AUTHOR QUERY FORM

Dear Author,

Any queries or remarks that have arisen during the processing of your manuscript are listed below and highlighted by flags in the proof. Please check your proof carefully and mark all corrections at the appropriate place in the proof (e.g., by using on-screen annotation in the PDF file) or compile them in a separate list.

Uncited references: References that occur in the reference list but not in the text-please position each reference in the text or delete it from the list.

Missing references: References listed below were noted in the text but are missing from the reference list-please make the list complete or remove the references from the text.

AQ1	Please provide a section name here.
AQ2	Is this the Web site you want? Please review.
AQ3	Ref not cited in text; OK to delete?

Electronic file usage
Sometimes we are unable to process the electronic file of your article and/or artwork. If this is the case, we have proceeded by:

☐ Scanning (parts of) your article ☐ Rekeying (parts of) your article

☐ Scanning the artwork

Thank you for your assistance.

c0001

Genetics and genomics of crop domestication

1

J.S. (Pat) Heslop-Harrison   Trude Schwarzacher

University of Leicester, Leicester, UK

Plants and Domestication

s0010

Scope

s0015

In this review of genetics and genomics related to plant bio- p0135
technology and agriculture, we consider the nature of species
that are grown as crops and used by mankind, or otherwise
associated with people. We will then review aspects of the
genetics and genome changes that have been associated with
crop plants and their domestication from their wild rela-
tives before speculating about some of the new opportunities
for plant biotechnology to meet the challenges faced in the
twenty-first century.

Domesticated crops

s0020

Domesticated crops are a subset of all plants. Domesticated p0140
species, whether plants or animals, are considered as those
grown by people for economic or other reasons, and that dif-
fer from their closest wild relatives. Domesticated species are
reliant on human intervention for their reproduction, nutri-
tion, health, planting, and dispersal. They are harvested with
the possibility that a different species will be planted in their
place. Additional characteristics selected for domestication
include size of harvested parts, yield or yield stability, and
quality for the use of the product. There are extensive genetic
differences in all of these characteristics between individuals
within a species as well as between species, and multiple char-
acteristics are selected at the time of domestication that make
the crop worth growing by farmers for millennia and now by
today's plant breeders.

p0145 Genomic techniques allow the underlying selection processes to be understood, exploited, and refined for crop improvement. Genomic scientists can now understand and improve the efficiency of exploitation of genes, genetic diversity, and controls present in crop species and their wild relatives. Domestication of plants, including selection of appropriate species and genetic changes, is one of the features of agriculture, but agriculture also requires knowledge beyond suitable genotypes (Janick, 2005), such as the planting, growing, protection, and harvest of the plants and the accurate timing of the various farming operations.

p0150 Domesticated plants are grown by the human population to meet a range of needs that can be summarized by the six "Fs": food, feed, fuel, fibers (and chemicals), flowers, and pharmaceuticals. Plants within each of these classes have substantial economic impact. Nevertheless, out of 400,000 species of flowering plants, less than 200 have been domesticated as food and feed plants, and just 12 species provide 75% of the food eaten (FAOStat, 2010). Very few of the 1000 gymnosperms, and arguably none of the 15,000 ferns and allies, have been domesticated. New knowledge of genetics and improved techniques of selection, hybridization, or gene transfer have the potential to enable more species to be domesticated.

p0155 As well as domesticated crop species, there are many spices, pharmaceutical (and medicinal), horticultural, and garden ("flowers") plants collected over the last millennia from the wild and cultivated on a small scale. These plants may be genetically similar and as diverse as their wild relatives, although one or a small number of genes may have been selected. Many of the selections require human intervention to survive, often because they are grown outside their natural climate range or have abnormalities that are regarded as attractive or useful but reduce plant fitness. However, with the exception of some hybrids, the limited changes mean they are not normally considered as domesticated.

Weeds

s0025

p0160 Weeds and invasive species are associated with human farming and habitation, although they are not normally considered as domesticated species. There has been limited genomic and genetic work on most of these species with notable exceptions, in particular *Arabidopsis thaliana*. Harlan and deWet (1965) defined a weed as "a generally unwanted organism that thrives in habitats disturbed by man"; like crops, weed species are extremely diverse, and have different strategies for survival. The effect of weeds on agriculture can be devastating, such as taking nutrition from the crop, making harvest difficult, or reducing the value and quality of the harvest.

p0165 Most crop plants will not establish themselves in an environment where weeds thrive and active intervention is needed to remove competition. In an extensive study of feral oil seed rape (canola or *Brassica napus*), Crawley and Brown (1995) showed the very high level of turnover of site occupancy on highway verges, with local extinction occurring within three years in the absence of new seeding and soil disturbance. In contrast, weeds can be notably persistent, with; for example, nettles (*Urtica dioica*) remain as markers of sites of habitation after hundreds of years in northern Europe in the absence of further habitation or evidence of crops. There are strong selection pressures on weeds to benefit from the human-created habitat at the farm (rather than plant breeder) level, working with potentially much larger and more widely distributed populations than breeders use. Weeds may mimic the growth forms or seeds of crops and are distributed or grow along with them. The selection is not applied to yield and quality characteristics, but on survival and population distribution or expansion, with key genes such as those for seed dormancy or dehiscence (see the section Genes of Domestication) potentially selected in the opposite direction from the seeds of a crop.

p0170 Crops can become weeds. In the *Brassicas*, for example, the same genotype may be a weed with low yield and poor characteristics in one environment, but a robust crop with desirable properties in another. Volunteers — plants from a previous crop on the same land — are a major challenge in growing many field crops. They thrive in the crop conditions; the economic damage from these weeds includes acting as reservoirs of crop-specific diseases over several seasons in a rotation.

p0175 Weeds have no harvest value in a crop, reducing yields, and making crop management difficult, so farmers have been improving their weed control methods since the start of agriculture. In advanced commercial farms weed control is an expensive part of the agronomy, while for smallholders and subsistence farmers, the continuous labor required can be one of the most tedious and demeaning operations for the people, usually women and children, that are involved. The removal and control of weeds is environmentally costly and involves burning, herbicides, deep plowing and multiple soil cultivations, processes leading to erosion, poor soil moisture conservation, use of large amounts of energy, loss of soil structure, uncontrolled fires, and smoke or pollution. Approaches to weed control have changed continuously over millennia, including use of fire, planting methods, and plowing. As well as the application of agronomic and technological approaches to limit weed spread, breeders must also consider the genetic characteristics of weeds and both the potential of a crop to become a weed and the ease of control of weeds within a new variety. This work interacts with making models of population biology based on the understanding of weed characteristics such as developmental plasticity or seed dormancy.

Invasive species

s0030

p0180 Another group of plants associated with humans are the invasive species. Along with habitat destruction, invasive species are often considered to be the major threat to biodiversity worldwide, although Gurevitch and Padilla (2004) pointed out that the cause and effect data are generally weak. Genetics and genomic research is required to understand the biology of invasives, so that the characteristics that led to uncontrolled displacement of native species can be avoided in the breeding of crops. The requirements of crops including high partitioning of the plant's resources to the harvestable product, non-distribution of seeds, and uniformity of growth

tend to mean that few domesticated crops have invasive characteristics. However, a number of horticultural plants and those introduced for their novelty value have caused problems both in agriculture and the wild in very diverse environments ranging from temperate and tropical, through fresh water, grasslands, and woods to uplands, with the species taking advantage of man-made or man-influenced habitats. Examples of invasive species causing significant problems include water hyacinth (*Eichhornia* species), *Rhododendron*, knotweed (*Fallopia japonica*; Bailey et al., 2007), kudzu (*Pueraria* spp.), and some ferns (bracken, *Pteridium*, and *Azolla* spp.). This is notable since ferns have not been domesticated as crops.

Model species and crop sciences

p0185 The diversity in growth forms, reproduction, and uses between the crops means that most crop scientists have focused their work on a single species, while fundamental studies adopted a small number of convenient models. During much of the twentieth century, major research or model species were crops because they could be easily obtained and grown worldwide, and laboratory protocols, resources, and background information were extensive. Spinach was used for many studies of photosynthesis (e.g., Bassham and Calvin, 1955), maize was used for genetics (e.g., McClintock et al., 1981), and carrot or tobacco was used for tissue culture. However, for genetics, a fast generation time, small plant size, and the ability to mutagenize populations were major advantages. Researchers including Kranz, Redei, and Koornneef (e.g., Koornneef et al., 1983) established *A. thaliana* as a model species in the 1970s, and, because of its small genome size (165 Mbp), *Arabidopsis* was chosen to be the first plant to have its DNA sequenced (Arabidopsis Genome Initiative, 2000). The ease of growing large numbers under controlled conditions and extensive scientific resources led to it becoming the model for plant research in many laboratories. Rice became the second plant genome to be sequenced, because of its status as one of the world's two major crops, relatively small genome size of 435 Mbp, and contrasting taxonomic position to *Arabidopsis* (e.g., Sasaki et al., 2002). A major justification of these sequencing projects was the suggestion that the gene content of all plants would be similar, a prediction that has largely held true (e.g., Figure 3 in Argout et al., 2011), although sequencing led to some surprises including the low total number of genes — typically 30,000 — found in all organisms.

p0190   With the advent of plant biotechnologies, genomics, mathematical modeling, and informatics, a large number of tools and results of general nature can be applied across most crops and potential crops (see review by Moose and Mumm, 2008). Few crop scientists are now restricted to work on one species and need to exploit approaches and results with other crops and model species. In the genetics and genomics field there are many parallels between species, making it essential to integrate information. Throughout history and prehistory, humans have been classifying plants, assessing their similarity to use as food or medicines, and avoiding or processing toxic plants long before the advent of agriculture.

A succession of techniques including morphological study, crossing, karyotype analysis, DNA sequence comparisons, and now whole genome sequences has established plant relationships. The Angiosperm Phylogeny Group (2009) presents a robust, monophyletic phylogeny showing relationships between all angiosperms; better understanding of the evolution and phylogeny is important for crop genetics because it shows the most closely related species to use to find valuable characteristics.

Understanding Domestication Processes

Evidence of relatives and processes of early domestication

p0195 The early processes of domestication can be inferred from examination of wild crop relatives and comparison with existing crops at the morphological, physiological, genetic, or DNA levels. Since farming and domestication is less than 10,000 years old, the archaeological record of the introduction of species into agriculture is rich (Zeder, 2006) and documents some aspects of the transition from hunter–gatherer societies to sedentary, farming-based communities. Indeed, the earliest hunter–gatherer cave paintings date from 32,000 years ago (Clottes, 2010) and in combination with archaeological evidence they show the pre-agricultural period. The domestication process happened independently in Southeast Asia and the Middle East, and soon after it is found in Asia, Europe, Africa, and the Americas after the retreat of the Pleistocene ice around 12,000 years ago. The domestication of all of the major crops now grown started at about the same time. Pictures of domesticated plants appear in Chinese and Arabic manuscripts up to 2000 years ago (Paris et al., 2009; Wang et al., 2008; Janick, 2005) and can be correlated with archaeological evidence. With the use of genetic markers to genotype crops and their relatives found in various locations, Salamini et al. (2002) reviewed how genetic markers traced the sites of domestication of cereals to wild populations of grasses in the Near East, and Gross and Olsen (2010) discussed that genetic inferences about geographical origins of crops and the number of independent domestication events are compatible with archaeological data.

p0200   Domestication of particular species, and the genetic characteristics that make them different from their wild relatives, are also associated with technology used in agriculture societies for planting, harvesting, threshing, transport, and storage; or long-lasting infrastructure like roads, habitations, and field organization; and domestic arrangements including specialized storage and preparation premises or cooking processes. All of these give additional information about the genetic changes from wild species since genotypes must complement the societal practices. In the first decade of the twenty-first century, genetic and genomic methods enabled examination of the processes of crop domestication, including both the identification of the genetic basis and its origin and the duration of domestication (Papa et al., 2007).

s0050 Genes of domestication

p0205 The "suite of traits" including seed dispersal, seed dormancy, gigantism in the harvested parts, determinate and synchronized growth, increased harvest index, and change in sweetness or bitterness have been called the "domestication syndrome" after Hammer (1984). These characteristics make a crop worth growing, and without them the difficulties of planting, cultivation, and poor harvest make them unrewarding to grow. It is likely that a combination of all of the characteristics must be present together for a species to reach the first stage of domestication, since most of these traits in some form are present in all domesticated crops. Doebley (2004) and Doebley et al. (2006) reviewed data showing that the differences in cultivars mean that wild progenitors of crops are not easily recognizable. Furthermore, many of these characters are so disadvantageous in non-cultivated situations that the crop will not establish in the wild (Crawley and Brown, 1995): indehiscent plants will not distribute seeds, whereas an annual plant bearing seeds without dormancy means the species would not survive one bad season.

p0210 Further evidence, at least in the cereals, for the importance of the small number of domestication syndrome genes comes from the similarity of changes in several domesticates known as convergent evolution. Paterson et al. (1995) showed that the same genes and gene pathways were involved in domestication of sorghum, rice, and maize. As with other genetic effects, many domestication characteristics are regulated by quantitative trait loci (QTLs) where several genes have effects (Varshney et al., 2006), and transcriptional regulators (rather than enzymatic or structural) genes (Doebley et al., 2006; Martin et al., 2010) are often involved.

s0055 Genetic variation and domestication

p0215 Genetically, any requirement for change of multiple characters simultaneously requires either an extremely unusual conjunction of genetic mutations or recombination, or selection and intercrossing to bring characters together over many generations. Clearly, the latter did not happen to any great extent, and genetic and genomic data collected over the last decade do suggest that the diversity of alleles present in domesticated species is lower than in their wild progenitors. This supports the domestication syndrome concept with a number of characteristics coming together at one time. This selection has left a "genomic signature" in all current crops, present thousands of generations later, and the loss of diversity compared to the wild species is seen as a "genetic bottleneck" (Doebley, 2004). Genetic analysis has shown that many of the gene alleles involved in the domestication syndrome are present within the gene pool of wild progenitors of crops, although with a low frequency, whereas other traits are apparently new mutations (Doebley et al., 2006; Huang et al.,

AQ1 2007; and see the section). One important approach to identifying genetic bottlenecks has been comparison of genomic regions neighboring key domestication traits with selectively neutral regions; reduced variation in linked genes suggests that the number of domestication syndrome genes is limited.

The "selective sweep" of the genome (Clark et al., 2004) with directional selection leads to reduced variation and linkage disequilibria (Anhalt et al., 2008, 2009) in the selected regions.

p0220 Whereas only a few plants have carried critical traits related to domestication and have been used for most subsequent breeding, the genetic bottleneck or "founder effect" will have reduced the diversity to a small number of gene alleles present in the original selected population (changing gene allele frequencies, eliminating rare alleles, and introducing linkage disequilibrium). It has widely been considered, especially on theoretical grounds, that genetic drift will have further reduced the diversity after domestication, given that the selection of a few hundred varieties at most for use in breeding represents a tiny population size. In many cultivated crops, the level of genetic polymorphism has been reduced by 60 to 90% in passing through the genetic bottleneck in cultivars compared to wild relatives (e.g., Buckler et al., 2001 in maize). Similarly, rice cultivars may include only 10 to 20% of the diversity present in the wild relatives (Zhu et al., 2007). Even with extensive data, it remains challenging to distinguish between the monophyletic and polyphyletic origin of a crop using molecular markers. As noted by Zhang et al. (2009), genetic marker data can indicate that the two cultivated rice subspecies, *indica* and *japonica*, either evolved independently at different times and sites (Tang et al., 2006), or had a monophyletic origin from a common wild rice that subsequently separated. The diversity restriction is not universal, and the polyphyletic origin of some polyploid crops has probably reduced the bottleneck effect: hexaploid bread wheat (AABBDD genome constitution) has much of the genetic diversity present in its progenitors (Dubcovsky and Dvorak, 2007) and originated recurrently with ancestral D-genomes (Caldwell et al., 2004), even if all the D-genome variation is not represented (Saeidi et al., 2008). Cifuentes et al. (2010) discussed the polyphyletic origin of canola (oilseed rape, *B. napus*), which incorporates variation from both the duplication of loci as a polyploid and from several ancestors of the tetraploid cultivars.

p0225 In domesticated species, artificial selection is the main evolutionary force because humans — farmers and more recently plant breeders — exert strong selection pressure compared to that from the environment where a species is established (Innan and Kim, 2008). These authors pointed out that artificial selection may act on alleles that may have been neutral variants before domestication, and the fixation of these may not remove DNA variation in the surrounding region, depending on the initial frequency of the beneficial alleles. The number of alleles selected during domestication, the population sizes, and the number of independent selection events will all affect the intensity of the selection bottleneck.

s0060 Genetic control related to diversity and speciation

p0230 While geographical isolation of populations stops gene flow within a species, it is far from the only effect that leads to separation of genotypes. Rieseberg and Blackman (2010)

have identified no less than 41 different genes that can lead to reproductive isolation of populations. Genetics related to plant evolution and isolation is a relatively recent research area, and it is clear that the identification of genes that effect reproductive behavior — recombination and interact with effects on fertility, leading to isolation and speciation (Heslop-Harrison, 2010) — may well show how some of the selective sweeps (Nielsen et al., 2005) have been driven during crop domestication. Understanding the genetic effects and genes that enable these processes may show how levels of diversity can be maintained within species, whether in wild ecosystems or crops.

s0065

Domestication of maize

p0235

One of the best understood examples of genetic and genomic changes during domestication comes from maize (*Zea mays* or corn in the Americas), where the seminal work of Doebley and colleagues (Doebley et al., 2006; Wright et al., 2005) identified the relatively few genes giving rise to the major physiological and morphological differences between maize and its closest wild ancestor, teosinte (represented by several *Zea* species). Maize, with naked grains in multiple rows and 10 to 100 times more kernels per ear, has a very different appearance from the branched teosinte, which has grains with a hard seed coat on inflorescences that shatter (disarticulate) when ripe and carried on multiple stalks. Among the first genes identified was *teosinte branched 1* (*tb1*), a transcriptional regulator that represses the branching (Doebley, 2004). The gene *teosinte glume architecture*, *tga1* (Wang et al., 2005), is a key single-gene that controls development of the hard coat around the kernel in teosinte. It was identified by high-resolution genetic mapping and map-based cloning. Doust (2007) more generally studied the developmental genetics of grass plant-architecture in genetic, evolutionary, and ecological contexts. He concluded that exploring the phylogenetic context of the crop grasses suggests new ways to identify and create combinations of morphological traits that will best suit future needs: knowledge of past events shows how future breeding can proceed.

p0240

Technically, works such as those previously mentioned have focused on making experimental hybrid populations for genetically mapping traits that can be identified as domestication related. Another group of researchers took a large-scale approach to characterizing how bottlenecks and artificial selection have altered genetic variation during domestication of teosinte to form maize using an unbiased, genome-wide approach. Wang et al. (2005; see also Vigouroux et al., 2005) measured single nucleotide polymorphism (SNP) levels in 774 genes, and found that the maize inbred lines had only 57% of the variation present in the teosinte sample, showing evidence for the genetic bottleneck. The genes could be divided into two classes based on the variation signatures at single nucleotides (SNPs): 2–5% of the genes were under selection during domestication and have been selected with 10 times the intensity of the selectively neutral genes where limited population size alone has reduced the variation. Yamasaki et al. (2005, 2007) sequenced 1095 maize genes

from various lines and identified eight genes with no variation between inbred maize lines, but with SNP variation in teosinte; six showed selection throughout the DNA sequence of the gene, while two had signatures of selection in the 3'portion of each gene. The functions of the genes, examined after the analysis, were "consistent with agronomic selection for nutritional quality, maturity, and productivity," although most had not been identified previously as being associated with their selection in the crop.

Domestication of legumes   s0070

Weeden (2007) examined the domestication of the pea (*Pisum sativum*), and identified approximately 20 genes or QTLs responsible for the domestication of it. Because of the availability of a range of germplasm from the pea, a time line for the "domestication syndrome" genes could be established. Domestication syndrome characters including indehiscent pods, seed dormancy, gigantism as seed weight, and earliness were seen in the most primitive lines, while dwarfing, harvest index, photoperiod-sensitivity and white flowering, along with additional seed weight traits, appeared much more recently. This is evidence for the model shown by Gross and Olsen (2010) that domestication is a two-stage process. First, is a rapid process that makes the crop worthwhile to grow, including the domestication syndrome traits that allow a crop to be reliably sown, cultivated, and harvested such as uniform seed germination and fruit ripening. This is then followed by a stage acquiring traits over a longer period that improves the crop.

p0245

A second finding of Weeden (2007) showed that, although the phenotypic characters are similar, the genes involved in pea domestication are different from those in the common bean, *Phaseolus*, contrasting with the conclusion showing convergent evolution in rice, maize, and sorghum (Paterson et al., 1995). Weeden is optimistic that the presence of multiple genes means that there are several ways for breeders to modify unwanted characters and avoid detrimental effects associated with some otherwise valuable alleles.

p0250

Several studies have investigated the genetic diversity and signatures of domestication in soybean, a species with a center of origin and domestication in South China. Guo et al. (2010) proposed a single origin with a moderately severe genetic bottleneck during domestication, showing that wild soybeans in South China have an unexploited and valuable gene pool for future breeding. However, Hyten et al. (2006) examined other populations, finding that there were several rounds of reduction of genetic diversity, following domestication in Asia to produce numerous Asian landraces and introduction of a few genotypes to North America. Notably, they found modern cultivars retained 72% of the sequence diversity present in the Asian landraces but lost 79% of rare alleles, with the major constrictions of diversity coming first from the domestication event, and secondly from the introduction of a small number of races to North America, while later breeding has had less effect.

p0255

Grasses tend to have inflorescences where all individuals flower together and the seeds reach maturity at a similar time,

p0260

which is certainly an advantage for agriculture. However, other wild plants flower and set seed over a long part of the crop season, making growing and harvest of the ripe seed difficult. In species such as soybean (*Glycine max*), determinacy of growth through the character of a terminal flower is an agronomically important trait associated with the domestication. Most soybean cultivars are classifiable into indeterminate and determinate growth habit, whereas *G. soja*, the wild progenitor of soybean, is indeterminate. Tian et al. (2010) took a candidate-gene approach to demonstrate that the determinate growth habit in soybean is controlled by a single gene homologous to *TFL1* (terminal flower) in *Arabidopsis*, which is a reasonable expectation. The genetics of the determinate habit has been known since the 1970s, and mapped more recently. There are, as expected from the known genetic background of soybean, four homologous copies for the determinate genes.

s0075 Yield traits

p0265 Yield, affected by gigantism and number of harvested units, is normally a quantitative trait with continuous variation and complex heritability. However, analysis and partitioning of yield components, combined with use of well-designed test crosses and large populations, is allowing key regions of the genome — in some cases now correlated with genes — to be identified. Genes increasing harvestable yield have been extensively studied using genetic and genomic approaches. Measurements of yield components, starting long before extensive use of genomic approaches, showed that, for example, rice yield includes traits such as grain number and grain weight, or duration and rate of grain-filling, and is regulated by multiple QTLs (Yano, 2001). Use of appropriate hybrid populations segregating for yield characteristics, such as biomass in forage grasses (ryegrass, *Lolium perenne*; Anhalt et al., 2009) or fruit yield in tomato, *Solanum lycopersicoides* (*Lycopersicon esculentum*; Cong et al., 2002), is showing that genetic regions on the map are responsible for a large part of the variation in yield observed. However, often a large number of genetic regions are identified: in tomato, no less than 28 different QTLs affecting fruit weight have been identified (Cong et al., 2002). QTL analysis is also of potential importance when identifying characters where the same gene affects different traits; this could indicate selection in opposite directions is unlikely to succeed (e.g., grain protein and yield or palatability/sweetness and insect resistance).

s0080 Hybrid Species and New Polyploids in Domestication

p0270 Most of the species previously discussed have a genetic structure similar to their wild relatives, such as fertility and reproduction through seeds. However, a group of crop species have a different genomic constitution from wild species, bringing together copies of genomes from different ancestral species that are not found normally in nature. This includes species that have different chromosome numbers from their relatives or are hybrids.

p0275 Among early domesticates, the banana is an interesting example. Wild, fertile, diploid bananas have small fruits and large seeds with very little fruit pulp that is eaten. However, almost all of the cultivated lines are sterile and parthenocarpic, producing fruits in the absence of seeds. This is a characteristic associated with triploidy, so cultivars have a chromosome constitution of $2n = 3x = 33$, while the fertile wild species are $2n = 2x = 22$. The ultimate origin of the accessions that have become cultivars is unclear, although there are many independent parthenocarpic cultivars selected by early farmers in South East Asia that are vegetatively propagated (Heslop-Harrison and Schwarzacher, 2007). De Langhe et al. (2010) presented evidence indicating the complex origin of the cultivars, involving intermediate hybrids and backcrossing, with additional clonal variation. They note that this gives extensive variation, but makes development of artificial breeding schemes difficult.

p0280 Polyploid cytotypes can be larger than their diploid progenitors, and this has lead to their selection and cultivation in blueberries and strawberries (the octaploid hybrid is cultivated, and small wild species are mostly diploid; Schulze et al., 2011). This advantage is not universal and fruit gigantism can be under the control of several genes. Cultivated grapes are much larger than their wild progenitors, but there is no recent genome duplication in their ancestry (French-Italian Public Consortium, 2007).

Post-Domestication Selection s0085

Modifications in crop characteristics s0090

p0285 Since early agriculture, crop improvement has been a continuous process over thousands of years, driven by the need for disease resistance, adaptation to new and changing climates, quality, ability to propagate and grow, and yield. Systematic breeding since the mid-twentieth century — required by the need for increased food production, new crop uses, a different socioeconomic environment, climate, and water use changes — and new quality requirements have accelerated breeding with unprecedented speed.

p0290 As well as the continuing increase in the world's population, people worldwide have moved into conurbations. In Europe, only 2% of the population is engaged in agriculture, a change seen in global trends, with 2008 being the first year when more than half the human population lived in cities.

p0295 Changes in demand have certainly lead to substantial changes in the nature of crops. This divorce of populations from local food production requires new infrastructure to bring agricultural production to the people, and perhaps new genotypes appropriate to the transport chain. Increased meat consumption, particularly of chicken and pork, is seen as a significant global trend, requiring not only farming of the animals, but production of the crops to feed these animals. Another significant global trend is the increase in plant oil production driven by consumer demand. Major crops such as oil palm, oilseed rape, and soybean show the greatest increase in production over the last 50 years. The increase in animal and

oil consumption is widely considered to have negative effects on human health, although the use of biotechnology allows modification of the fatty acid content of oilseed plants for health purposes or to produce nutritional fatty acids not normally found in crop plants (Damude and Kinney, 2008).

p0300   Over a similar period, several crops have shown substantial declines either globally or regionally. Notable examples would be fiber crops including hemp, jute, and flax that have been replaced by petrochemical-based ropes and textiles, while oats in northern Europe are no longer grown as feed for horses.

p0305   Mechanization of agriculture in the twentieth century in developed countries has directly led to landscape-wide effects such as larger fields, more land in cultivation, and changed water or erosion management. Improved crop protection and artificial nitrogenous fertilizer application has also led to substantial changes in landscape and crop mixes, largely removing the need for fallowing or cover crops. Along with labor costs (which are very high in crops where mechanization has been limited), farming practices have certainly changed both the mix of crops grown and the requirements from the genetics of the varieties in a complex relationship between consumer income, tastes, import availability, and price.

p0310   Is biotechnology affecting the species and acreage of crops? Two crops showing substantial production declines regionally have been restored to production with transgenic varieties. Lines of cotton carrying insect resistance with the Bt-toxin have lead directly to an increase (or stability from a declining trend) in acreage of cotton. Papaya, where Hawaiian production fell by 40% and moved from the Oahu island because of devastation by the papaya ringspot virus, has been restored to production by the introduction of the resistance transgenic variety from 1998 (http://www.hawaiipapaya.com/rainbow.htm). It is predictable that intensive production of other crops will follow similar patterns (King et al., 2003, 2004) in the future with either stoppage of production or introduction of transgenic varieties.

AQ2

s0095   ## New Domestication

s0100   ### Domesticated species

p0315   About 10% of all plants species are suitable for food, and a higher proportion has been used as species or flavors, or has medicinal value. However, most species have not met the first requirements for successful domestication. The small number of plants that have been domesticated, along with their repeated domestication in a number of localities, might suggest that no other species are suitable for domestication, at least for production of substantial amounts of food. Are there other plants that have appropriate genetical attributes to become a crop? It may be that the genetics was complex or allelic diversity was restricted so that it never become worthwhile to cultivate the plant at the early period of domestication, and later in history, refinement of genetic traits had already given the major crops a significant advantage so others could not compete in terms of ease of husbandry, harvest, nutritional value, or some other quality. There are prospects

for domestication of new species, although in practice the number of new species introduced to agriculture in the last century on a significant scale is very limited. On the FAO list of crops, only kiwi fruit and the hybrid cereal triticale are significant additions in the last 50 years. As well as the new crop classes mentioned earlier, there are species currently collected from the wild that are being prospected for domestication. Overexploitation is threatening the survival of several medicinal plants and spices, providing an incentive for increasing the scope of domestication.

p0320   Our understanding of the domestication processes in the major crops can now be applied to advancing domestication in species that have not been domesticated (Heslop-Harrison, 2002). Although the farmer and consumer demands are higher than ever, the use of molecular methods and genomics covered in this volume are likely to mean that relatively rapid selection of unimproved wild species is possible, including some fruits and nuts with limited breeding (Heslop-Harrison, 2004). Appropriate diversity can be found within germplasm collections and multiple DNA markers can combine traits and select from huge populations of plants or transformation can introduce single-genes.

s0105   ### Lost crops

p0325   Major crops were domesticated early in the history of agriculture, and the rarity of introductions was discussed earlier. It is also valuable to consider crops that were cultivated more extensively in prehistory — meeting the early requirements for being worthwhile crops, but not being as useful, or improved at the rate of other species more recently. Proso millet (*Panicum miliaceum*) was, like wheat, widely grown, as is demonstrated by archaeological evidence (Hunt et al., 2008). However, since this early period, its production has reduced to less than 1% of the production of wheat. It is notable for being the most water-efficient grass (Heyduck et al., 2008), but there is minimal genetic work (not even the ancestors of the tetraploid have been defined clearly), so it is one example where further work on its genetics will be valuable. A number of legumes, now minor, may also be interesting "lost crops."

s0110   ### Trees and biofuels

p0330   Trees provide fuel and fiber for construction, utensils, and paper. Planting and coppicing of trees has long been practiced, with selection of trees for regeneration and yield. However, the selection and characterization of trees for construction timber or fiber has been less systematic because of the availability of timber trees in native forests, the long timescales involved in cultivation, lack of continuous revenue, and large capital costs involved. The selection methods and genetics being applied to trees are rapidly changing, and application of genomics and marker technology has potential to improve tree characteristics (Neale, 2007; Gailing et al., 2009)

p0335   For biofuels, the twenty-first century has seen the establishment of several genomics-based research programs that are looking at both the improvement of existing crops and

introduction of new crops with high biomass yields suitable for fuel production as well as new ways of processing using microorganisms to break down biomass (Heaton et al., 2008; Rubin, 2008; Somerville et al., 2010)

s0115 Genetics and breeding for new uses: Ecosystem services

p0340 Most plants planted and grown on a large scale by man provide products that are traded and used or have horticultural or amenity value. However, as well as conservation management of wild environments, it is probable that there will be increased selective breeding for plants that improve the local or wider environment. During the twentieth century, plants have been used for habitat restoration of mining sites, and the selection of appropriate species and genotypes of trees, shrubs, and grasses for colonization of these sites has, along with improvements in earth handling and planting methods, made enormous improvements in the landscapes of derelict areas (Richardson, 1975). Elsewhere, plantings are used to stabilize soils or sands and prevent erosion or drift. As well as the poor substrates, mining wastes may be contaminated by heavy metals, and Bradshaw and colleagues carried out extensive work on the selection of genotypes for land remediation, which are now widely applied (Antonovics et al., 1971).

t0010
AQ2

Table 1.1 Key features of selected model species and major and minor crops related to the organization of their genomes

Crop	Species	Family	1C Genome size (Mb)[c]	2n Chromosome number	Ploidy level	Life form	Life span	Climatic range
Maize	*Zea mays*[a]	Poaceae	2670	20	2x (4x)	Herb	Annual	Temperate
Wheat	*Triticum aestivum*	Poaceae	17,000	42	2	Herb	Annual	Temperate
Rice	*Oryza sativa* ssp. *indica*[a]	Poaceae	420	24	2	Herb	Annual	Tropical
Rice	*Oryza sativa* ssp. *japonica*[a])	Poaceae	466	24	2	Herb	Annual	Temperate to tropical
Sorghum	*Sorghum bicolor*[a]	Poaceae	730	20	2	Herb	Annual or perennial	Tropical to subtropical
Barley	*Hordeum vulgare*	Poaceae	5400	14	2		Annual	Temperate
Rye	*Secale cereale*	Poaceae	8100	14	2		Annual	Temperate
Pearl millet	*Pennesitum glaucum*	Poaceae	2620	14	2	Herb	Annual	Tropical
Foxtail millet	*Setaria italica*	Poaceae	513	18	2	Herb	Annual	Temperate to subtropical
Sugar cane	*Saccharum* sp.	Poaceae	3960 (80 chr)	80–128	8x or more	Giant herb	Perennial	Warm temperate to tropical
Potato	*Solanum tuberosum*	Solanaceae	2050	48	4	Herb	Perennial	Temperate
Tomato	*Solanum lycopersicum*	Solanaceae	1000	24	2	Herb	Perennial, grown as annual crop	Temperate to subtropical
Cassava	*Manihot esculenta*	Euphorbiaceae	807	36		Woody shrub	Perennial	Tropical to subtropical
Soybean	*Glycine max*[a]	Fabaceae	1100	40	2	Annual herb	Annual	Temperate to subtropical
Groundnut or peanut	*Arachis hypogaea*	Fabaceae	2807	40	4	Herb	Annual	Tropical to warm temperate

Beyond use for land remediation, other plant species provide "ecosystem services" such as waste decomposition, water purification, hydrology improvement through root systems, fencing, or hedging. It is certain that the uses of plants to provide these services will increase as their value is recognized through economic methods (e.g., the Millenium Ecosystem Assessment, 2010, undertaken in connection with the United Nations system). New ecosystem values are likely to be introduced, including "carbon capture."

p0345   The human uses of plants to provide ecosystem services have currently involved selection of appropriate genotypes from the wild. However, there is a genetic basis for the properties needed, and systematic breeding, including use of

biotechnology (Chory et al., 2000), can improve their performance enough so there is enough improvement to balance the research and breeding costs.

Features of Domesticated Genomes   s0120

Table 1.1 summarizes the remarkable diversity in fundamen-   p0350
tal characteristics of some major and minor crops and some other comparator species including features of their genome organization and size (see discussion in Heslop-Harrison and Schwarzacher, 2011). The crops have mostly been domesticated and then selected by farmers and breeders over several

Predominant breeding system	Propagation	Parts used	Commodity	Nutritional use	World production 2009 (Million tonnes)	Storage
Cross-pollinating / monoecious although self-fertile	Seed	Grain endosperm/ leaves	Cereal/foliage	Starch, protein, oil	817	Dry/years
Self-pollinating	Seed	Grain endosperm	Cereal	Starch, protein	682	Dry/years
Self-pollinating	Seed	Grain endosperm	Cereal	Starch, protein	679	Dry/years
Self-pollinating	Seed	Grain endosperm	Cereal	Starch, protein		Dry/years
Self-pollinating occasionally out-crossing	Seed	Grain endosperm/ leaves	Cereal/foliage	Starch, protein	62	Dry/years
Self-pollinating	Seed	Grain endosperm	Cereal	Starch, protein	150	Dry/years
Out-crossing	Seed	Grain endosperm	Cereal	Starch, protein	18	Dry/years
Out-crossing	Seed	Grain endosperm	Cereal	Starch, protein	32 (millet)	Dry/years
Self-pollinating occasionally out-crossing	Seed	Grain endosperm/ leaves	Cereal/foliage	Starch, protein	32 (millet)	Dry/years
Cross-pollinating	Stem cutting	Stalks	Sugar	Sugar	1683	Processed
Not true breeding	Vegetatively (tuber)	Tuber	Vegetable	Starch	330	Months
Self-incompatible, self-fertile in some cultivars	Seed	Fruit	Vegetable	Dietary fiber, antioxidants[b]	141	Fresh
Out-crossing	Vegetatively (stem)	Root	Vegetable	Starch	241	Days/only in ground
Self-pollinating	Seed	Seed (cotyledon)	Protein and oil	Protein, oil	222	One year
Self-pollinating	Seed	Pods with seed (cotyledon)	Vegetable, oil	Protein, fat, nutrient rich	36	One year

(*Continued*)

t0010
AQ2

Table 1.1 (Continued)

Crop	Species	Family	1C Genome size (Mb)[c]	2n Chromosome number	Ploidy level	Life form	Life span	Climatic range
Alfalfa	*Medicago sativa*	Fabaceae	841	16/32	2x/4x	Herb	Annual	Warm temperate
Oil palm	*Elaeis guineensis*	Arecaceae	1800	32	2	Tree	Perennial	Tropical
Date palm	*Phoenix dactylifera*	Arecaceae	929	36	2	Tree	Perennial	Tropical
Coconuts	*Cocos nucifera*	Arecaceae	3472	32	2	Tree	Perennial	Tropical
Sugar beet	*Beta vulgaris*	Amaranthaceae	1223	18 or 36	2x or 4x	Herb	Biennial	Temperate
Banana and plantain	Banana cultivars	Musaceae		33	3x	Giant Herb	Perennial	Tropical
Wild banana	*Musa acuminata*	Musaceae	550	22	2	Giant herb	Perennial	Tropical
Sweet potato	*Ipomoea batatas*	Convolvulaceae	1467	60		Herb/vine	Perennial	Tropical to warm temperate
Onion	*Allium cepa*	Alliaceae	16382	16	2	Herb	Biennial	Temperate
Rapeseed	*Brassica napus*	Brassicaceae	1125	38	4	Herb	Annual	Temperate
Cabbage	*Brassica oleracea*	Brassicaceae	758	18	2	Herb	Annual	Temperate
Arabidopsis	*Arabidopsis thaliana*	Brassicaceae	165	10	2	Ephemeral	Annual	Temperate
Oranges (citrus)	*Citrus × sinensis*	Rutaceae	611	18	2	Tree	Perennial	Warm temperate to subtropical
Apples	*Malus domesticus*	Rosaceae	327	34/51	2x/3x	Tree	Perennial	Temperate
Strawberry	*Fragaria × ananassa*	Rosaceae	597	56	8	Herb	Perennial	Temperate
Wild strawberry	*Fragaria vesca*[a]	Rosaceae	240	14	2	Herb	Perennial	Temperate
Grape	*Vitis vinifera*[a]	Vitaceae	490	38	2	Shrub	Perennial	Temperate to warm temperate
Cucumber	*Cucumis sativus*[a]	Cucurbitaceae	367	14	2	Herb	Annual	Temperate
Olive	*Olea europea*	Oleaceae	1907	46	2	Tree	Perennial	Warm temperate
Lettuce	*Lactuca sativa*	Asteraceae	2590	18	2	Herb	Annual or biannual	Temperate

Predominant breeding system	Propagation	Parts used	Commodity	Nutritional use	World production 2009 (Million tonnes)	Storage
Out-crossing and self -pollinating	Seed	Leaves	Foliage			Fresh/ processed
Out-crossing/monoecious	F1 seed	Fruit (mesocarp and kernel)	Oil	Oil	207	Months
Out-crossing/dioecious	Seed/ cuttings	Fruit (mesocarp)	Fruit	Fruit	7	One year
Out-crossing	Seed	Seed (endosperm)	Fruit	Fruit, fiber	4	
Out-crossing, occasionally self- fertile	Seed	Root	Root vegetable	Sugar	229	Month/ processed
Sterile	Suckers/tissue culture	Fruit	Fruit	Starch tropics, temperate)[b]	130	Weeks
Cross-pollinating but self-fertile	Suckers	Fruit/leaves	Fruit/fiber	Starch		📝
Out-crossing	Seed	Tuber	Vegetable	Starch	108	Week
Cross-pollinating but self-fertile	Seed/bulb	Bulb	Vegetable	Flavoring, starch	72	Months
Out crossing/ self-incompatible	Seed	Seed	Oil	Oil	62	One year
Out crossing/ self-incompatible	Seed	Leaves	Vegetable	Antioxidants/ vitamins[b]	71	Fresh
Self-pollinating	Seed					
Self fertile, some self-pollinating"	Grafting	Fruit	Fruit	Dietary fiber, vitamins[b]	68 (124 Citrus)	Weeks
Out-crossing/ self-incompatible	Grafting	Fruit	Fruit	Dietary fiber, vitamins[b]	72	Months
Hybrid	Runners	Fruit	Fruit	Dietary fiber, vitamins[b]	4	Fresh
Out-crossing	Seeds, runners	Fruit	Fruit	Dietary fiber, vitamins[b]		
Dioecious	Grafting	Berry	Fruit, wine	📝	67	Fresh/ processed
Self-fertile, some self-pollinating	Seed	Fruit	Vegetable	Dietary fiber, vitamins[b]	39	Fresh
Out-crossing/ self-incompatible		Fruit	Vegetable/oil	Oil	18	Months
Self-fertile	Seed	Leaf	Leaf vegetable	Dietary fiber, vitamins[b]	24	Fresh

(*Continued*)

t0010

Table 1.1 (Continued)

Crop	Species	Family	1C Genome size (Mb)[c]	2n Chromosome number	Ploidy level	Life form	Life span	Climatic range
Celery	*Apium graveolens*	Apiaceae	1050	22	2	Herb		Temperate
Papaya	*Carica papaya*[a]	Caricaceae	367	18	2	Tree	Perennial	Tropical
Saffron	*Crocus sativus*	Iridaceae	5770	24	3x	Herb	Perennial	Temperate to mediterranean
Cotton	*Gossypium hirsutum*	Malvaceae	2347	52	4	Woody shrub		Warm temperate
Poplar	*Populus trichocarpa*[a]	Salicaceae	550	38	2	Tree	Perennial	Temperate
Human	*Homo sapiens*	Hominidae	3200	46	2		Perennial	

These crops have been selected intensively by farmers at both the level of choice of species (including new species or hybrids hardly known outside agriculture) and for characteristics including harvestable yield and propagation from diversity within each species. However, few common features related to genome size, chromosome number, and ploidy emerge from the table.

[a]Genome sequenced, public and published by 2011.

[b]Five-a-day: fruit or vegetable with range of properties making it a healthy food, not normally eaten for energy or protein.

[c]1C is the unreplicated haploid DNA content; most DNA contents from angiosperm genome size database, Bennett and Leitch (2011); some from sequencing consortia.

thousand years from the approximately 400,000 plant species. Cereals dominate the list of production figures, and it is clear that the exploitation of the seed has been very important. As a high-energy, harvestable, desiccated, storable, transportable, and robust part of the plant, people have been able to exploit the requirements of the plant for propagation.

p0355 Comparative analysis is extremely informative in most of biology. All of the domesticated species in Table 1.1 share at least some key characteristics related to domestication and selection: an imbalance of parts compared to the wild forms with the harvested part being larger; selection against the dispersal mechanisms most common in the wild (including shattering of pods or inflorescences, dropping of fruits, continuous fruiting, and delayed germination of seeds); ability to establish quickly in single-species stands; or reduction in bitter or other compounds in the harvested part.

p0360 However, the data in Table 1.1 suggest, perhaps surprisingly, few features of large-scale genome organization that have evidence of selection (Heslop-Harrison and Schwarzacher, 2011), despite the intensive selection of genic characteristics as discussed previously. Of the top three cereals, rice has a very small genome and wheat a very large genome; wheat is a hexaploid, rice is diploid, and maize is an ancient tetraploid. Related to selection, it is also notable that the breeding systems or propagation methods (Dwivedi et al., 2010; Charlesworth, 2006), affecting heterozygosity and gene allele population genetics in agriculture, are diverse. For example, among the Poaceae, maize is out-crossing, rye is self-incompatible, and sugarcane vegetatively propagated,

while other major cereals are self-fertile. Thus it seems there are no "rules" about genome structure — size, number of chromosomes, or ploidy — for plants selected as crops.

Superdomestication

s0125

p0365 Breeding of new plant varieties requires genetic variation. This can come from wild collections of germplasm (see Heslop-Harrison, 2002), where extensive seed or plant collections are available for most major crops with allelic variation present in most genes. Many genes with the same function are present in different species, and transgenic approaches mean that genes can be transferred from one species to another. Individual genes that are desirable in a crop can be transferred between species: the gene making the Cry toxin from *Bacillus thuringiensis* giving resistance to many lepidopteran pests in Bt crops is not found in any plants but has been transferred from the bacterium. It is also possible to engineer entire biosynthetic pathways that are missing in one species and desirable for cultivation (e.g., "golden rice" includes the pathway for beta-carotene synthesis in the endosperm, Ye et al., 2000; or to alter oil properties, Damude and Kinney, 2008). New mutations can also be identified as beneficial for crop plants; some of the first genetic changes in domestication were selected by farmers from new mutations. Radiation or chemical mutagenesis has also been helpful in the generation of new genetic variation, and the FAO/IAEA mutant variety database (2010) shows that more than 3000 plant mutant cultivars have been released

Predominant breeding system	Propagation	Parts used	Commodity	Nutritional use	World production 2009 (Million tonnes)	Storage
Self-fertile	Seed	Petiole, root or seed	Vegetable	Dietary fiber, vitamins, spice[b]		Fresh, weeks
cross-pollinating, self-pollinating or parthenocarpic		Fruit	Fruit	Dietary fiber, vitamins[b]	10	Week
Sterile	Bulb	Stigma	Spice	Flavoring		One year
Self-pollinating, but out-crossing possible	Seed	Seed	Fiber/textiles		64	Decades
Out-crossing/ dioecious	Seed	Trunk	Timber/fiber			Decades
Dioecious				Eats and uses the rest		Decades

commercially worldwide by 2010. The range of characteristics covers nearly all breeding traits and has proved useful for correcting weaknesses in existing varieties or generating new characters where there is no accessible variation in germplasm.

p0370   We can expect that synthetic gene construction, random or site-directed mutagenesis, outside the plant cell may increase further availability of gene alleles for specific crop requirements. The current use of green fluorescent proteins (GFP) in plants may provide an indication of the power of this method. The GFP gene used in plant research as a marker for gene expression is modified from that in its source, the jellyfish *Aequorea victoria*, to make it more stable in plants, and several modifications to the coding sequence give variants with different colors and much increased brightness (Chiu et al., 1996). Better understanding of the genetic pathways involved in crops through systems biology (e.g., Kim et al., 2008, 2010) will also be valuable for identifying improvement targets.

p0375   With the understanding of both the genetics and the genomics of crop species, we are now able to develop new crop genotypes incorporating designed characteristics (Vaughan et al., 2007). Farmers will be able to deliver appropriate crops to a growing population by exploitation of appropriate technology and use of the gene pool (Tanksley and McCouch, 1997) — the range of genes present in organisms — and perhaps beyond through synthetic biology approaches. Biotechnology and understanding the behavior of the plant genome provides a range of tools and options that allow crop "superdomestication" (the planning of requirements of new characteristics in our crops).

p0380   In most plants, conventional crossing programs have followed the paradigm of intercrossing pairs of optimum varieties and then selecting progeny following inbreeding for several generations that performed better than either parent. This is summed up in the mantra of "cross the best with the best and hope for the best." In the twenty-first century, crop improvement is accelerating through the use of genetic maps and DNA markers to identify useful variant alleles of genes, to plan recombination between desirable traits, to combine different resistance genes, and accelerate selection, particularly for quantitative traits. Plant breeding is an increasingly targeted and quantitative process.

p0385   An important meta-study by van de Wouw et al. (2010) addressed whether there is a continuing reduction in genetic diversity in crop species. Jarvis and Hodgkin (1999) recognized hybridization with undomesticated lines in many species, thus increasing the diversity in the variation available to plant breeders. Analyzing a large amount of research reported in many papers, with a range of cereal (e.g., Huang et al., 2007), leguminous, and other crops, van de Wouw et al. (2010) showed that in the last century there has been no overall decline in genetic diversity in varieties released over each decade, suggesting that introduction of new germplasm has kept pace with the loss of diversity through inbreeding.

p0390   This volume cannot be divorced from social, economic, and political areas, not least because research can foresee future challenges or problems, and can indicate options for their solution. Farming, whether for food, fuel, or fibers, never assists biodiversity, uses water, leads to erosion, uses crop

protection chemicals, and uses fertilizers. Many of the most pressing problems of mankind are related to plants and the environment, whether for health, food security, or response to climate change. Based on socioeconomic factors, including changes in national and global trade patterns, recognition of requirements for sustainability, nutritional and health needs, and developing crops suitable for changed climates, targets can be set for new crop varieties and occasionally introduction of new crops. Appropriate technologies can then be applied to deliver solutions.

References

Angiosperm Phylogeny Group, An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society*, *161*(2), 105–121.

Anhalt, U. C. M., Heslop-Harrison, J. S., Byrne, S., Guillard, A., & Barth, S. (2008). Segregation distortion in *Lolium*: Evidence for genetic effects. *Theoretical Applied Genetics*, *117*, 297–306. doi: 10.1007/s00122-008-0774-7

Anhalt, U. C. M., Heslop-Harrison, J. S., Piepho, H. P., Byrne, S., & Barth, S. (2009). Quantitative trait loci mapping for biomass yield traits in a Lolium inbred line derived F2 population. *Euphytica*, *170*, 99–107. doi: 10.1007/s10681-009-9957-9

Antonovics, J., Bradshaw, A. D., & Turner, R. G. (1971). Heavy metal tolerance in plants. In J. B. Cragg (Ed.), *Advances in ecological research* (pp. 2–86). London: Academic Press.

Arabidopsis Genome Initiative, Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, *408*(6814), 796–815. Available from: <http://dx.doi.org/10.1038/35048692/>

Argout, X., Salse, J., Aury, J. M., Guiltinan, M. J., Droc, G., & Gouzy, J., et al. (2011, February). The genome of *Theobroma cacao*. *Nature Genetics*, *43*(2), 101–108. Available from: <http://dx.doi.org/10.1038/ng.736/>

Bailey, J. P., Bímová, K., & Mandák, B. (2007). The potential role of polyploidy and hybridisation in the further evolution of the highly invasive *Fallopia taxa* in Europe. *Ecological Research*, *22*, 920–928.

Bassham, James A., & Calvin, M. (1955). *Photosynthesis*. Lawrence Berkeley National Laboratory. Lawrence Berkeley National Laboratory. LBNL Paper UCRL-2853. Retrieved from: <http://escholarship.org/uc/item/0j6008b4/>

Buckler, E. S., IV, Thornsberry, J. M., & Kresovich, S. (2001, June). Molecular diversity, structure and domestication of grasses. *Genetical Research*, *77*(3), 213–218. Available from: <http://view.ncbi.nlm.nih.gov/pubmed/11486504/>

Caldwell, K. S., Dvorak, J., Lagudah, E. S., Akhunov, E., Luo, M. C. C., & Wolters, P., et al. (2004, June). Sequence polymorphism in polyploid wheat and their D-genome diploid ancestor. *Genetics*, *167*(2), 941–947. Available from: <http://dx.doi.org/10.1534/genetics.103.016303/>

Charlesworth, D. (2006, September 5). Evolution of plant breeding systems. *Current Biology*, *16*, R726–R735. DOI 10.1016/j.cub.2006.07.068

Chiu, W. I., Niwa, Y, Zeng, W., Hirano, T., Kobayashi, H., & Sheen, J. (1996, March). Engineered GFP as a vital reporter in plants. *Current Biology*, *6*(3), 325–330. Available from: <http://www.cell.com/current-biology/abstract/S0960-9822(02)00483-9/>

Chory, J., Ecker, J. R., Briggs, S., Caboche, M., Coruzzi, G. M., & Cook, D., et al. (2000). National Science Foundation-sponsored workshop report: "The 2010 project functional genomics and the virtual plant. a blueprint for understanding how plants are built and how to improve them". *Plant Physiology*, *123*, 423–426. doi:10.1104/pp.123.2.423

Cifuentes, M., Eber, F., Lucas, M. O., Lode, M., Chevre, A. M., & Jenczewski, E. (2010, July). Repeated polyploidy drove different levels of crossover suppression between homoeologous chromosomes in *Brassica napus* allohaploids. *The Plant Cell*, *22*(7), 2265–2276. Available from: <http://dx.doi.org/10.1105/tpc.109.072991/>

Clark, R. M., Linton, E., Messing, J., & Doebley, J. F. (2004, January). Pattern of diversity in the genomic region near the maize domestication gene *tb1*. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(3), 700–707. Available from: <http://dx.doi.org/10.1073/pnas.2237049100/>

Clottes, J. 2010. Chauvet Cave (ca. 30,000 B.C.). (2000). In *Heilbrunn Timeline of Art History*. New York: The Metropolitan Museum of Art. <http://www.metmuseum.org/toah/hd/chav/hd_chav.htm/> Accessed 09.10.

Cong, B., Liu, J., & Tanksley, S. D. (2002, October). Natural alleles at a tomato fruit size quantitative trait locus differ by heterochronic regulatory mutations. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(21), 13606–13611. Available from: <http://dx.doi.org/10.1073/pnas.172520999/>

Crawley, M. J., & Brown, S. L. (1995). Seed limitation and the dynamics of feral Oilseed Rape on the M25 motorway. *Proceedings of the Royal Society of London Series B: Biological Sciences*, *259*(1354), 49–54.

Available from: <http://dx.doi.org/10.1098/rspb.1995.0008/>

Damude, H. G., & Kinney, A. J. (2008). Enhancing plant seed oils for human nutrition. *Plant Physiology*, *147*, 962–968. doi:10.1104/pp.108.121681

De Langhe, E., Hřibová, E., Carpentier, S., Doležel, J., & Swennen, R. (2010, December). Did backcrossing contribute to the origin of hybrid edible bananas? *Annals of Botany*, *106*(6), 849–857. Available from: <http://dx.doi.org/10.1093/aob/mcq187/>

Doebley, J. (2004). The genetics of maize evolution. *Annual Review of Genetics*, *38*(1), 37–59. Available from: <http://dx.doi.org/10.1146/annurev.genet.38.072902.092425/>

Doebley, J. F., Gaut, B. S., & Smith, B. D. (2006, December). The molecular genetics of crop domestication. *Cell*, *127*(7), 1309–1321. Available from: <http://dx.doi.org/10.1016/j.cell.2006.12.006/>

Doust, A. (2007, October). Architectural evolution and its implications for domestication in grasses. *Annals of Botany*, *100*(5), 941–950. Available from: <http://dx.doi.org/10.1093/aob/mcm040/>

Dubcovsky, J., & Dvorak, J. (2007, June). Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science*, *316*(5833), 1862–1866. Available from: <http://dx.doi.org/10.1126/science.1143986/>

Dwivedi, S., Perotti, E., Upadhyaya, H., & Ortiz, R. (2010, December). Sexual and apomictic plant reproduction in the genomics era: Exploring the mechanisms potentially useful in crop plants. *Sexual Plant Reproduction*, *23*(4), 265–279. Available from: <http://dx.doi.org/10.1007/s00497-010-0144-x/>

FAOstat (2010). *Production data relating to food and agriculture*. <http://faostat.fao.org/>.

FAO/IAEA mutant variety database (2010). <http://mvgs.iaea.org/aboutMutantVarieties.aspx/>

French-Italian Public Consortium for Grapevine Genome Characterization, The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, *449*, 463–468. doi:10.1038/nature06148

Gailing, O., Vornam, B., Leinemann, L., & Finkeldey, R. (2009, December). Genetic and genomic approaches to assess adaptive genetic variation in plants: Forest trees as

a model. *Physiologia Plantarum*, *137*(4), 509–519.

TI1   Gross, B. L., & Olsen, K. M. (2010, September). Genetic perspectives on crop domestication. *Trends in Plant Science*, *15*(9), 529–537. Available from: <http://dx.doi.org/10.1016/j.tplants.2010.05.008/>

Guo, J., Wang, Y., Song, C., Zhou, J., Qiu, L., & Huang, H., et al. (2010, September). A single origin and moderate bottleneck during domestication of soybean (Glycine max): Implications from microsatellites and nucleotide sequences. *Annals of Botany*, *106*(3), 505–514. Available from: <http://dx.doi.org/10.1093/aob/mcq125/>

Gurevitch, J., & Padilla, D. (2004, September). Are invasive species a major cause of extinctions? *Trends in Ecology & Evolution*, *19*(9), 470–474. Available from: <http://dx.doi.org/10.1016/j.tree.2004.07.005/>

Hammer, K. (1984). Das Domestikations-syndrom. *Kulturpflanze*, *32*, 11–34.

Harlan, J., & de Wet, J. (1965 January). Some thoughts about weeds. *Economic Botany*, *19*(1), 16–24. Available from: <http://dx.doi.org/10.1007/BF02971181/>

Heaton, E. A., Dohleman, F. G., & Long, S. P. (2008). Meeting US biofuel goals with less land: The potential of *Miscanthus*. *Global Change Biology*, *14*, 2000–2014.

Heslop-Harrison, J. S. (2002). Exploiting novel germplasm. *Australian Journal of Agricultural Research*, *53*(8), 873–879.

Heslop-Harrison, J. S. (2004). Biotechnology of fruit and nut crops: Introduction. In R. E. Litz (Ed.), *Biotechnology of Fruit and Nut Crops* (pp. xix–xxiv). CABI.

Heslop-Harrison, J. S. (2010). Genes in evolution: The control of diversity and speciation. *Annals of Botany*, *106*(3), 437–438. Available from: <http://dx.doi.org/10.1093/aob/mcq168/>

Heslop-Harrison, J. S., & Schwarzacher, T. (2007). Domestication, genomics and the future for banana. *Annals of Botany*, *100*(5), 1073–1084. doi: 10.1093/aob/mcm191

Heslop-Harrison, J. S., & Schwarzacher, T. (2011). Organisation of the plant genome in chromosomes. *The Plant Journal*, *66*(1), 18–33. Available from: <http://dx.doi.org/10.1111/j.1365-313X.2011.04544.x/>

Heyduck, R. F., Baltensperger, D. D., Nelson, L. A., & Graybosch, R. A. (2008). Yield and agronomic traits of waxy proso in the central Great Plains. *Crop Science*, *48*(2), 741+. Available from: <http://dx.doi.org/10.2135/cropsci2007.02.0081/>

Huang, X. Q., Wolf, M., Ganal, M. W., Orford, S., Koebner, R. M. D., & Röder, M. S. (2007). Did modern plant breeding lead to genetic erosion in european winter wheat varieties? *Crop Science*, *47*(1), 343+. Available from: <http://dx.doi.org/10.2135/cropsci2006.04.0261/>

Hunt, H. V., Vander Linden, M., Liu, X., Motuzaite-Matuzeviciute, G., Colleridge, S., & Jones, M. K. (2008). Millets across Eurasia: Chronology and context of early records of the genera *Panicum* and *Setaria* from archaeological sites in the Old World. *Vegetation History and Archaeobotany*, *17*, 5–18.

Hyten, D. L., Song, Q., Zhu, Y., Choi, I. Y., Nelson, R. L., & Costa, J. M., et al. (2006 November). Impacts of genetic bottlenecks on soybean genome diversity. *Proceedings of the National Academy of Sciences*, *103*(45), 16666–16671. Available from: <http://dx.doi.org/10.1073/pnas.0604379103/>

Innan, H., & Kim, Y. (2008 July). Detecting local adaptation using the joint sampling of polymorphism data in the parental and derived populations. *Genetics*, *179*(3), 1713–1720. Available from: <http://dx.doi.org/10.1534/genetics.108.086835/>

Janick, J. (2005). The origin of fruits, fruit growing, and fruit breeding. *Plant Breeding Reviews*, *25*, 255–320.

Jarvis, D. I., & Hodgkin, T. (1999). Wild relatives and crop cultivars: Detecting natural introgression and farmer selection of new genetic combinations in agroecosystems. *Molecular Ecology*, *8*, S159–S173.

Kim, J., Kim, T. -G., Jung, S. H., Kim, J. -R., Park, T., & Heslop-Harrison, P., et al. (2008). Evolutionary design principles of modules that control cellular differentiation: Consequences for hysteresis and multistationarity. *Bioinformatics*, *24*(13), 1516–1522. doi: 10.1093/bioinformatics/btn229

Kim, T. H., Kim, J., Heslop-Harrison, P., & Cho, K. H. (2010). Evolutionary design principles and functional characteristics based on kingdom-specific network motifs. *Bioinformatics* <http://dx.doi.org/10.1093/bioinformatics/btq633/>

King D., Heslop-Harrison J. S., & 23 other members of the GM Science Review Panel. (2003). GM science review: An open review of the science relevant to GM crops and food based on the interests and concerns of the public. <www.gmsciencedebate.org.uk/>. 296pp

King D., Heslop-Harrison J. S., & 25 other members of the GM Science Review Panel. (2004). GM science review: Second Report. An open review of the science relevant to GM crops and food based on the interests and concerns of the public. Full text from: <http://gmsciencedebate.org.uk/>. 116pp

Koornneef, M., van Eden, J., Hanhart, C. J., Stam, P., Braaksma, F. J., & Feenstra, W. J. (1983, July). Linkage map of Arabidopsis thaliana. *Journal of Heredity*, *74*(4), 265–272. Available from: <http://jhered.oxfordjournals.org/content/74/4/265.abstract/>

Martin, C., Ellis, N., & Rook, F. (2010, October). Do transcription factors play special roles in adaptive variation? *Plant Physiology*, *154*(2), 506–511. Available from: <http://dx.doi.org/10.1104/pp.110.161331/>

McClintock, B., Kato, T. A., & Blumenschein, A. (1981). Chromosome constitution of races of maize. Its significance in the interpretation of relationships between races and varieties in the Americas. Chapingo, Mexico: Colegio de Postgraduados.

Millenium Ecosystem Assessment. (2010). Guide to the Millennium Assessment Reports. <http://www.maweb.org/en/Index.aspx/>

Molnar, I., Cifuentes, M., Schneider, A., Benavente, E., & Molnar-Lang, M. (2011). Association between simple sequence repeat-rich chromosome regions and intergenomic translocation breakpoints in natural populations of allopolyploid wild wheats. *Annals of Botany* <http://dx.doi.org/10.1093/aob/mcq215/>   AQ3

Moose, S. P., & Mumm, R. H. (2008). Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiology*, *147*, 969–977. doi: 10.1104/pp.108.118232

Neale, D. B. (2007, December). Genomics to tree breeding and forest health. *Current Opinion in Genetics & Development*, *17*(6), 539–544.

Nielsen, R., Williamson, S., Kim, Y., Hubisz, M. J., Clark, A. G., & Bustamante, C. (2005, November). Genomic scans for selective sweeps using SNP data. *Genome Research*, *15*(11), 1566–1575. Available from: <http://dx.doi.org/10.1101/gr.4252305/>

Papa, R., Bellucci, E., Rossi, M., Leonardi, S., Rau, D., & Gepts, P., et al. (2007, October). Tagging the signatures of domestication in common bean (*Phaseolus vulgaris*) by means of pooled DNA samples. *Annals of Botany*, *100*(5), 1039–1051. Available from: <http://dx.doi.org/10.1093/aob/mcm151/>

Paris, H. S., Daunay, M. C., & Janick, J. (2009, June). The Cucurbitaceae and Solanaceae illustrated in medieval manuscripts known as the Tacuinum Sanitatis. *Annals of Botany*, *103*(8), 1187–1205. Available from: <http://dx.doi.org/10.1093/aob/mcp055/>

Paterson, A. H., Lin, Y. R., Li, Z., Schertz, K. F., Doebley, J. F., & Pinson, S. R. M., et al. (1995). Convergent domestication of cereal crops by independent mutations. *Science*, *269*, 1714–1718.

Richardson, J. A. (1975). Physical problems of growing plants on colliery waste. M. J. Chadwick & G. T. Goodman (Eds.), *The ecology of resource degradation and renewal. The fifteenth symposium of the british ecological society*, Leeds, 10-12 July 1973 (pp. 275-286). Oxford: Blackwell Scientific Publications.

Rieseberg, L. H., & Blackman, B. K. (2010, September). Speciation genes in plants. *Annals of Botany*, *106*(3), 439–455. Available from: <http://dx.doi.org/10.1093/aob/mcq126/>

Rubin, E. M. (2008, August 14). Genomics of cellulosic biofuels. *Nature*, *454*, 841–845. doi: 10.1038/nature07190

Saeidi, H., Rahiminejad, M. R., & Heslop-Harrison, J. S. (2008). Retroelement insertional polymorphisms, diversity and phylogeography within diploid, D-Genome Aegilops tauschii (Triticeae, Poaceae) sub-taxa in Iran. *Annals of Botany*, *101*(6), 855–861. doi:10.1093/aob/mcn042

Salamini, F., Ozkan, H., Brandolini, A., Scafer-Pregl, R., & Martin, W. (2002). Genetics and geograph of wild cereal domestication in the near East. *Nature Reviews Genetics*, *3*, 429.

Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T., & Katayose, Y., et al. (2002, November). The genome sequence and structure of rice chromosome 1. *Nature*,

420(6913), 312–316. Available from: <http://dx.doi.org/10.1038/nature01184/>

Schulze, J., Stoll, P., Widmer, A., & Erhardt, A. (2011). Searching for gene flow from cultivated to wild strawberries in Central Europe. *Annals of Botany*, *107*(4), 699–707. Available from: <http://dx.doi.org/10.1093/aob/mcr018/>

Somerville, C. R. (2010, August 13). Feedstocks for lignocellulosic biofuels. *Science*, 790–792. DOI: 10.1126/science.1189268.

Tang, T., Lu, J., Huang, J., He, J., McCouch, S. R., & Shen, Y., et al. (2006, November). Genomic variation in rice: Genesis of highly polymorphic linkage blocks during domestication. *PLoS Genetics*, *2*(11), e199 + . Available from: <http://dx.doi.org/10.1371/journal.pgen.0020199/>

Tanksley, S. D., & McCouch, S. R. (1997, August). Seed banks and molecular maps: Unlocking genetic potential from the wild. *Science*, *277*(5329), 1063–1066. Available from: <http://dx.doi.org/10.1126/science.277.5329.1063/>

Tian, Z., Wang, X., Lee, R., Li, Y., Specht, J. E., & Nelson, R. L., et al. (2010, May). Artificial selection for determinate growth habit in soybean. *Proceedings of the National Academy of Sciences*, *107*(19), 8563–8568. Available from: <http://dx.doi.org/10.1073/pnas.1000088107/>

Varshney, R. K., Hoisington, D. A., & Tyagi, A. K. (2006). Advances in cereal genomics and applications in crop breeding. *Trends in Biotechnology*, *24*(11), 490–499. doi: 10.1016/j.tibtech.2006.08.006

Vaughan, D. A., Balázs, E., & Heslop-Harrison, J. S. (2007). From crop domestication to super-domestication. *Annals of Botany*, *100*(5), 893–901. doi: 10.1093/aob/mcm224

Vigouroux, Y., Mitchell, S., Matsuoka, Y., Hamblin, M., Kresovich, S., & Smith, J.

S. C., et al. (2005, March). An analysis of genetic diversity across the maize genome using microsatellites. *Genetics*, *169*(3), 1617–1630. Available from: <http://dx.doi.org/10.1534/genetics.104.032086/>

Wang, H., Nussbaum-Wagler, T., Li, B., Zhao, Q., Vigouroux, Y., & Faller, M., et al. (2005, August). The origin of the naked grains of maize. *Nature*, *436*(7051), 714–719. Available from: <http://dx.doi.org/10.1038/nature03863/>

Wang, J. X., Gao, T. G., & Knapp, S. (2008 December). Ancient chinese literature reveals pathways of eggplant domestication. *Annals of Botany*, *102*(6), 891–897. Available from: <http://dx.doi.org/10.1093/aob/mcn179/>

Weeden, N. F. (2007 October). Genetic changes accompanying the domestication of pisum sativum: Is there a common genetic basis to the 'Domestication Syndrome' for Legumes?. *Annals of Botany*, *100*(5), 1017–1025. Available from: <http://dx.doi.org/10.1093/aob/mcm122/>

Wouw, M. van de, van Hintum, T., Kik, C., van Treuren, R., & Visser, B. (2010, April). Genetic diversity trends in twentieth century crop cultivars: A meta analysis. *TAG Theoretical and Applied Genetics Theoretische und angxewandte Genetik*, *120*(6), 1241–1252. Available from: <http://dx.doi.org/10.1007/s00122-009-1252-6/>

Wright, S. I., Bi, I. V., Schroeder, S. G., Yamasaki, M., Doebley, J. F., & McMullen, M. D., et al. (2005 May). The effects of artificial selection on the maize genome. *Science*, *308*(5726), 1310–1314. Available from: <http://dx.doi.org/10.1126/science.1107891/>

Yamasaki, M., Tenaillon, M. I., Vroh, Bi I, Schroeder, S. G., Sanchez-Villeda, H., & Doebley, J. F., et al. (2005 November). A large-scale screen for artificial selection in maize identifies candidate agronomic loci

for domestication and crop improvement. *The Plant Cell*, *17*(11), 2859–2872. Available from: <http://dx.doi.org/10.1105/tpc.105.037242/>

Yamasaki, M., Wright, S. I., & McMullen, M. D. (2007, October). Genomic screening for artificial selection during domestication and improvement in maize. *Annals of Botany*, *100*(5), 967–973. Available from: <http://dx.doi.org/10.1093/aob/mcm173/>

Yano, M. (2001). Genetic and molecular dissection of naturally occurring variation. *Current Opinion in Plant Biology*, *4*, 130–135.

Ye, X., Al-Babili, S., Klöti, A., Zhang, J., Lucca, P., & Beyer, P., et al. (2000). Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science*, *287*(5451), 303–305. doi:10.1126/science.287.5451.303

Zeder, M. A. (2006). *Documenting domestication: New genetic and archaeological paradigms*. University of California Press.

Zhang, Y., Wang, J., Zhang, X., Chen, J. Q., Tian, D., & Yang, S. (2009 April). Genetic signature of rice domestication shown by a variety of genes. *Journal of Molecular Evolution*, *68*(4), 393–402. Available from: <http://dx.doi.org/10.1007/s00239-009-9217-6/>

Zhu, Q., Zheng, X., Luo, J., Gaut, B. S., & Ge, S. (2007, March). Multilocus analysis of nucleotide variation of *Oryza sativa* and its wild relatives: Severe bottleneck during domestication of rice. *Molecular Biology and Evolution*, *24*(3), 875–888. Available from: <http://dx.doi.org/10.1093/molbev/msm005/>

Uncited references

p0400 Molnar et al. (2011); Somerville (2010))