, Mitchell SE, Hamblin MT, Sun H, Bowers JE, Paterson AH, Aquadro CF, Kresovich S: **Diversity and selection in sorghum: simultaneous analyses using simple sequence repeats**. *Theor Appl Genet* 2005, **111**:23-30.

28. Aguilar-Melendez A, Morrell PL, Roose ML, Kim S-C: **Genetic diversity and structure in semiwild and domesticated chiles (*Capsicum annuum*; Solanaceae) from Mexico**. *Am J Bot* 2009, **96**:1190-1202.

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

30. Eyre-Walker A, Gaut RL, Hilton H, Feldman DL, Gaut BS: **Investigation of the bottleneck leading to the domestication of maize**. *Proc Natl Acad Sci U S A* 1998, **95**:4441-4446.

31. Allaby RG, Fuller DQ, Brown TA: **The genetic expectations of a**
●   **protracted model for the origins of domesticated crops**. *Proc Natl Acad Sci U S A* 2008, **105**:13982-13986.
This study investigated whether multilocus, neutral markers could lead to the inference of monophyly, if a true domestication actually involved two independent events. Through extensive modeling and simulations, they found that the signatures of multiple independent domestications can frequently be lost through genetic drift that occurred during a 'protracted' bottleneck. This resolves previous inconsistencies in certain crops, where genetic data suggested single origins while archeological evidence suggested multiple origins of domesticates.

32. Chrispeels M, Sadava D: *Plants, Genes, and Crop Biotechnology*. Sudbury: Jones and Bartlett Publishers; 2003.

33. Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, York TL, Polato NR, Olsen KM, Nielsen R, McCouch SR *et al.*: **Genome-wide patterns of nucleotide polymorphism in domesticated rice**. *PLoS Genet* 2007, **3**:1745-1756.

34. Tian F, Stevens NM, Buckler ES: **Tracking footprints of maize**
●   **domestication and evidence for a massive selective sweep on chromosome 10**. *Proc Natl Acad Sci U S A* 2009, **106(Suppl 1)**:9979-9986.
This work discovered a striking case of selective sweep on maize chromosome 10, covering 1.1 Mb and more than 15 genes, and only retaining 3.6% of the silent site diversity in maize relative to teosinte. The authors suggested that positive selection for domestication traits and other adaptive traits are perhaps responsible for shaping this unusual pattern.

35. Palaisa K, Morgante M, Tingey S, Rafalski A: **Long-range patterns of diversity and linkage disequilibrium surrounding the maize Y1 gene are indicative of an asymmetric selective sweep**. *Proc Natl Acad Sci U S A* 2004, **101**:9885-9890.

36. Hamblin MT, Casa AM, Sun H, Murray SC, Paterson AH, Aquadro CF, Kresovich S: **Challenges of detecting directional selection after a bottleneck: lessons from *Sorghum* bicolor**. *Genetics* 2006, **173**:953-964.

37. Lu J, Tang T, Tang H, Huang J, Shi S, Wu CI: **The accumulation of deleterious mutations in rice genomes: a hypothesis on the cost of domestication**. *Trends Genet* 2006, **22**:126-131.

38. Bjornerfeldt S, Webster MT, Vila C: **Relaxation of selective constraint on dog mitochondrial DNA following domestication**. *Genome Res* 2006, **16**:990-994.

39. Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A *et al.*: **The *Sorghum* bicolor genome and the diversification of grasses**. *Nature* 2009, **457**:551-556.

40. Bowers JE, Arias MA, Asher R, Avise JA, Ball RT, Brewer GA, Buss RW, Chen AH, Edwards TM, Estill JC *et al.*: **Comparative physical mapping links conservation of microsynteny to chromosome structure and recombination in grasses**. *Proc Natl Acad Sci U S A* 2005, **102**:13206-13211.

41. Paterson AH: **What has QTL mapping taught us about plant domestication?** *New Phytol* 2002, **154**:591-608.

42. Chapman MA, Pashley CH, Wenzler J, Hvala J, Tang S, Knapp SJ,
●   Burke JM: **A genomic scan for selection reveals candidates for genes involved in the evolution of cultivated sunflower (*Helianthus annuus*)**. *Plant Cell* 2008, **20**:2931-2945.
This study screened 492 EST-derived loci from a panel of wild, primitive, and improved sunflower lines and applied selection scan on this gene set based on multiple criteria. A total of 36 genes showed evidence of selection either during domestication or the subsequent crop improvement. Furthermore, they showed that the genes with signatures of selection often colocalize with known QTLs, and specific gene candidates were identified for more in-depth functional analyses.

43. Doebley J: **The genetics of maize evolution**. *Annu Rev Genet* 2004, **38**:37-59.

44. Paterson AH, Lin YR, Li ZK, Schertz KF, Doebley JF, Pinson SRM, Liu SC, Stansel JW, Irvine JE: **Convergent domestication of cereal crops by independent mutations at corresponding genetic-loci**. *Science* 1995, **269**:1714-1718.

45. Chardon F, Virlon B, Moreau L, Falque M, Joets J, Decousset L, Murigneux A, Charcosset A: **Genetic architecture of flowering time in maize as inferred from quantitative trait loci**

meta-analysis and synteny conservation with the rice genome. *Genetics* 2004, **168**:2169-2185.

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

47. Yamamoto T, Yonemaru J, Yano M: **Towards the understanding of complex traits in rice: substantially or superficially?** *DNA Res* 2009, **16**:141-154.

48. van Leeuwen H, Kliebenstein DJ, West MA, Kim K, van Poecke R, Katagiri F, Michelmore RW, Doerge RW, St Clair DA: **Natural variation among *Arabidopsis thaliana* accessions for transcriptome response to exogenous salicylic acid**. *Plant Cell* 2007, **19**:2099-2110.

49. Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales J, Fish LJ, Worland AJ, Pelica F *et al.*: **'Green revolution' genes encode mutant gibberellin response modulators**. *Nature* 1999, **400**:256-261.

50. Vollbrecht E, Springer PS, Goh L, Buckler ESt, Martienssen R: **Architecture of floral branch systems in maize and related grasses**. *Nature* 2005, **436**:1119-1126.

51. Johnson DS, Mortazavi A, Myers RM, Wold B: **Genome-wide mapping of in vivo protein–DNA interactions**. *Science* 2007, **316**:1497-1502.

52. Fields S, Song O: **A novel genetic system to detect protein–protein interactions**. *Nature* 1989, **340**:245-246.

53. Shang J, Tao Y, Chen X, Zou Y, Lei C, Wang J, Li X, Zhao X,
● Zhang M, Lu Z *et al.*: **Identification of a new rice blast resistance gene, Pid3, by genomewide comparison of paired nucleotide-binding site — leucine-rich repeat genes and their pseudogene alleles between the two sequenced rice genomes**. *Genetics* 2009, **182**:1303-1311.
This study did a genome-wide analysis of the rice NBS-LRR gene family with a focus on the pseudogene members. They found that the *Pid3* locus became pseudogenized in the *japonica* varieties after *indica*–*japonica* split, thus conferring rice blast susceptibility only to the *japonica* geno-types. The study also showed how *in silico* comparative genomics (in this case screening for nonsense mutations in one subspecies versus another) can quickly identify candidate gene loci that may be responsible for varietal and species differences.

54. McNally KL, Childs KL, Bohnert R, Davidson RM, Zhao K, Ulat VJ,
● Zeller G, Clark RM, Hoen DR, Bureau TE *et al.*: **Genomewide SNP variation reveals relationships among landraces and modern varieties of rice**. *Proc Natl Acad Sci U S A* 2009, **106**:12273-12278.
Large scale genotyping in 20 diverse rice varieties and landraces using high density tiling arrays yielded 160 000 nonredundant SNPs. On the basis of the distributions of these SNPs, introgressions are evident that reflect the breeding history and relationships among the selected geno-types. Interestingly, the introgressed regions often correlate with genomic regions known to be responsible for breeding traits. The study culminated in a public SNP resource (OryzaSNP) useful in high-resolution genotyping to study gene–trait association.

55. Huang X, Feng Q, Qian Q, Zhao Q, Wang L, Wang A, Guan J, Fan D,
●● Weng Q, Huang T *et al.*: **High-throughput genotyping by whole-genome resequencing**. *Genome Res* 2009, **19**:1068-1076.
Genotyping by resequencing of RIL individuals between a cross of *japonica* and *indica* using Illumina technology was order of magnitude faster in data collection and more accurate in determining the recombi-nation breakpoints than PCR-based genotyping. Reference genome sequences of both parents permitted SNPs to be called with less ambiguity. After segmenting the chromosomes based on the recombina-tion breakpoints, the authors derived a map based on 2334 recombina-tion bins for the 150 RILs. The bin map proved accurate for QTL mapping and has potential for gene cloning.

56. Rounsley S, Marri PR, Yu Y, He R, Sisneros N, Goicoechea JL, Lee SJ, Angelova A, Kudrna D, Luo M *et al.*: **De Novo next generation sequencing of plant genomes**. *Rice* 2009, **2**:35-43.

57. Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ,
●● Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC *et al.*: **The genetic architecture of maize flowering time**. *Science* 2009, **325**:714-718.
This work combines the merits of linkage and association mapping in a population through a nested association mapping (NAM) design. This population was assembled by pooling the RILs from 25 crosses, with each cross sharing a common parent B73. This is an extremely useful resource for QTL mapping in maize, in which various quantitative traits can be evaluated in multiple crosses. Buckler *et al.* did indeed exploit this resource and scored the flowering time for nearly a million individual plants in four environments and in two years. They found many QTLs with small additive effects on the flowering time. This observation is in contrast with the genetic architectures of flowering time in other studied organisms such as *Arabidopsis* and rice, which are often dominated by a few loci of relatively large effects.

58. Cavanagh C, Morell M, Mackay I, Powell W: **From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants**. *Curr Opin Plant Biol* 2008, **11**:215-221.

59. Heaton EA, Dohleman FG, Long SP: **Meeting US biofuel goals with less land: the potential of *Miscanthus***. *Global Change Biol* 2008, **14**:2000-2014.

60. Hu FY, Tao DY, Sacks E, Fu BY, Xu P, Li J, Yang Y, McNally K, Khush GS, Paterson AH *et al.*: **Convergent evolution of perenniality in rice and sorghum**. *Proc Natl Acad Sci U S A* 2003, **100**:4050-4054.

61. Ming R, Del Monte TA, Hernandez E, Moore PH, Irvine JE, Paterson AH: **Comparative analysis of QTLs affecting plant height and flowering among closely-related diploid and polyploid genomes**. *Genome* 2002, **45**:794-803.

62. Paterson AH, Schertz KF, Lin YR, Liu SC, Chang YL: **The weediness of wild plants: molecular analysis of genes influencing dispersal and persistence of johnsongrass, *Sorghum halepense* (L.) Pers**. *Proc Natl Acad Sci U S A* 1995, **92**:6127-6131.

63. National Research Council: **Lost Crops of Africa**. Washington, DC: National Academy Press; 1996/2006/2008.

64. Tang H, Bowers JE, Wang X, Ming R, Alam M, Paterson AH: **Synteny and collinearity in plant genomes**. *Science* 2008, **320**:486-488.

65. Bhave MR, Lawrence S, Barton C, Hannah LC: **Identification and molecular characterization of shrunken-2 cDNA clones of maize**. *Plant Cell* 1990, **2**:581-588.

66. Wang ZY, Zheng FQ, Shen GZ, Gao JP, Snustad DP, Li MG, Zhang JL, Hong MM: **The amylose content in rice endosperm is related to the post-transcriptional regulation of the waxy gene**. *Plant J* 1995, **7**:613-622.