Chapter 12
Molecular Genetic Basis of the Domestication Syndrome in Cereals

Tao Sang and Jiayang Li

12.1 Introduction

Plant and animal domestication that was initiated approximately 10,000 years ago led to the dramatic evolution of human society and rapid speciation of plants and animals co-evolving with humans (Diamond 2002). The emergence of these new species and remarkable new traits fueled a series of scientific discoveries. The observation of rapid and drastic phenotypic changes under artificial selection stimulated at least partly Darwin's thinking of the origin of species under natural selection (Darwin 1859). Through experimental crosses and subsequent analyses of crop species, particularly pea plants, Mendel discovered the basic rules of genetics. With the arrival of the genomics era, recent studies yielded considerable new insights into the molecular basis of domestication traits and population genetic mechanisms underlying the domestication processes.

Cereal crops, including wheat, rice, maize, barley, sorghum, oats, and millets, provide the primary source of human calories. Of approximately 1.4 billion hectares or ~10 % of the terrestrial ecosystems converted to cropland, about half or ~0.7 billion hectares are currently used for producing cereals. The top three cereal crops, maize, rice, and wheat, grow on ~0.55 billion hectares (http://faostat.fao.org).

Cereals belong to the grass family, Poaceae, which consists of ~10,000 species worldwide. The domestication of cereals occurred independently in different

T. Sang (✉)
State Key Laboratory of Systematic and Evolutionary Botany, Key Laboratory of Plant Resources and Beijing Botanical Garden, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China
e-mail: sang@ibcas.ac.cn

J. Li
State Key Laboratory of Plant Genomics and National Center for Plant Gene Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China
e-mail: jyli@genetics.ac.cn

P. K. Gupta and R. K. Varshney (eds.), *Cereal Genomics II*,
DOI: 10.1007/978-94-007-6401-9_12,

continents, e.g., wheat and barley in Middle East ~10,000 years ago, rice and foxtail millet in China ~8,000 years ago, maize in Central America ~7,000–9,000 years ago, and sorghum and pearl millet in Africa ~4,000 years ago (Salamini et al. 2002; Doebley et al. 2006). With regard to the seemingly coincidental initiation of cereal domestications, an increasingly popular hypothesis is that they shared at least one common driving force, i.e., climate change following the last glacial maximum (Sage 1995; Richerson et al. 2001; Cunniff et al. 2008). The consequential change of global vegetation could have provided opportunities for humans to explore expanding grassland and developed a more reliable food source for feeding growing populations that faced increasing pressure of food shortage.

Despite the fact that the cereal crops were independently domesticated from distantly related grass species, the phenotypic modifications associated with the domestications were strikingly similar. The suite of phenotypic changes that transformed wild grasses into food crops are known as the domestication syndrome (Harlan 1992; Hancock 2004). This includes the following major trends: reduced grain shattering, improved threshing ability, weakened seed dormancy, reduced tiller number or shoot branches, synchronized grain maturation, lightened hull and seed colors, increased tiller erectness, enlarged panicles, and increased grain size and weight. Through these changes driven by artificial selection, plants became adapted to new environments and relied on humans for survival and reproduction.

In this chapter, we review the identification and molecular and population genetic analyses of genes involved in the development of the cereal domestication syndrome. With the literature review, we revisit the generalized theories and hypotheses concerning crop domestication. We then comment on the current status and future prospects of discovering domestication-related genes. From what has been learned about the molecular and population genetics of cereal domestication, we discuss how this information may be utilized to facilitate the domestication of new crops, especially perennial grasses potentially serving as lignocellulosic energy crops.

12.2 Domestication Syndrome and Genetic Analyses

Mature grains of wild grasses detach easily when disturbed by wind or animals so that their seeds can be dispersed promptly. The easy shattering of wild plants, however, makes grain harvest difficult, especially when they mature over a relatively long period of time during which substantial grain loss occurs as a result of strong wind or storms. Thus the reduction of grain shattering was necessary for effective harvest. Threshing was developed in association with the reduction in shattering and allowed grains or seeds to be separated and recovered from straw. There are two primary ways of threshing. In rice, mature grains are detached from harvested panicles by mechanistic forces, which is more or less equivalent to the partial reduction of shattering. The hulls are subsequently removed by milling. In wheat and barley, hulls open and separate from seeds at grain maturation so that seeds can be shaken out with moderate forces and collected.

Seed dormancy is an evolutionary safety strategy of wild grasses adapted to unstable environments. It allows seeds to remain in seed banks through the winter and germinate under favorable climatic conditions in the following spring. It also distributes seed germination over a prolonged period of time to avoid unpredicted harmful conditions such as flood and frost. However, this unwanted delay or uneven germination makes crops grow and mature at different rates, which in turn causes problems in crop management and harvest.

The loss of hulls and their appendices such as the stony fruitcases of maize kernels and awns of rice made food preparation easier but weakened protection against seed predators and ability of seed dispersal. The changes of hull colors, usually from dark to straw-white, might have been helpful for non-shattering grains to avoid predation at maturity while hull color becomes undistinguishable from that of withering straws (Zhu et al. 2011). Together, these phenotypic modifications changed the landscape of adaptation, resulting in well-adapted crop plants in the agricultural field, which could no longer survive in nature without human protection or assistance for reproduction.

The modification of plant architecture is another critical aspect of cereal domestication. This followed a common trend of reduction in tiller number or branches so that each tiller became stronger and capable of supporting a higher grain yield. Tillers also grew erect to allow a larger number of tillers to be compacted in a unit field. Panicles became more highly branched and capable of bearing a larger number of grains. These changes were clearly driven by selection for higher yield.

Domestication is evolution under artificial selection. The domestication syndrome highlights phenotypic changes that shifted the adaptive optima of wild grasses in their natural habitats to that of the new crops species in the agricultural field. To gain the mechanistic understanding of the process, we need to identify the genetic basis of these changes, which includes the number and chromosomal locations of loci/genes responsible for a domestication transition, the phenotypic effect of these loci/genes, and ultimately the casual mutations.

There are two main approaches for studying the genetic basis of domestication traits. A straightforward approach is to cross cultivars with their wild progenitors or relatives and subsequently conduct a quantitative trait locus (QTL) analysis. This so called “top-down” approach has been wildly adopted in crops because plants are relatively easy to hybridize. Numerous QTLs potentially underlying domestication transitions in almost all cereal crops have been reported. Another seemingly promising approach is to perform a genome-wide screening for signatures of artificial selection. Through a comparison of the distribution of nucleotide polymorphism between cultivars and their wild progenitors, loci/genes that presumably experienced selective sweeps are considered to be candidates involved in domestication. This is called a “bottom-up” approach (Ross-Ibarra et al. 2007).

In the first edition of *Cereal Genomics* published in 2004, domestication traits and corresponding QTLs were discussed (Pozzi et al. 2004). While the detected domestication-related QTLs continued to accumulate at a fast pace, a large portion of them have been deposited in user-friendly databases such as Gramene

(Youens-Clark et al. 2011). Over the past few years, the explosive development of grass genome resources and tools has considerably accelerated the molecular cloning of domestication-related QTLs, which have yielded markedly new insights into the genetic and evolutionary mechanisms of crop domestication. In addition, the advent of the next-generation sequencing technologies began to show a great potential of identifying candidate genes using the bottom-up approach (e.g., Xia et al. 2009). Thus here we choose not to update the growing list of identified QTLs, rather focus on the cloned QTLs/genes for a better understanding of cereal domestication.

12.3 Genes Underlying Domestication Syndrome

To date, more than a dozen of genes directly relevant to the development of the cereal domestication syndrome have been cloned (Table 12.1). Although many more genes identified in cereals were functionally related to the domestication traits, there was little evidence that they were directly responsible for the changes occurred during domestication. Domestication genes considered in this chapter share several features. First, the majority of them were cloned through the top-down approach, i.e., QTL analyses of crosses between cultivars and their wild relatives or between cultivars of independent origins, followed by fine mapping. Second, there was population genetic evidence indicating that the genes were under artificial selection. Finally, all but *tb1* were identified after 2005, which was apparently a result of recent advances in cereal genomics.

12.3.1 Shattering and Threshing

During cereal domestication, reduction in shattering was coupled with the maintenance or gain of threshing ability. The two traits together ensured grains to be effectively harvested and subsequently recovered. In rice, the balance of shattering and threshing was struck by weakening the function of the *sh4* gene that regulates the formation and function of the abscission zone from which a mature grain detaches from the pedicle (Li et al. 2006a). *Sh4* is the major shattering QTL in rice, explaining ~69 % of phenotypic variance between a traditional *indica* cultivar and the annual wild progenitor, *O. nivara* (Li et al. 2006b). The gene encodes a putative transcription factor. The causal mutation was a single nucleotide substitution leading to an amino acid substitution from lysine to asparagine in the predicted DNA binding domain. The substitution of the neutral for the positively charged amino acid, which presumably weakened but did not knock out the gene function, caused the incomplete development and partial function of the abscission zone (Li et al. 2006a). This disabled the natural detachment of grains necessary for seed dispersal in the wild species, but still allowed manual separation of grains from pedicels in cultivars during harvest.

Table 12.1 Genes controlling domestication traits of cereals

Domestication trait	Gene	Crop	Gene encoding	Causal mutation[a]	Origin	Fixation[b]	Year of publication
Shattering, threshing	*sh4*	Rice	Transcription factor	Trans-modification	Single	Fixed	2006
	qSH1	Rice	Transcription factor	Cis- modification	Single	Not fixed	2006
	Q	Wheat	Transcription factor	Trans-modification	Single	Fixed	2006
	nud	Barley	Transcription factor	Loss-of-function	Single	Fixed	2008
Plant architecture, inflorescence structure	*tb1*	Maize	Transcription factor	Cis-modification	Single	Fixed	1997
	prog1	Rice	Transcription factor	Tran- and Cis- modification	Single	Fixed	2008
	vrs1	Barley	Transcription factor	Loss-of-function	Multiple	Not fixed	2007
	Q	Wheat	Transcription factor	Trans-modification	Single	Fixed	2006
Grain/seed cover, size, and coloration	*tga1*	Maize	Transcription factor	Tran-modification	Single	Fixed	2005
	Bh4	Rice	Amino acid transporter	Loss-of-function	Multiple	Fixed	2011
	rc	Rice	Transcription factor	Loss-of-function	Multiple	Not fixed	2006
	rd	Rice	Enzyme	Loss-of-function	Multiple	Not fixed	2006
	phr1	Rice	Enzyme	Loss-of-function	Multiple	Not fixed	2008
	qSW5	Rice	Unknown	Loss-of-function	Single	Not fixed	2008
	GS3	Rice	Unknown	Loss-of-function	Single	Not fixed	2006

[a]The types of mutations considered here include loss-of-function and functional modification. For functional modification, a mutation occurring in the coding and regulatory region of a transcription factor is called trans- and cis-modification, respectively

[b]Domestication allele(s) that were found in all cultivars surveyed so for are considered to be fixed

Another shattering QTL, *qSH1*, accounting for ~69 % of phenotypic variance between *indica* and temperate *japonica* cultivars, was also cloned (Konishi et al. 2006). The causal mutation was a nucleotide substitution in the regulatory element located ~12 kb upstream of the coding region of a homeobox gene, which altered the level and pattern of the gene expression and disrupted the development of the abscission zone. The mutation was found in a portion of temperate *japonica* rice.

In barley, the derivation of non-shattering phenotype, also known as non-brittle rachis, was controlled primarily by two tightly linked loci, *btr1* and *btr2*. The homozygous recessive genotype at one of the loci, *btr1btr1/Btr2Btr2* or *Btr1Btr1/btr2btr2*, confers the non-brittle phenotype. Cultivars from the western parts of the world have predominantly the *btr1btr1/Btr2Btr2* genotype, while most of eastern cultivars have the *Btr1Btr1/btr2btr2* genotype. Although neither locus has been identified at the genic level, phylogenetic analysis of DNA sequences of the flanking regions showed that the eastern and western cultivars formed their own groups, indicating independent origins of non-brittle rachis from the eastern and western regions (Azhanguvel and Komatsusa 2007). The double homozygous recessive genotype *btr1btr1/btr2btr2*, however, has not been found in any barley cultivars and the linkage has never been broken up in experimental crosses (Komatsuda et al. 2004). It was hypothesized that *btr1* and *btr2* might be different mutations of the same gene (Sang 2009).

In tetraploid wheat with AABB genomic constitution, non-brittle rachis is controlled largely by two loci, Br_2 and Br_3, located in the homoeologous regions of group 3 chromosomes (Watanabe et al. 2002). They are potentially orthologous loci between the AA and BB genomes of the diploid parents. In hexaploid bread wheat, there is an additional brittle rachis locus, Br_1, also mapped to the orthologous location of group 3 chromosome, 3D, of the DD genome (Nalam et al. 2006; Watanabe et al. 2006). Furthermore, comparative mapping showed that this chromosomal region of wheat might be orthologous to that of barley containing *btr1* and *btr2* (Nalam et al. 2006; Pourkheirandish and Komatsuda 2007). The region is not orthologous to either of those harboring the rice shattering genes, *sh4* or *qSH1* (Li and Gill 2006; Sang 2009).

Unlike rice where grains are recovered from straw through threshing, seeds of free-threshing barley and wheat are directly removed from hulls that remain on straws. This required additional mutations that allowed easy release of seeds from hulls. The allele, *nud*, conferring free-threshing was cloned in barley (Taketa et al. 2008). In the wild progenitors of barley, the gene encodes an ethylene response factor that regulates lipid biosynthesis in the seed coat, which produces adhesive lipid between seed coats and hulls. A 17 kb deletion in the chromosome region containing *Nud* is responsible for the disruption of the lipid layer and consequently easy releasing of seeds from hulls in cultivated barley.

In wheat, free-threshing was achieved through the appearance of softened and easily separable hulls. Hulls of the free-threshing cultivars could open easily to release seeds under moderate forces such as beating or grinding. Genetic analysis between durum wheat and the wild progenitor of emmer wheat identified four QTLs for free-threshing (Simonetti et al. 1999). Of these, two with large

effect, each accounting for ~25 % of phenotypic variation, were *Tg* on the short arm of chromosome 2B and *Q* on the long arm of chromosome 5A. The free-threshing alleles, *tg* and *Q*, are partially recessive and partially dominant, respectively. The free-threshing tetraploid wheat has a genotype of $tgtg^{2B}QQ^{5A}$. In the hexaploid bread wheat, there is an additional recessive mutation at the *Tg* locus of the DD genome required for free-threshing, resulting in the genotype of $tgtg^{2B}tgtg^{2D}QQ^{5A}$ (Jantasuriyarat et al. 2004; Nalam et al. 2007).

Molecular cloning of *Q* showed that it is a gene belonging to the AP2 family of transcription factors (Simons et al. 2006). The *Q* allele had a higher level of transcription than the wild type allele, *q*, in spikes, leaves, and roots. The coding regions of the two alleles differed by an amino acid substitution, which was responsible for an increased abundance of homodimer of Q protein when tested in yeast. This mutation in the coding region, together with regulatory mutations potentially including a substitution at the microRNA binding site (Chuck et al. 2007), led to the gain-of-function mutation of *Q* that conferred the free-threshing phenotype. Interestingly, *Q* also contributes to the toughness of rachis or reduced shattering.

12.3.2 Plant Architecture and Inflorescence Structure

Maize has undergone the most drastic morphological modifications among all cereals. It involved the development of a single stalk from highly branched shoots of the wild progenitor, teosinte. A gene, *tb1*, controlling the difference was cloned using a maize mutant resembling the shoot branching pattern of teosinte (Doebley et al. 1997). It was confirmed that *tb1* was allelic to the major QTL underlying the architectural transition from teosinte to maize. The gene is a member of the TCP family of transcriptional regulators involved in the transcriptional regulation of cell cycle genes. In maize, *tb1* confers apical dominance by repressing the outgrowth of axillary meristems and branch elongation through its repressive effect on the cell cycle (Doebley et al. 2006). The causal mutations are located in the regulatory regions of the gene that alter the pattern and level of gene expression (Wang et al. 1999). The maize allele of *tb1* was highly expressed in the axillary buds whereas the teosinte allele showed no sign of expression (Hubbard et al. 2002).

In rice, a similar but less dramatic change occurred in plant architecture. In comparison to the wild progenitors, *O. nivara* and *O. rufipogon*, cultivated rice has fewer and more erect tillers. This architectural change has allowed cultivars to more effectively capture solar radiation and to be planted more densely in the field, both of which contribute to higher yield. *Prostrate growth 1* (*prog1*), responsible for this transition, was cloned using near-isogenic lines of rice cultivars that contained a small region of the short arm of chromosome 7 from *O. rufipogon* (Jin et al. 2008; Tan et al. 2008). The gene encodes a zinc-finger nuclear transcription factor and is predominantly expressed in the axillary meristems, from which tiller buds form. The causal mutation included primarily an amino acid substitution that weakens or disrupts the function of the gene and possibly those in its regulatory region as well.

Transgenic experiments showed that cultivated rice containing the wild-type allele from *O. rufipogon* had not only a larger number of more prostrate tillers but also shorter tillers with panicle bearing fewer primary and secondary branches and thus fewer grains. The pleiotropic effect of the gene matched almost perfectly the effect of a set of QTLs identified from a cross between an *indica* cultivar and *O. nivara* (Li et al. 2006b) and a cross between a *japonica* cultivar and *O. rufipogon* (Onishi et al. 2007). These QTLs, overlapped with *prog1* on the short arm of chromosome 7, had the largest effect on almost all morphological traits, including plant height, tiller number, tiller angle, and the number of primary and secondary branches of a panicle (Li et al. 2006b). If *prog1* is indeed allelic to the QTL for all of these traits, it represents one of the most important genes involved in the improvement of plant architecture and yield during rice domestication. It is amazing that artificial selection could have targeted a gene with such a wide range of pleiotropic effects rather than multiple genes each affecting various aspects of plant architecture, such as those individually controlling tiller number and angle or panicle branches (e.g., Li et al. 2003; Ashikari et al. 2005; Li et al. 2007; but see Jiao et al. 2010; Miura et al. 2010).

Another remarkable example of pleiotropic effect is the *Q* gene in wheat. While it is largely responsible for the development of free-threshing, it also contributes to non-brittle rachis, shorter culms, and shorter and denser spikes. The pleiotropic effect involves not only shattering and threshing necessary for effective harvest, but also plant architecture and inflorescence structure important for yield. Thus *Q* has been considered to be a super domestication gene (Faris et al. 2006).

During barley domestication, the appearance of six-rowed ears was a key innovation that substantially increased yield. On the ears of wild progenitors, each spike serving as a seed dispersal unit consists of three spikelets, of which the two lateral ones are reduced with only awns left to assist the dispersal of fully developed central spikelet. This trait remains the same in the domesticated two-rowed barley, while the awns are lost in the lateral spikeltes. In more advanced cultivars, the two lateral spikelets become fully developed so that the number of rows of grains is tripled.

Another barley gene, *Vrs1*, that controls the development of the lateral spikelets was cloned (Komatsuda et al. 2007). It encodes an HD-ZIP containing transcription factor expressed specifically in the lateral-spikelet primordia and suppresses the development of the lateral rows. The loss-of-function mutation of *Vrs1* allows further development of the lateral rows and gives rise to six-rowed barley. Three mutations in the coding region independently disrupt the function of the gene (Komatsuda et al. 2007).

12.3.3 Grain/Seed Cover, Size, and Coloration

The loss of grain/seed cover and coloration is another component of the domestication syndrome. The most remarkable is the loss of fruitcases during maize domestication. In teosinte, kernels are enclosed by stony fruitcases derived from modified

cupules and glumes. Teosinte ears disarticulate at maturity and the fruitcases become the units of seed dispersal. In maize, the fruitcases do not form so that naked kernels are readily edible. Cupules and glumes become a part of maize cobs on which kernels remain undetached at maturity. *Teosinte glume architecture1* (*tga1*), a major QTL controlling fruitcase formation, has been cloned (Wang et al. 2005). It is a member of the squamosa-promoter binding protein family of transcription regulators. In maize and teosinte, the gene is expressed in the inflorescence meristem of a developing ear, the spikelet primordia, and the adaxial junction of the spikelet and the inflorescence axis, the region where cupules and glumes develop. The functional difference between the maize and teosinte alleles of *tga1* appears to be due to a single amino acid substitution.

Changes in hull colors were widespread during cereal domestication. A common trend was the change from dark colors to the color that mimics withering straws, so-called straw-white. Two major QTLs were found to be responsible primarily for the change of hull color in rice (Gu et al. 2005). The one with larger effect, *bh4*, was recently cloned using a near-isogenic line of an *indica* cultivar containing a small region of chromosome 4 from *O. rufipogon* (Zhu et al. 2011). The wild-type allele in *O. rufipogon* encodes an amino acid transporter that is expressed specifically in the developing hulls. In different rice cultivars examined, it was found that two deletions and a nucleotide substitution in the coding region of the gene were independently responsible for the truncation of the BH4 protein and consequently for the loss of black hull color.

The color of seed coats or pericarps, although invisible during harvest, was also a target of artificial selection. The wild progenitors of rice have dark (brown to red) pericarps, whereas the pericarps of cultivated rice are predominantly white. Genetic analyses have shown that the color of rice pericarps is controlled primarily by two loci, *Rc* and *Rd*. Mutations at both loci have been identified. *Rc* encodes a bHLH protein that presumably regulates anthocyanin biosynthesis in the seed coat (Sweeney et al. 2006). Two mutations in exon 6 of the gene could independently be responsible for the loss of pigmentation in the pericarps. A 14-bp deletion was found in nearly 98 % of rice cultivars with white pericarps and a nucleotide substitution resulting in a premature stop codon was found in the remaining white rice (Sweeney et al. 2007).

Rd encodes dihydroflavonol-4-reductase, an enzyme involved in anthocyanin biosynthesis. The presence of premature stop codons in the first and second exons disrupts the function of the enzyme. When the lose-of-function alleles are denoted as *rc* and *rd*, the *Rc/Rd* and *Rc/rd* genotypes produce red and brown pericarps, respectively. The *rc/Rd* and *rc/rd* genotypes produce white pericarps (Furukawa et al. 2007).

Another color-related trait that experienced artificial selection is the darkening of hulls and pericarps of *indica* rice cultivars after prolonged storage. This, however, does not occur in *japonica* cultivars. The difference between the two types of cultivars is controlled by a single gene, *Phr1*, which encodes a polyphenol oxidase (Yu et al. 2008). The survey of *indica* cultivars identified three independent mutations, including two deletions and one insertion in the coding region that disrupted the function of the gene.

Larger and heavier grains are obviously favored by farmers. Numerous QTLs controlling grain size and weight have been identified. One of them, *qSW5*, with

the largest effect on the difference in grain width between a pair of *indica* and *japonica* cultivars appeared to have been targeted by artificial selection (Shomura et al. 2008). Molecular characterization of *qSW5* has indicated that in *indica* cultivars with narrower grains, the gene determines the number of cells in the outer glumes of rice flowers. In *japonica* cultivars with wider grains, a large deletion knocks out the gene function, which allows for the development of wider grains with additional rows of cells in the outer glumes.

The molecular cloning of a major QTL, *GS3*, controlling grain length provided additional insights into the changes of grain size and shape during rice domestication (Fan et al. 2006; Takano-Kai et al. 2009; Mao et al. 2010). The wild-type allele of *GS3* in the wild species and in the cultivars with grains of medium length is a negative regulator of grain size. Cultivars with long grains carry an allele with a loss-of-function mutation. It was targeted by artificial selection that favored long grains during rice cultivation.

It is interesting to note that none of the derived alleles responsible for an increase in rice grain width or length has been driven to fixation by artificial selection. This reflects the existence of highly diverse grain shapes and sizes among rice cultivars. It seems that while selection for larger grains was generally favored during rice cultivation, it was not one of the key driving forces of rice domestication. There are two reasons for this. One is that grain size and grain number on a panicle are often negatively correlated because the source for grain filling is limited (Wang et al. 2011). A trade-off between them may be beneficial for higher yield. The other reason is that in addition to grain size, variable grain shapes were probably selected by rice consumers and growers (Takano-Kai et al. 2009).

12.4 Gene Evolution and Domestication Processes

Because crops were derived from wild species under artificial selection, the cloning and evolutionary analyses of domestication genes should shed light on the history and processes of domestication. This type of information has been shown to be valuable for addressing questions such as how complex the genetic basis of a domestication trait could be? Whether a crop or a domestication trait originated once or multiple times? How long a domestication process could have lasted? In this section, we review the phylogenetic and population genetic analyses of the domestication-related genes identified in rice, maize, barley, and wheat, from which we attempt to gain a better understanding of domestication processes.

12.4.1 Rice

Rice, with a small genome and high-quality genome sequences (IRGSP 2005), has become a model system for studying plant function. It is thus not surprising that a much larger number of domestication genes have been cloned in rice than in other

cereals. Prior to the cloning of domestication genes, analyses of multiple neutral molecular markers suggested that rice was domesticated at least twice, with *indica* and *japonica* cultivars originated independently from wild species in different geographic locations (Cheng et al. 2003; Ma and Bennetzen 2004; Vitte et al. 2004; Zhu and Ge 2005; Londo et al. 2006; but see Molina et al. 2011a, b). The initial examination of the distribution of *sh4* alleles was apparently contradictory to the view of independent domestication. The finding that all cultivars shared the mutation responsible for the non-shattering phenotype indicated that there was a single origin of the domestication allele and possibly a single origin of cultivated rice (Lin et al. 2007).

Two models were proposed for the reconciliation (Sang and Ge 2007a). The snowballing model considers a single origin of a rice cultivar containing a set of essential domestication alleles including *sh4*. This original cultivar then spread and hybridized with wild populations, which gave rise to new cultivars with divergent genomic background. This process also led to an increase in genetic diversity of rice cultivars including their diversification into two groups of cultivars, *indica* and *japonica*. In the combination model, rice cultivars are believed to have been domesticated independently and cultivars of different origins possessed distinct sets of domestication genes. After these cultivars spread and hybridized, the best set of domestication genes were selected and combined into the modern cultivars while the genomic diversity was largely maintained.

A recent study based on the gene markers from three rice chromosomes claimed a strong support for the single origin of domesticated rice (Molina et al. 2011a). This result was challenged because the study seemed to have underestimated the likelihood of independent domestications (Ge and Sang 2011). While the models of rice domestication have yet to be tested, it is almost certain that in either case the gene flow between cultivars and wild progenitors was an important part of the process of rice domestication. The spreading of valuable domestication alleles through introgression was also documented at another domestication related locus, *Rc*, which primarily controls the pericarp color. It was shown that the common lost-of-function allele of the gene that accounted for 98 % of white-pericarp in rice originated in *japonica* rice and subsequently spread into *indica* cultivars through introgression (Sweeney et al. 2007).

Insight into the domestication of rice was also gained through population genetic analysis of *sh4*, which estimated the rate of fixation of the non-shattering allele in cultivated rice. A severe reduction of DNA sequence polymorphism in cultivated rice was observed at the *sh4* locus, suggesting that the allele experienced a selective sweep under strong artificial selection. Because the allele allowed for well balanced shattering and threshing, the selection could be strong enough to drive its fixation in a short period of ~100 years (Zhang et al. 2009).

However, it was recently found that in the archeological sites in China, the frequency of non-shattering phenotype increased relatively slowly, suggesting that rice domestication, as evaluated on the basis of the development of a non-shattering trait, might have lasted for two to three millennia (Fuller et al. 2009).

Nevertheless, this apparent discrepancy in the rate of rice domestication between genetic and archeological data can be reconciled under certain circumstances. One possibility is that *sh4* did not arise or did not have a chance to spread to the archeological sites where cultivars with inferior non-shattering alleles were present (Zhang et al. 2009). If this turns out to be the case, one can conclude that population genetic analyses of domestication genes coupled with archeological evidences would give a more complete picture of rice domestication. The fixation of a critical domestication allele could be rapid locally, but prolonged for the crop as a whole.

The similar situation was later found in another domestication gene, *prog1*, which was most likely allelic to the QTL controlling a suite of morphological changes leading to better architecture and higher yield of cultivated rice. Although there has not been a thorough phylogenetic and population genetic analysis of *prog1* in comparison to *sh4*, based on genetic analyses reported from two independent studies it seems clear that the domestication allele *prog1* with modified functions originated once and had been fixed in all rice cultivars examined (Jin et al. 2008; Tan et al. 2008).

The shared pattern between *sh4* and *prog1*, each underlying a key component of domestication syndrome in rice, reinforces the notion that introgression played an essential role in the rapid fixation of domestication alleles giving rise to superior phenotypes. It was strong artificial selection that led to the spreading and fixation of the most desirable set of domestication alleles. Meanwhile, natural selection on hybrids with distinct genomic backgrounds would help maintain local adaptation and genetic diversity of cultivars (Sang and Ge 2007b).

The two genes controlling hull and pericarp colors, *Bh4* and *Rc*, share another pattern. Cultivars have alleles with loss-of-function mutations, with one allele being predominant and others occurring at low frequencies and having independent origins (Sweeney et al. 2007; Zhu et al. 2011). These low-frequency alleles have been maintained in cultivars probably because they are functionally indistinguishable from the common alleles and have not been wiped out by selective sweep associated with the common alleles. Given such a low nucleotide polymorphism of the common white-hull allele, it seems to have originated recently, possibly even more recent than *sh4*. If the change of hull color from black to straw-white was indeed to mimic straw color and avoid bird predation, it makes sense that this change occurred after the non-shattering phenotype was widely established (Zhu et al. 2011).

Like *bh4* and *rc*, the derived alleles of *GS3* and *SW5* also had loss-of-function mutations. However, the frequencies of the loss-of-function alleles are much lower, consistent with the hypothesis that longer and/or wider grains were not universally favored in cultivated rice. The change in grain shape and increase in grain size must have met various limitations such as source availability and consumer preference. Although both alleles experienced artificial selection and increased in frequency during rice domestication, they were obviously different from those domestication alleles that carried clearly selective advantages in one direction.

12.4.2 Maize

Phylogenetic analyses of genome-wide neutral markers were in agreement with the archeological evidence, pointing to a single domestication of maize near the central Balsas river valley of southern Mexico approximately 7,000–9,000 years ago (Matsuoka et al. 2002; Doebley 2004). The identification of two major domestication genes, *tb1* and *tg1*, each regulating a key phenotypic transition lend additional support to this conclusion.

The story of *tg1* is probably the most straightforward of all domestication genes. A single nonsynonymous substitution was primarily responsible for the loss of stony fruitcases that protected teosinte kernels. The allele was derived ~10,000 years ago and was fixed in maize under strong artificial selection (Wang et al. 2005). The estimated early origin of this allele also suggests that an absence of fruitcase in kernels for easier human consumption was an important early step of maize domestication. The single domestication event allowed *tg1* to be fixed more easily than *sh4* in rice.

The gene *tb1* was the first cloned QTL for domestication of crops. The search for the causal mutation, however, took much longer than any other domestication genes later identified. After almost a decade of persistent effort, the mutation(s) were narrowed down to a region ~58–69 kb 5′ to the coding region of *tb1* (Clark et al. 2004, 2006). It was also confirmed that the mutation(s) were in the regulatory region, which changed *tb1* expression and plant architecture. Despite the early detection of strong artificial selection in the 5′ upstream region of *tb1*, the strikingly distant location of the causal mutation(s) from the gene itself complicated the fine mapping process.

Although the domestication history of maize seems to be the best characterized of all cereals, there are still many questions left to be addressed with continuing effort to identify and analyze domestication genes. It is still unclear whether a single mutation or several mutations of independent origins gave rise to *tb1* in maize. With genome sequences available, cloning additional important domestication genes of maize will be of great interest for comparison with rice domestication. The extent to which introgression between cultivars and teosinte might have influenced maize domestication has just begun to be revealed (van Heerwaarden et al. 2011).

12.4.3 Barley and Wheat

Barley and wheat, belonging to the same subfamily, Pooideae, were domesticated from the Fertile Crescent ~10,000 years ago. Map-based QTL cloning is relatively difficult in barley and wheat due to their large genome sizes. Despite this, essential domestication QTLs such as *btr* and *br* controlling brittle rachis were mapped on chromosomes, although their characterization at the molecular level is still awaited. The fact that these QTLs are mapped to the orthologous chromosomal

regions between barley and wheat suggests that the orthologous genes might have been targeted by artificial selection for non-brittle rachis in these crops (Sakuma et al. 2011). This will be an interesting hypothesis to test once the QTLs are cloned (Paterson et al. 1995).

For the free-threshing trait, the major QTL, *Nud*, was cloned from barley, and a secondary QTL, *Q*, was cloned from wheat. Each of these two domestication genes had a single origin and spread into cultivars of independent origins (Simons et al. 2006; Taketa et al. 2008). For barley, a growing body of phylogenetic evidences suggested that it contained cultivars with distinct genomic backgrounds derived from wild progenitors in two places, the Fertile Crescent and another area 1,500–3,000 km farther east (Morrell and Clegg 2007). This situation resembles the case of rice domestication where two types of cultivars (indica and japonica) with distinct genomic backgrounds share the domestication genes of single origins. It appears that the models of rice domestication could at least partly apply to other cereals.

Free-threshing wheat includes both tetraploid and hexaploid cultivars. With a gain-of-function mutation, one *Q* locus could confer the phenotype in the polyploid wheat. The free-threshing condition may have developed in the AABB-genome wheat with the genotype $tgtg^{2B}QQ^{5A}$, which then served as the tetraploid parent of hexaploid bread wheat. In this scenario, an additional recessive mutation at *Tg* of the DD genome would be sufficient. Alternatively, *Q* was selected initially in the hexaploid wheat and then spread into tetraploid wheat through introgression (Faris et al. 2006; Simons et al. 2006). Either scenario supports the importance of gene flow during cereal domestication.

12.4.4 Generalization

Above we reviewed and discussed recent advances in the identification and analyses of genes controlling the domestication syndrome of cereals. The rapid progress made along this line over the past few years substantially enhanced our understanding of crop domestication. This provides us with the opportunity to evaluate certain general conclusions drawn on domestication mechanisms and processes.

Several decades ago, Beadle proposed the “one-gene, one-trait” hypothesis as an explanation for the finding that one chromosome region often accounted for a major domestication transition from teosinte to maize (Beadle 1939; Doebley 2001). The cloning of domestication genes in cereals supports this hypothesis to a large extent (Sang 2009). A single gene or even a single mutation in many cases controlled a critical domestication transition. These include *tb1* and *tg1* in maize, *sh4* and *prog1* in rice, and *nud* in barley, each primarily responsible for the development of a key component of the domestication syndrome. It was remarkable that the one-gene, one-trait hypothesis derived from maize holds for rice and barley despite the fact that each of these crops might have had multiple origins. Models developed for rice provided a mechanistic explanation for the one-gene, one-trait

hypothesis (Sang and Ge 2007a, b). They suggested that a highly favorable domestication allele could be quickly fixed even for crops of multiple origins by strong artificial selection combined with introgression.

Whether a domestication allele is fixed in a crop or not probably reflects the strength of artificial selection. It is noticeable that the majority of alleles conferring reduction in coloration and increase in grain length or width are not fixed in rice. These traits are not essential for cultivation; actually in some cases, wild-type alleles were selected for the maintenance of phenotypic diversity as exemplified by with red-pericarp and variable grain shapes among rice cultivars. Furthermore, these alleles, including *rc*, *rd*, *phr1*, *qSW5*, and *GS3*, were derived from loss-of-function mutations that often had multiple origins.

It was previously suggested, based on the review of domestication-related genes in a broader range of crops, that key domestication genes tended to be regulatory genes with modified functions in crops, whereas genes controlling varietal variation were broader in functional categories and more frequently underwent loss-of-function mutations during domestication (Doebley et al. 2006). The growing number of domestication-related genes cloned in cereals supports the hypothesis and further suggests that essential domestication alleles tend to have single origins while alleles controlling varietal variation often have multiple origins (Table 12.1).

Taken together, a critical domestication transition necessary for cultivation tends to be controlled by a single mutation of a single gene. The domestication genes are often transcription factors with modified functions resulting from mutations in either regulatory or coding regions (Table 12.1). The functional modification could cause a cascade of downstream effects that substantially altered a trait but had relatively little deleterious effects on plants (Doebley and Lukens 1998). The selection on these alleles was so strong that could have driven them to fixation in cultivars of independent origins. Genes for further improvement of domestication traits were not absolutely necessary for cultivation, and therefore were more diverse in functional categories and more often had independent loss-of-function mutations. The domestication alleles of these genes are often not fixed in all cultivars due to either relatively weak selection or diversifying selection for varietal variation.

12.5 Recent Advances

The cloning and characterization of genes underlying the domestication syndrome of cereals yielded intriguing insights into the molecular and population genetics of crop domestication. It is expected that the pace of cloning domestication QTLs will increase with rapid advances in grass genomics. With additional cereal genomes, especially those of sorghum and foxtail millet, being recently sequenced, the effort to identify domestication-related genes from these two crops will begin to bear fruits and consequently allow a comparison of the genetic basis of domestication in a wider range of cereal crops.

Another major advance in genomics over the past few years was the development of the next-generation sequencing technologies. This will bring substantial changes to the molecular genetic studies of domestication traits. QTL cloning may be considerably accelerated with the use of genotyping methods relying on next-generation sequencing. It has been shown that resequencing the genomes of recombinant populations at low genome coverage could markedly speed up genotyping and improve the resolution of QTL mapping (Huang et al. 2009). QTL maps with higher precision and resolution will facilitate the process of gene cloning (Wang et al. 2011).

Rapid genome resequencing also opened great opportunities for identifying domestication-related genes through a direct analysis of genome-wide nucleotide polymorphisms between crops and their wild progenitors. Severe reduction of nucleotide polymorphism in cultivars is an indication of selective sweep and these regions potentially contain genes that experienced artificial selection. This "bottom-up" approach has been applied to maize using PCR-based markers (Vigouroux et al. 2002; Wright et al. 2005), although the full potential of SNPs identified from whole-genome resequencing has not been tested in crops. The resequencing data seem to be most accessible for rice whose small and homozygous genome offers an excellent system for exploring the potential of this approach (He et al. 2011).

Despite the seemingly promising potential, we have yet to see a body of publications emerging from this idea. Apparently, several technical issues need to be addressed. First, it requires the sequences of a reference genome for a crop system. In cereals, this requirement is met better than any other groups of crops. High-quality genome sequences are available for rice, maize, and sorghum (Paterson et al. 2009; Schnable et al. 2010), and will soon be available for foxtail millet (Doust et al. 2009). Sequencing the large genomes of barley and wheat is still challenging, but should become increasingly realistic with the emergence of the third-generation sequencing technologies.

Second, high-density SNPs generated by whole-genome resequencing provide great opportunities for narrowing down candidate genes anywhere in the genome, but it presents considerable challenges for the detection of regions that indeed experienced selective sweep. With the whole genome under examination, the possibility of detecting false positives increases dramatically. Various stochastic factors combined with population demography are very likely to cause difficulties in setting appropriate statistical threshold for maximizing the detection of domestication-related genes and meanwhile minimizing false positives. New population genomic models and analytical methods are needed in order to take full advantage of resequencing data to single out regions under reasonably strong artificial selection.

Finally, a few sampling strategies need to be considered. Appropriate genome coverage of resequencing should be determined for specific cases. There is a possibility that a relatively low coverage is sufficient for generating a high-density haplotype map through data imputation as long as the sample size is large enough (Huang et al. 2010). The difficulty remains however for the imputation in crops and their wild relatives whose genomes are heterozygous. The same is true for

taxonomic sampling strategy. A suitable number of cultivars and wild relatives that effectively cover the entire diversity of crops and wild relatives needs to be determined. There should be ways to identify individuals derived from introgression between cultivars and their wild relatives. Including the hybrids in the analysis may violate the assumptions of population genetic models. Furthermore, unlike the top-down approach that starts with traits known to be important for domestication, candidate genes identified from genome screening need to be tested for their phenotypic effect and for their role in artificial selection.

12.6 Summary and Outlook

The independent domestication of cereal crops was triggered at least partly by the climate change following the last glacial maxima. Today the world is facing an anthropogenic climate change that will presumably have serious consequences. This is primarily due to rapid consumption of fossil fuels since the Industrial Revolution. The quick depletion of fossil energy and the consequential climate changes has forced the world to develop renewable sources of energy. Of various possible sources, bioenergy is the one that holds the greatest potential for replacing fossil oil. The large-scale production of bioenergy will have to rely heavily on energy crops.

Crops that are currently used for producing bioethanol and biodiesel are those previously domesticated for food, sugar or vegetable oil, such as maize, sugarcane, soybean, and rapeseed. They are known as first-generation energy crops. However, it has been increasingly realized that these annual crops cannot be sustainable solutions to energy shortage or climate change because of their low net energy output, low potential for greenhouse gas mitigation, and inability to utilize marginal land (Fargione et al. 2008; Robertson et al. 2008; Searchinger et al. 2008). Under these circumstances, the concept of second-generation energy crops emerged. These are dedicated energy crops that can be grown on marginal land with relatively little energy input involved in plantation, irrigation, fertilization, harvest, and transportation. Meanwhile, they should yield high biomass and have a strong ability of carbon sequestration. With these features, second-generation energy crops are capable of providing a major sustainable source of renewable energy with little or even negative greenhouse gas emission (Heaton et al. 2008; Karp and Shield 2008; Oliver et al. 2009).

Grasses are likely to play a leading role as potential second-generation energy crops. Therefore, what we learned from cereal crop domestication can be used for guiding us in bring about the next round of crop domestication for energy and environment security. Two perennial grasses, switchgrass and *Miscanthus*, native to North America and Asia respectively, have already emerged as the top candidates for second-generation energy crops in the northern temperate regions of the world (Somerville et al. 2010; Sang and Zhu 2011). They share the features of being C4 perennials with high water and nutrient use efficiencies and are adapted to a wide range of climates. Similar to cereal crops, the domestication of these energy crops also tends to target at a small portion of grass species native to different continents.

Unlike cereal domestication, where artificial selection focused on planting, harvest efficiencies, grain yield, and harvest indices, the domestication of lignocellulosic energy crops will focus on sustainable production under unfavorable soil and climate conditions (Sang 2011). With regard to the domestication traits, they may share characteristics such as strong biotic and abiotic resistance, maximized length of vegetative growth as permitted by local climates, the highest possible water, nutrient, and radiation use efficiencies, at least partial sterility that minimizes seed production, and modified lignocellulosic properties for effective biorefining.

In addition to the domestication syndrome, the duration of domestication of dedicated energy crops will have to be much shorter than that of cereals. Instead of centuries to millennia taken for domesticating a food crop, the domestication of energy crops needs to be completed within decades in order to address problems of energy shortage and anthropogenic climate change facing the world. Fortunately, the knowledge gained from the study of cereal domestication provides us with encouraging clues to rapid domestication. Because a key domestication transition could result from selection for a single mutation, it is possible that a desirable trait develops quickly in a new crop, if a large population with ample natural and induced mutations is subjected to strong artificial selection. We can then shape up the domestication syndrome by combining a suite of desirable traits through carefully designed experimental crosses. Genetic and genomic studies that identify valuable loci and genes will considerably facilitate and speed up the process of domestication through molecular breeding and biotechnology.

References

Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles E, Qian Q, Kitano H, Matuoka M (2005) Cytokinin oxicase regulates rice grain production. Science 309:741–745

Azhanguvel P, Komatsusa T (2007) A phylogenetic analysis based on nucleotide sequence of a marker linked to the brittle rachis locus indicates a diphyletic origin of barley. Ann Bot 100:1009–1015

Beadle GW (1939) Teosinte and the origin of maize. J Hered 30:245–247

Cheng C, Motohashi R, Tsuchimoto S, Fukuta Y, Ohtsubo H, Ohstubo E (2003) Polyphyletic origin of cultivated rice: based on the interspersion pattern of SINEs. Mol Biol Evol 20:67–75

Chuck G, Meeley R, Irish E, Sakai H, Hake S (2007) The maize *tasselseed4* microRNA controls sex determination and meristem cell fate by targeting *Tasselseed6/indeterminate spkelet1*. Nat Genet 39:1517–1521

Clark RM, Linton E, Messing J, Doebley JF (2004) Pattern of diversity in the genomic region near the maize domestication gene *tb1*. Proc Natl Aced Sci USA 101:700–707

Clark RM, Wagler TN, Quijada P, Doebley JF (2006) A distant upstream enhance at the maize domestication gene *tb*1 has pleiotropic effects on plant and inflorescent architecture. Nat Genet 38:594–597

Cunniff J, Osborne CP, Ripley BS, Charles M, Jones G (2008) Response of wild C4 crop progenitors to subambient CO_2 highlights a possible role in the origin of agriculture. Glob Change Biol 14:576–587

Darwin CR (1859) On the origin of species by means of natural selection. Jone Murray, London

Diamond J (2002) Evolution, consequences and future of plant and animal domestication. Nature 418:700–707

Doebley JF (2001) George Beadle's other hypothesis: one-gene, one-trait. Genetics 158:487–493
Doebley JF (2004) The genetics of maize evolution. Ann Rev Genet 38:37–59
Doebley JF, Lukens L (1998) Transcriptional regulators and the evolution of plant form. Plant Cell 10:1075–1082
Doebley JF, Stec A, Hubbard L (1997) The evolution of apical dominance in maize. Nature 386:485–488
Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. Cell 127:1309–1321
Doust AN, Kellogg EA, Devos KM, Bennetzen JL (2009) Foxtail millet: a sequence-driven grass model system. Plant Physiol 149:137–141
Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X, Zhang Q (2006) *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. Theor Appl Genet 112:1164–1171
Fargione J, Hill J, Tilman D, Polasky S, Hawthorne P (2008) Land clearing and the biofuel carbon debt. Science 319:1235–1238
Faris JD, Simons KJ, Zhang Z, Gill BS (2006) The wheat super domestication gene *Q*. Wheat Information Service—Frontiers of Wheat Bioscience 100:129–148. (http://www.shigen.nig.ac.jp/wheat/wis/No100/100.html)
Fuller DQ, Qin L, Zheng Y, Zhao Z, Chen X, Hosoya LA, Sun GP (2009) The domestication process and domestication rate in rice: spikelet bases from the lower Yangtze. Science 323:1607–1610
Furukawa T, Maekawa M, Oki T, Suda I, Lida S, Shimada H, Takamure I, Kadowaki K (2007) The Rc and Rd genes are involved in proanthocyanidin synthesis in rice pericarp. Plant Journal 49:91–102
Ge S, Sang T (2011) Inappropriate model rejects independent domestications of *indica* and *japonica* rice. Proc Natl Acad Sci USA 108:E75
Gu XY, Kianian SF, Hareland GA, Hoffer BL, Foley ME (2005) Genetic analysis of adaptive syndromes interrelated with seed dormancy in weedy rice (*Oryza sativa*). Theor Appl Genet 110:1108–1118
Hancock JF (2004) Plant evolution and the origin of crop species, 2nd edn. CABI Publishing, Cambridge
Harlan JR (1992) Crops and man, 2nd edn. American Society of Agronomy and Crop Science Society of America, Madison
He Z, Zhai W, Wen H, Tan T, Wang Y, Lu X, Greenburg AJ, Hudson RR, Wu C-I, Shi S (2011) Two evolutionary histories in the genome of rice: the roles of domestication genes. PLoS Genet 7:e1002100
Heaton EA, Flavell RB, Mascia PN, Thomas SR, Dohleman FG, Long SP (2008) Herbaceous energy crop development: recent progress and future prospects. Curr Opin Biotech 19:202–209
Huang XH, Feng Q, Qian Q, Zhao Q, Wang L, Wang AH, Guan JP, Fan DL, Weng QJ, Huang T, Dong GJ, Sang T, Han B (2009) High-throughput genotyping by whole-genome resequencing. Genome Res 19:1068–1076
Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z, Li M, Fan D, Guo Y, Wang A, Wang L, Deng L, Li W, Lu Y, Weng Q, Liu K, Huang T, Zhou T, Jing Y, Li W, Lin Z, Buckler ES, Qian Q, Zhang Q, Li J, Han B (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. Nat Genet 42:961–967
Hubbard L, McSteen P, Doebley J, Hake S (2002) Expression patterns and mutant phenotype of *teosinte branched1* correlate with growth suppression in maize and teosinte. Genetics 162:1927–1935
International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. Nature 436:793–800
Jantasuriyarat C, Vales MI, Watson CJW, Riera-Lizarazu O (2004) Identification and mapping of genetic loci affecting the free-threshing habit and spike compactness. Theor Appl Genet 108:261–273
Jiao Y, Wang Y, Xue D, Wang J, Yan M, Liu G, Dong G, Zeng D, Lu Z, Zhu X, Qian Q, Li J (2010) Regulation of *OsSPL*14 by OsmiR156 defines ideal plant architecture in rice. Nat Genet 42:541–544

Jin J, Huang W, Gao J-P, Yang J, Shi M, Zhu M-Z, Luo D, Lin H-X (2008) Genetic control of rice plant architecture under domestication. Nat Genet 40:1365–1369

Karp A, Shield I (2008) Bioenergy from plants and the sustainable yield challenge. New Phytol 179:15–32

Komatsuda T, Maxim P, Senthil N, Mano Y (2004) High-density AFLP map of nonbrittle rachis 1 (*btr*1) and (*btr*2) genes in barley (*Hordeum vulgare* L.). Theor Appl Genet 109:986–995

Komatsuda T, Pourkheirandish M, He C, Azhaguvel P, Kanamori H, Perovic D, Stein N, Graner A, Wicher T, Tagiri A et al (2007) Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. Proc Natl Acad Sci USA 104:1424–1429

Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M (2006) An SNP caused loss of seed shattering during rice domestication. Science 312:1392–1396

Li W, Gill BS (2006) Multiple genetic pathways for seed shattering in the grasses. Funct Integr Genomics 6:300–309

Li X, Qian Q, Fu Z, Wang Y, Xiong G, Zeng D, Wang X, Liu X, Teng S, Hiroshi F, Yuan M, Luo D, Han B, Li J (2003) Control of tillering in rice. Nature 422:618–621

Li C, Zhou A, Sang T (2006a) Rice domestication by reducing shattering. Science 311:1936–1939

Li C, Zhou A, Sang T (2006b) Genetic analysis of rice domestication syndrome with the wild annual species, *Oryza nivara*. New Phytol 170:185–194

Li P, Wang Y, Qian Q, Fu Z, Wang M, Zeng D, Li B, Wang X, Li J (2007) *LAZY1*controls rice shoot gravitropism through regulating polar auxin transport. Cell Res 17:402–410

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

Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA (2006) Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. Proc Natl Acad Sci USA 103:9578–9583

Ma J, Bennetzen JL (2004) Rapid recent growth and divergence of rice nuclear genomes. Proc Natl Acad Sci USA 101:12404–12410

Mao H, Sun S, Yao J, Wang C, Yu S, Xu C, Li X, Zhang Q (2010) Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. Proc Natl Acad Sci USA 107:19579–19584

Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez J, Buckler E, Doebley J (2002) A single domestication for maize shown by multilocus microsatellite genotyping. Proc Natl Acad Sci USA 99:6080–6084

Miura K, Ikeda M, Matsubara A, Song X, Ito M, Asano K, Matsuoka M, Kitano H, Ashikari M (2010) OsSPL14 promotes paniclebranching and higher grain productivity in rice. Nat Genet 42:545–549

Molina J, Sikora M, Garud N, Flowers JM, Rubinstein S, Rynolds A, Huang P, Jackson SA, Schaal BA, Bustanante CD, Boybo AR, Purugganan MD (2011a) Molecular evidence for a single evolutionary origin of domesticated rice. Proc Natl Acad Sci USA 108:8351–8356

Molina J, Sikora M, Garud N, Flowers JM, Rubinstein S, Rynolds A, Huang P, Jackson SA, Schaal BA, Bustanante CD, Boybo AR, Purugganan MD (2011b) Reply to Ge and Sang: a single origin of domesticated rice. Proc Natl Acad Sci USA, doi/10.1073/pnas.1112466108

Morrell PL, Clegg MT (2007) Genetic evidence for a second domestication of barley (*Hordeum vulgare*) east of the Fertile Crescent. Proc Natl Acad Sci USA 104:3289–3294

Nalam VJ, Vales MI, Watson CJW, Kianian SF, Riera-Lizarazu O (2006) Map-based analysis of genes affecting the brittle rachis character in tetraploid wheat (*Triticum turgidum* L.). Theor Appl Genet 112:373–381

Nalam VJ, Vales MI, Watson CJW, Johnson EB, Riera-Lizarazu O (2007) Map-based analysis of genetic loci on chromosome 2D that affect glume tenacity and threshability, components of the free-threshing habit in common wheat (*Triticum aestivum* L). Theor Appl Genet 116:135–145

Oliver RJ, Finch JW, Taylor G (2009) Second generation bioenergy crops and climate change: a review of the effects of elevated atmospheric CO_2 and drought on water use and the implications for yield. GCB Bioenergy 1:97–114

Onishi K, Horiuchi Y, Ishigoh-Oka N, Takagi K, Ichikawa N, Maruoka M, Sano Y (2007) A QTL cluster for plant architecture and its ecological significance in Asian wild rice. Breeding Sci 57:7–16

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

Paterson AH et al (2009) The Sorghum bicolor genome and the diversification of grasses. Nature 457:551–556

Pourkheirandish M, Komatsuda T (2007) The importance of barley genetics and domestication in a global perspective. Ann Bot 100:999–1008

Pozzi C, Rossini L, Vecchietti A, Salamini F (2004) Gene and genome changes during domestication of cereals. In Gupta PK, Varshney RK (eds) Cereal genomics, pp 165–198

Richerson PJ, Boyd R, Bettinger RL (2001) Was agriculture impossible during the Pleiostocene but mandatory during the Holocene? A climate change hypothesis. Amer Antiq 66:387–411

Robertson GP, Dale VH, Doering OC, Hamburg SP, Melillo JM, Wander MM, Parton WJ, Adler PR, Barney JN, Cruse RM, Duke CS, Fearnside PM, Follett RF, Gibbs HK, Goldember J, Dladenoff DJ, Ojima D, Palmer M, Sharpley A, Wallace L, Weathers KC, Wiens JA, Wilhelm WW (2008) Sustainable biofuels redux. Science 322:49–50

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

Sage R (1995) Was low atmospheric CO_2 during the Pleistocene a limiting factor for the origin of agriculture? Glob Change Biol 1:93–106

Sakuma S, Salomon B, Komatsuda T (2011) The domestication syndrome genes responsible for the major changes in plant form in the Triticeae crops. Plant Cell Physiol 52:738–749

Salamini F, Ozkan H, Brandolini A, Schafer-Pregl R, Marin W (2002) Genetics and geography of wild cereal domestication in the Near East. Nat Rev Genet 3:420–441

Sang T (2009) Genes and mutations underlying domestication transitions in grasses. Plant Physiol 149:63–70

Sang T (2011) Toward the domestication of lignocellulosic energy crops: learning from food crop domestication. J Integr Plant Biol 53:96–104

Sang T, Ge S (2007a) The puzzle of rice domestication. J Integr Plant Biol 49:760–768

Sang T, Ge S (2007b) Genetics and phylogenetics of rice domestication. Cur Opin Genet Dev 17:533–538

Sang T, Zhu W-X (2011) China's bioenergy potential. GCB Bioenergy 3:79–179

Schnable et al (2010) The B73 maize genome: complexity, diversity, and dynamics. Science 326:1112–1115

Searchinger T, Heimlich R, Houghton RA, Dong F, Elobeid A, Fabiosa J, Tokgoz S, Hayes D, Yu TH (2008) Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land use change. Science 319:1238–1244

Shomura A, Izawa T, Ebana K, Ebitani T, Kanegae H, Konishi S, Yano M (2008) Deletion in a gene associated with grain size increased yields during rice domestication. Nat Genet 40:1023–1028

Simonetti MC, Bellomo MP, Laghetti G, Perrino P, Simeone R, Blanco A (1999) Quantitative trait loci influencing free-threshing habit in tetraploid wheats. Genet Res Crop Evol 46:267–271

Simons KJ, Fellers JP, Trick HN, Zhang Z, Tai YS, Gill BS, Faris JD (2006) Molecular characterization of the major wheat domestication gene *Q*. Genetics 172:547–555

Somerville C, Yongs H, Taylor C, Davis SC, Long SP (2010) Feedstocks for lignocellulosic biofuels. Science 329:790–792

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

Sweeney MT, Thomson MJ, Cho YG, Park YJ, Williamson SH, Bustamante CD, McCouch SR (2007) Global dissemination of a single mutation conferring white pericarp in rice. PLoS Genet 3:e133

Takano-Kai N, Jiang H, Kubo T, Sweeney M, Matsumoto T, Kanamori H, Padhukasahasram B, Bustamante C, Yoshimura A, Doi K, McCouch S (2009) Evolutionary history of GS3, a gene conferring grain length in rice. Genetics 182:1323–1334

Taketa S, Amano S, Tsujino Y, Sato T, Saisho D, Kakeda K, Nomura M, Suzuki T, Matsumoto T, Sato K et al (2008) Barley grain with adhering hulls is controlled by an ERF family transcription factor gene regulating a lipid biosynthesis pathway. Proc Natl Acad Sci USA 105:4062–4067

Tan L, Li X, Liu F, Sun X, Li C, Zhu Z, Fu Y, Cai H, Wang X, Xie D, Sun C (2008) Control of a key transition from prostrate to erect growth in rice domestication. Nat Genet 40:1360–1364

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

Vigouroux Y, McMullen M, Hittinger CT, Houchins K, Kresovich S, Matsuoka Y, Doebley J (2002) Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. Proc Natl Acad Sci USA 99:9650–9655

Vitte C, Ishii T, Lamy F, Brar D, Panaud O (2004) Genomic paleontology provides evidence for two distinct origins of Asian rice (*Oryza sativa* L.). Mol Gen Genet 272:504–511

Wang RL, Stec A, Hey J, Lukens L, Doebley JF (1999) The limits of selection during maize domestication. Nature 398:236–239

Wang H, Nussbaum-Wagler T, Li BL, Zhao Q, Vigouroux Y, Faller M, Bomblies K, Lukens L, Doebley JF (2005) The origin of the naked grains of maize. Nature 436:714–719

Wang L, Wang AH, Huang XH, Zhao Q, Dong GJ, Qian Q, Sang T, Han B (2011) Mapping 49 quantitative trait loci at high resolution through sequencing-based genotyping of rice recombination inbred lines. Theor Appl Genet 122:327–340

Watanabe N, Sugiyama K, Yamagishi Y, Sakata Y (2002) Comparative telosomic mapping of homoeologous genes for brittle rachis in tetraploid and hexaploid wheats. Hereditas 137:180–185

Watanabe N, Fujii Y, Kato N, Ban T, Martinek P (2006) Microsatellite mapping of the genes for brittle rachis on homoeologous group 3 chromosomes in tetraploid and hexaploid wheats. J Appl Genet 47:93–98

Wright SI, Vroh Bi I, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD, Gaut BS (2005) The effects of artificial selection on the maize genome. Science 308:1310–1314

Xia Q et al (2009) Complete resequencing of 40 genomes reveals domestication events and genes in silkworm (*Bombyx*). Science 326:433–436

Youens-Clark K, Buckler E, Casstevens T, Chen C, DeClerck G, Derwent P, Dharmawardhana P, Jaiswal P, Kersey P, Karthikeyan AS, Lu J, McCouch SR, Ren L, Spooner W, Stein JC, Thomason J, Wei S, Ware D (2011) Gramene database in 2010: updates and extensions. Nucleic Acids Res 39:D1085–D1094

Yu Y, Tang T, Qian Q, Wang Y, Yan M, Zeng D, Han B, Wu C-I, Shi S, Li J (2008) Independent losses of function in a polyphenol oxidase in rice: Differentiation in grain discoloration between subspecies and the role of positive selection under domestication. Plant Cell 20:2946–2959

Zhang L, Zhu Q, Wu Z, Ross-Ibarra J, Gaut BS, Ge S, Sang T (2009) Selection on grain shattering genes and rates of rice domestication. New Phytol 184:708–720

Zhu Q, Ge S (2005) Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. New Phytol 167:249–265

Zhu B, Si L, Wang Z, Zhou Y, Zhu J, Shangguan Y, Lu D, Fan D, Li C, Lin H, Qian Q, Sang T, Zhou B, Minobe Y, Han B (2011) Genetic control of a transition from black to straw-white seed hull in rice domestication. Plant Physiol 155:1301–1311