



ELSEVIER

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in
Plant Biology

Understanding rice domestication and implications for cultivar improvement

Tao Sang^{1,2} and Song Ge¹

Considerable insights were recently gained into the history and process of rice domestication. It becomes increasingly clear that artificial and natural selections coupled with extensive introgression have shaped the genomes of cultivated rice. The interplay of these evolutionary forces gave rise to the cultivated species, *Oryza sativa*, with divergent genomic backgrounds from two wild species, *O. rufipogon* and *O. nivara*, governed by a set of domestication alleles which had originated primarily at one location of initial cultivation. The mechanistic understanding of domestication suggests that the combination of quantitative trait locus mapping, genome-wide association study, and genome scan will be effective means for discovering potentially valuable alleles from the cultivated and wild species. The accumulation and appropriate sampling of germplasm collections for these analyses should effectively grow the useful allele pool, which combined with molecular breeding may get to a point literally triggering the re-domestication of rice varieties for sustainable food production

Addresses

¹ State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

² Key Laboratory of Plant Resources and Beijing Botanical Garden, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

Corresponding author: Sang, Tao (sang@ibcas.ac.cn)

Current Opinion in Plant Biology 2013, 16:xx–yy

This review comes from a themed issue on **Genome studies and molecular genetics**

Edited by **Qifa Zhang** and **Rod Wing**

1369-5266/\$ – see front matter, © 2013 Elsevier Ltd. All rights reserved.

<http://dx.doi.org/10.1016/j.pbi.2013.03.003>

Introduction

Food crops have gone through many stages of development. The domestication of major cereal crops began approximately 10 000 years ago in different continents, which initiated a long-term evolutionary experiment giving rise to a number of new species adapted to the agricultural system [1,2]. At the blink of evolution eyes in the history of flower plants, these crop species have become amazingly successful and occupied about five percent of the earth's land surface. With the development of high-yield varieties as well as modern agricultural

technologies, the crops are now supporting an ever-large human population with ever-low frequencies of famine.

However, this success did not come without costs. While continuing conversion of natural ecosystem into cropland and increasing uses of synthetic fertilizer and pesticide push up the crop yield, we apparently are running out of options for meeting the growing needs for food. Meanwhile, the over uses of cropland with heavy fertilization have severely reduced the quality and productivity of the cropland. With the high-quality cropland and other natural resources for food crop production getting in the trend of decrease, finding ways to ensure sustainable food production has become an increasingly important and challenging problem for research [3]. In this paper, we approach the problem from an evolutionary point of view, especially with a focus on the origin and evolution of Asian cultivated rice, *Oryza sativa*. By reviewing what we have learned from studying the processes and mechanisms of rice domestication and improvement, we hope to extract instructive information to help breeding rice varieties for sustainable production (Figure 1).

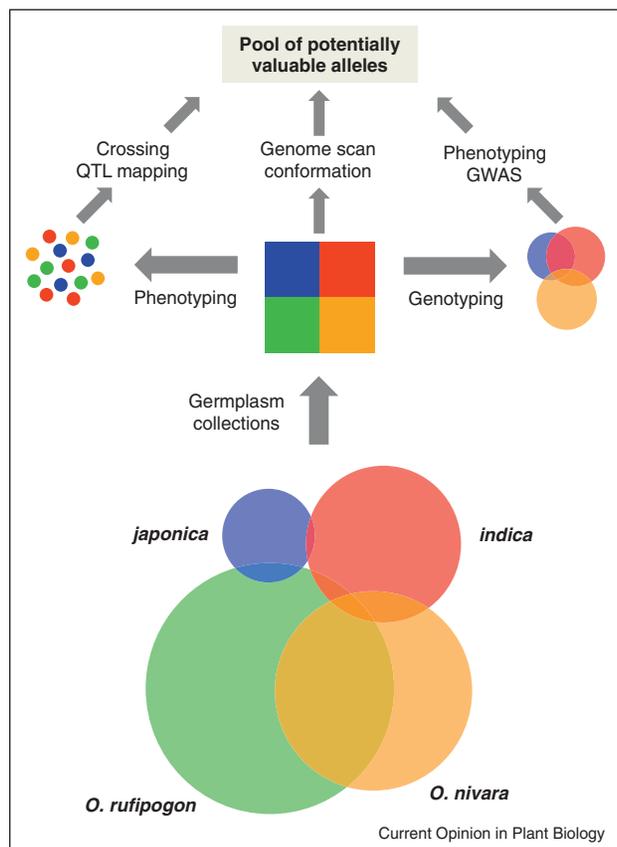
The origin and evolution of cultivated rice was subjected to considerable debate over the past several years. The debate centered on whether the two major rice cultivars, *O. sativa* ssp. *indica* and ssp. *japonica*, were derived from a single ancestor or domesticated independently at different locations [4–6]. Other evolutionary questions revolving this issue included the origin, spreading, and fixation of domestication related alleles. Answers to these questions are directly relevant to what we can learn from the domestication mechanisms and how we apply this knowledge to the future improvement of rice cultivars [7]. Below we will begin with the review of recent progress toward understanding the origin and evolution of cultivated rice and then provide perspectives on rice breeding.

The origin of *indica* and *japonica* cultivars

It has long been recognized that there are two distinct groups of rice cultivars, namely *indica* and *japonica* subspecies of *O. sativa*, which differ in a series of morphological and physiological traits. The difference at the genomic level was also revealed, with multiple lines of evidence suggesting that the genomic cores of the two subspecies have diverged for 200 000 years or longer [8–10]. The fact that the divergence time markedly predates the time of rice domestication suggests from the phylogenetic point of view that *indica* and *japonica* cultivars were domesticated independently from differentiated

2 Genome studies and molecular genetics

Figure 1



Approaches for discovering valuable alleles for molecular breeding of rice varieties for sustainable production. Circles in the lower portion of the figure indicate two subspecies of cultivated rice, *O. sativa* ssp. *japonica* and *O. sativa* ssp. *indica*, and two wild species *O. rufipogon* and *O. nivara*. Sizes of the circles approximate relative genetic diversity of each entity. Overlapping of the circles illustrates (but not intend to precisely represent) shared genomic identity due to demography and/or gene flow. Accessions of cultivars (modern and traditional, elite and landrace) and wild species are preserved in the germplasm collection centers. From these collections, those with desirable phenotypes such as tolerant to biotic and abiotic stresses, can be identified as cross parents of genetic populations developed for QTL mapping and gene cloning (toward left); diverse cultivars and *O. nivara* populations can be genotyped (e.g. through low-coverage genome sequencing) to establish panels for phenotyping and GWAS separately or at least partly together (toward right); candidate genes can be identified by genome scan and have to be substantiated by functional conformation (upward).

ancestral populations of the wild progenitors. This viewpoint had gained growing acceptance until the cloning of key domestication genes.

The first surprise was that the *sh4* gene underlying the reduction of grain shattering had a single origin [11,12]. To reconcile the apparently conflict evidence coming from genomic divergence versus domestication-gene uniformity, two models of rice domestication were proposed [13]. The snowballing model hypothesized that there was a single domestication leading to the development of an

early cultivar with a suite of key domestication traits including the reduction of grain shattering. This early cultivar then hybridized with wild populations with diverged genomes, leading to the fixation of the same critical domestication alleles in cultivars of distinct genomic background. Alternatively, *indica* and *japonica* were domesticated independently from diverged wild progenitors. The subsequent hybridization between them led to the introgression and fixation of the same set of domestication alleles in cultivars with distinct genomic background. This was called the combination model.

No matter which model turns out to be more close to the real domestication processes, there is one thing in common, that is, hybridization or introgression played an essential role in shaping the present genetic and genomic structure of cultivated rice [13–15,16]. It was the interplay of artificial and natural selections that balanced the fixation of domestication alleles and the maintenance of genetic diversity of cultivated rice [16,17]. Strong artificial selection on highly beneficial domestication alleles drove them quickly across the genomic boundaries and consequently fixed in the entire cultivar gene pool. Population genetic simulation estimated that the highly beneficial *sh4* allele could have been fixed over a time period as short as 100 years [18]. This is similar to the breeding programs practiced in the modern agriculture except that modern rice breeding was done purposely from the chosen cross parents whereas domestication practice done thousands of years ago was largely unconscious.

During the domestication processes, natural selection must have played an essential role in maintaining the genetic diversity of rice cultivars. It is conceivable that the genomic divergence between *indica* and *japonica* was maintained and perhaps reinforced by natural selection in different geographic regions. Regional differences in climates and soil conditions to which wild progenitors of *indica* and *japonica* genomic background adapted could have been ecological factors driving the divergence.

There is a partial reproductive barrier between *indica* and *japonica*. It is not known whether this barrier existed between the genome donors of the cultivars or established later on to help maintaining the identities of the cultivars. Given that there had been extensive gene flow at the early stage of rice domestication that facilitated the fixation of domestication alleles across the entire cultivated gene pool and there has not been clearly documented postzygotic isolation between populations of the wild progenitors, it seems likely that the partial reproductive barrier was progressively established during rice domestication to probably reinforce cultivar diversification. If so, the questions of why and how the reproductive isolation was strengthened by either natural or artificial selection or both would be of great interest. The cloning and further

evolutionary analyses of genes involved in hybrid sterility between the two subspecies should begin to provide answers to these questions [19,20,21,22].

Of cereal crops, there are examples of single and multiple origins, for example, maize and barley, respectively [23,24]. The origin of rice was complicated by the fact that the same set of key domestication alleles presented in the diverged genomic background. We now know that extensive introgression followed by artificial and natural selection was responsible for this apparently complex pattern. However, the question of whether the snowballing or combination model was more close to the reality had remained unanswered. A recent study that conducted a large-scale survey of domestication loci lent a support to the snowballing model and further suggested that *japonica* was domesticated at first in southern China and the subsequent introgression of the domestication alleles into wild populations westward gave rise to *indica* [25]. The conclusion was drawn based on the result that a combined analysis of conformed and candidate regions of major domestication sweeps grouped both cultivars with *O. rufipogon* populations in southern China.

While the evidence is convincing, there is still room for further substantiating the hypothesis. As the number of conformed domestication genes grows, each of the genes can be phylogenetically analyzed to trace its origin. Depending on the nature and proportion of these genes traced back to the wild populations and how strongly the relationship is supported, the hypotheses of single versus multiple origins can be fully evaluated. Such an evaluation should be conducted in the context of genetic and ecological divergence of wild progenitors, which will help us to bring together the whole picture of rice domestication.

The divergence of wild progenitors

The wild progenitors of cultivated rice were *O. nivara* and *O. rufipogon*, two most closely related wild species with the current distribution from southeastern Asia to India. *O. nivara* was often regarded as an annual ecotype of *O. rufipogon* in the literature partly because the correct identification of the species was not always possible or reliable for seeds distributed from the germplasm centers [11,15,26]. Nevertheless, the two taxa are morphologically, physiologically, and ecologically distinct in the natural habitats, experimental field, and greenhouses. *O. rufipogon*, adapted to the stable, deep-water habitat, is perennial, predominantly outcrossing, and photoperiod sensitive [27,28]. *O. nivara* was derived from *O. rufipogon* by adapting to the seasonally dry habitat in regions with a clear monsoon season. In order to promptly complete the life cycle in the unstable habitat, *O. nivara* went through a series of adaptive steps and became an annual, self-fertilized, and photoperiod insensitive species [27]. Unlike *O. rufipogon* that relies primarily on vegetative

propagation for reproduction, *O. nivara* reproduces sexually and produces a larger number of bigger and heavier seeds. In most of these characteristics, *O. nivara* is rather similar to the cultivated rice and sometimes viewed as weedy rice derived from crosses between *O. rufipogon* and cultivars. This, however, was not supported by the recent evolutionary studies of weedy rice [29,30].

Thus, domestication of rice from *O. nivara* would require less phenotypic transitions than from *O. rufipogon*. Genetic studies of crosses between *indica* and *O. nivara* did not detect segregation in traits such mating systems, suggesting the possibility of direct domestication of *indica* from *O. nivara* [31]. With the results of genome-wide analysis indicating the domestication of *japonica* rice directly from *O. rufipogon* in southern China where there is no natural distribution of *O. nivara* at the present time, the role that *O. nivara* played in rice domestication remains unclear. There are several possibilities.

First, *O. nivara* contributed little to rice domestication. The phenotypic similarity between *O. nivara* and *O. sativa* was selected in parallel by nature and human and the targeted genes were largely the same for major transitions such as the grain size and mating system. Second, *O. nivara* served as the progenitor of *indica*. However, because *O. nivara* had numerous properties of cultivars, artificial selection was much weaker than that for domesticating *japonica* from *O. rufipogon*. In fact, grains from the natural populations of *O. nivara* were still collected as food in certain areas of India [32]. As a result, the majority of key domestication alleles could have arisen during *japonica* domestication. The presumably semi-domesticated *indica* became the primary recipient of the domestication alleles when *japonica* was brought by immigrants and grown in the same or near field. In this case, there should be at least a small portion of domestication alleles originated from *O. nivara*, which might have been overwhelmed by *japonica* alleles in the combined analysis of multiple chromosomal regions containing domestication genes [25]. The single domestication hypothesis can be further substantiated by ruling out this alternative through separate phylogenetic analyses of domestication genes with *O. nivara* included.

Genes targeted by artificial selection during rice domestication

Evolutionary analyses of domestication related genes should yield mechanistic insights into the efforts that turned wild species into cultivars. These efforts were probably unconsciously in most of the cases. With advances in molecular breeding and biotechnology, we can apply this knowledge to conscious efforts to improve rice cultivars. A half century ago, a 'one gene, one trait' hypothesis was proposed for key domestication traits in maize [33]. With the cloning of genes underlying these critical phenotype transitions, the hypothesis gained

4 Genome studies and molecular genetics

support from studying a number of key domestication traits not only in maize but also in rice [2]. For domestication traits such as non-shattering and tiller erectness, QTL explaining the majority of phenotypic variance harbor causal mutations from the *sh4* and *prog1* genes, respectively [18,34,35].

That a major domestication transition of rice was controlled primarily by one gene is especially striking given that rice cultivars have distinct genomic backgrounds. This further highlights the dynamic processes of rice domestication. If domesticated rice is defined to be one with the set of critical domestication alleles, evolutionary forces, including gene flow, artificial selection, and natural selection, must have acted together to give rise to this major food crop grown in the diverse climatic conditions. The way that rice was domesticated through rapidly spreading beneficial alleles and maintaining genetic diversity at the same time can be borrowed by modern rice breeding.

Genes underlying critical domestication traits often encode transcription factors, and the causal mutations are more likely functional modification rather than the loss or gain of function [1,2]. The functional modification of transcription factors can lead to a cascade of downstream changes that are capable of substantially altering an important trait. The degree of modification needs to be well balanced to optimize an agronomic trait. For example, the weakened function of *sh4* allowed grains to stay on straws at the time of harvest and to be subsequently separated from straws during threshing.

Genes underlying the loss of coloration contributed to a part of the domestication syndrome in cereal crops. The loss involves both hulls and pericarps, with the former changed from black to straw white and the latter changed from red to white. The change of hull colors was controlled by two genes, *Bh4* and *Phr1*, while the change of pericarp colors was controlled by *Rc* and *Rd* genes [36–39]. Interestingly, they share two features different from *sh4* and *prog1*, that is, multiple loss-of-function mutations of each coloration gene were selected during domestication. Of the multiple mutations of a gene, there are one predominant allele and other low-frequency alleles of independent origins [36,40]. Thus, for loss-of-function mutations that are functionally equivalent, selective sweep was not as strong or complete as critical functional modification of *sh4* or *prog1*.

Grain size was also a target of artificial selection. Two genes, *GS3* and *SW5*, were responsible for the increase in grain length and width, respectively [41,42]. A single loss-of-function mutation of each gene was selected during rice domestication. Unlike genes controlling other aspects of the domestication syndrome discussed above, the derived alleles of *GS3* and *SW5* conferring longer and

wider grains occurred in much lower frequencies in cultivated rice. Additionally, the derived alleles have higher frequencies in *japonica* than *indica*, which was explained by their origins in *japonica*. Even within *japonica*, the frequencies of the domestication alleles of the two genes are variable, with that of *GS3* occurring frequently in tropical *japonica* but rarely in temperate *japonica* and that of *SW5* the other way around [43,44]. There are several explanations for this. First, the increase in grain size ran into some limitation such as limited source availability in cultivars with relatively low biomass or a tradeoff between grain size and grain number. Second, the preference of diverse grain shapes led to resistance to the spreading of these alleles in certain cultivars. Third, there were additional genes contributing to grain size increase during domestication.

qSH1 is another gene involved in the reduction of shattering [45]. However, the non-shattering allele did not spread in all cultivars like *sh4*, but was found only in a portion of temperate *japonica* [18]. In addition to *sh4*, this allele further reduces shattering, and consequently requires stronger threshing force to separate and recover grains from straw. This may be beneficial for cultivars that would otherwise suffer significant yield loss by relatively easy shattering under certain growing conditions. However, it could be disadvantageous for cultivars that lost more grains due to incomplete threshing caused by this allele. Thus, even though an allele can enhance a domestication trait, it may only be selectively advantageous in certain genomic background, under certain climatic conditions, or for certain agricultural practice.

A somewhat surprising finding in domestication genes was that the semidwarf allele, *sd1*, responsible for the green revolution taking place a half century ago was also targeted by artificial selection during rice domestication [46]. An allele reducing tiller length experienced a selective sweep in *japonica*. The wild progenitor, *O. rufipogon*, growing in the deep-water habitat is much taller than the cultivated rice. The tillers of *O. rufipogon* lodge easily and new tillers and roots grow out from the internodes, which provides the means of vegetative propagation of the species [27]. This trait is obviously disadvantageous for cultivation and it makes sense during rice domestication that the semidwarf allele was selected and became fixed in *japonica*. It is remarkable that the alleles of the similar function were targeted twice in the history of rice evolution, once by early farmers who might have selected an allele unconsciously during domestication and again by modern rice breeders who purposely spread another allele and led to the revolutionary improvement of yield [47,48].

It was probably also surprising that such a beneficial allele was barely spread into the landraces of *indica*. A partial reproductive barrier was hypothesized to have prevented the introgression of the allele from *japonica* to *indica* [46].

This, however, does not seem to be an adequate explanation because several other domestication alleles were fixed or reached high frequencies in both cultivars. Another hypothesis was that having *sd1* allele and consequently a reduced level of growth hormone, gibberellin, might be adaptively disadvantageous for *indica* landraces planted in the rainfed upland [48]. There is one additional possibility. The wild species, *O. nivara*, which might have played a role in the origin of *indica*, was a short annual plant that did not have the lodging habit [31]. Thus, it was possible that *sd1* from *japonica* was indeed disadvantageous for *indica* cultivars under the ancient cultivation conditions [26].

Implications for rice improvement

One important lesson that we can learn from the evolutionary analyses of rice domestication is the conscious and unconscious maintenance of genetic diversity of cultivars. No matter whether rice domestication initiated at one or multiple locations, diverse cultivars were developed and maintained throughout the history of rice cultivation. Particularly, *indica* has a diverse genomic background, which conceivably came from spreading the common set of domestication alleles into the wild populations [25^{••}]. Looking at this from another angle, it is the process that increased the genetic diversity of cultivars. This somewhat snowballing of genetic diversity that maintained and widened local adaptation was most likely to have facilitated the widespread of rice cultivars [16[•]].

Thus, it should be kept in mind in the future rice breeding that genetic diversity of cultivars should be maintained or increased. Given the sophisticated seed distribution mechanisms at the present time, it is possible that a small number of cultivars with superior performance in a few traits such as yield can quickly increase in the area of cultivation. If this trend continues and further narrows the genetic diversity of modern cultivars, we may end up with lacking solutions to problems caused by the appearance of new pests and diseases and the change of local climates. Also because rice is a food crop that has been grown in a wide range of climates and field conditions, the future breeding practice should always consider regional factors and diversify cultivars accordingly.

With rapid advances in genome technologies, molecular design and marker-assisted breeding are opening enormous opportunities for crop improvement. One of the most instructive lessons learnt from the evolutionary analyses of rice domestication is that a mutation modifying the function of a transcription factor was responsible primarily for the gain of an important domestication trait. This is especially encouraging for the application of molecular breeding as long as a suitable mutation can be identified for improving an agronomic trait. Such mutation may come from several sources, including wild progenitors, landraces, and even modern cultivars. The

development of effective means to discover and utilize these mutations is critical to the success of molecular breeding [49,50].

A widely practiced approach is making crosses between diverged cultivars, especially between *indica* and *japonica*, and subsequently conduct QTL mapping for traits of interest [51]. This has been an effective approach for the identification of numerous valuable alleles for rice improvement. The limitation of this approach was that genotyping and fine mapping were time consuming. The second-generation sequencing technologies have begun to lift this limitation [52,53[•]]. If the efficiency of gene cloning from a mapping population is substantially improved, the matter then is to evaluate extensively cultivars including landraces for valuable traits such as high water and nutrient use efficiencies and resistance to disease and pests. This way of identifying useful mutations is most suitable for those that have clear phenotype but remain in low frequencies probably because the cultivars carrying them perform poorly in many other agronomic traits so that these cultivars have not been selected for breeding elite rice varieties. Rediscovery of these cultivars from germplasm centers through extensive phenotypic evaluation followed by effective gene cloning should substantially enhance our ability to breed green super rice [54].

If the alleles have moderate frequencies in traditional cultivars and landraces, genome-wide association studies (GWAS) may be utilized to directly identify valuable genes. One basic requirement for GWAS is to genotype all accessions sampled in the study. This has become increasingly feasible for large sample sizes as low-coverage whole genome sequencing is proven a promising method [55]. In order to take full advantage of this approach, it is necessary to establish a set of cultivars representing the entire diversity of cultivated rice including traditional and modern cultivars [56[•],57[•]]. Once this set is genotyped by sequencing, they can be evaluated for any potentially useful traits, which are then subjected to GWAS. It should be a worthwhile effort for the international rice community to establish such a genotyped collection and distribute germplasm for GWAS of various traits worldwide.

The phenotypic evaluation of traits valuable for breeding varieties such as green super-rice to meet the current and future needs of sustainable agriculture should also be extended to the wild progenitors of rice, *O. rufipogon* and *O. nivara*. Because they cross easily with rice cultivars, useful genes can be incorporated into cultivated rice through molecular breeding. Once valuable traits are identified in certain accessions of the wild species, they can be either crossed with cultivars for gene cloning or directly narrowed down in the sampled wild accession using GWAS. The application of the GWAS in wild

6 Genome studies and molecular genetics

species, however, is not as straightforward as in cultivars because *O. rufipogon* is an outcrosser and has a high level of heterozygosity. They would have to be vegetatively maintained once genotyped, which increases practical difficulty.

Fortunately, the annual species, *O. nivara*, is predominantly self-fertilized and suitable for GWAS. Thus, it will be interesting to assemble a panel of *O. nivara* accessions for GWAS, especially because the wild species has higher genetic diversity than cultivars and consequently the potential to harbor valuable genes for rice improvement [49,58]. Whether the wild and cultivated panels can be integrated for GWAS depends on whether the subpopulation structure formed as a result of artificial selection can be effectively dealt with in the analyses [56*,59].

A somewhat different approach to identify candidate genes responsible for domestication and cultivar differentiation is genome scan for selection signals [60]. By identifying genomic regions with significantly reduced genetic diversity of cultivars relative to the wild progenitors, candidate genes targeted by artificial selection may be located in the regions [12,61*]. Candidate genes responsible for cultivar group differentiation may also be identified through similar comparison between those groups. This is a rapid approach for identifying potentially useful genes, especially when powered by the availability of dense SNP markers. However, function conformation of the candidate genes requires additional work.

In conclusion, with rapid advances in DNA sequencing technologies, evolutionary genomic analyses are becoming increasingly powerful for discovering valuable alleles from cultivated rice and its wild progenitors. The incorporation of these alleles into cultivars through molecular breeding holds enormous potential for rice improvement. With the expansion of the allele pool and its intensified utilization, we should approach a point where a large number of alleles from the wild species and landraces can be incorporated for developing cultivars meeting the requirement of sustainable agriculture. This may be viewed as the re-domestication of rice toward adequate and sustainable production of the world's staple food.

Acknowledgements

The work was supported by the Key Program of the National Natural Science Foundation of China (91131902) and the Knowledge Innovation Program of the Chinese Academy of Sciences (KSCX2-EX-QR-1) to TS and the National Natural Science Foundation of China (91231201) to SG.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
 - of outstanding interest
1. Doebley JF, Gaut BS, Smith BD: **The molecular genetics of crop domestication.** *Cell* 2006, **127**:1309-1321.
 2. Sang T: **Genes and mutations underlying domestication transitions in grasses.** *Plant Physiol* 2009, **149**:63-70.
 3. Godfray HCJ, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, Pretty J, Robinson S, Thomas SM: **Camilla Toulmin: food security: the challenge of feeding 9 billion people.** *Science* 2010, **327**:812-818.
 4. Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA: **Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*.** *Proc Natl Acad Sci U S A* 2006, **103**:9578-9583.
 5. Molina J, Sikora M, Garud N, Flowers JM, Rubinstein S, Rynolds A, Huang P, Jackson SA, Schaal BA, Bustanante CD *et al.*: **Molecular evidence for a single evolutionary origin of domesticated rice.** *Proc Natl Acad Sci U S A* 2011, **108**:8351-8356.
 6. Ge S, Sang T: **Inappropriate model rejects independent domestications of *indica* and *japonica* rice.** *Proc Natl Acad Sci U S A* 2011, **108**:E755.
 7. Hufford MB, Xu X, van Heerwaarden J, Pyhajarvi T, Chia J-M, Cartwright RA, Elshire RJ, Glaubitz JC, Guill KE, Kaeppeler SM *et al.*: **Comparative population genomics of maize domestication and improvement.** *Nat Genet* 2011, **44**:808-811.
 8. Ma J, Bennetzen JL: **Rapid recent growth and divergence of rice nuclear genomes.** *Proc Natl Acad Sci U S A* 2004, **101**:12404-12410.
 9. Vitte C, Ishii T, Lamy F, Brar D, Panaud O: **Genomic paleontology provides evidence for two distinct origins of Asian rice (*Oryza sativa* L.).** *Mol Gen Genet* 2004, **272**:504-511.
 10. Zhu Q, Ge S: **Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes.** *New Phytol* 2005, **167**:249-265.
 11. Li C, Zhou A, Sang T: **Rice domestication by reducing shattering.** *Science* 2006, **311**:1936-1939.
 12. Lin Z, Griffith ME, Li X, Zhu Z, Tan L, Fu Y, Zhang W, Wang X, Xie D, Sun C: **Origin of seed shattering in rice (*Oryza sativa* L.).** *Planta* 2007, **226**:11-20.
 13. Sang T, Ge S: **The puzzle of rice domestication.** *J Integr Plant Biol* 2007, **49**:760-768.
 14. Kovach MJ, Sweeney MT, McCouch SR: **New insights into the history of rice domestication.** *Trends Genet* 2007, **23**:578-587.
 15. Zhao K, Wright M, Kimball J, Eizenga G, McClung A, Kovach M, Tyagi W, Ali ML, Tung C-W, Reynolds A *et al.*: **Genomic diversity and introgression in *O. sativa* reveal the impact of domestication and breeding on the rice genome.** *PLoS One* 2010, **5**:e10780.
 16. He Z, Zhai W, Wen H, Tang T, Wang Y, Lu X, Greenberg AJ, Hudson RR, Wu C-I, Shi S: **Two evolutionary histories in the genome of rice: the role of domestication genes.** *PLoS Genet* 2011, **7**:e1002100.
- Using Illumina Solexa and SOLiD platforms, this study sequenced *indica* and *japonica* subspecies of cultivated rice and the wild progenitor. By focusing on low diversity regions (LDRs) shared by the two subspecies, the authors found that the genealogical histories of these LDRs were distinct from their divergent genomic backgrounds. The results suggested that some of the overlapping LDRs bearing domestication genes might have originated only once and became fixed through introgression in genomes diverged before rice domestication.
17. Sang T, Ge S: **Genetics and phylogenetics of rice domestication.** *Cur Opin Genet Dev* 2007, **17**:533-538.
 18. Zhang L, Zhu Q, Wu Z, Ross-Ibarra J, Gaut BS, Ge S, Sang T: **Selection on grain shattering genes and rates of rice domestication.** *New Phytol* 2009, **184**:708-720.
 19. Long Y, Zhao L, Niu B, Su J, Wu H, Chen Y, Zhang Q, Guo J, Zhuang C, Mei M *et al.*: **Hybrid male sterility in rice controlled by interaction between divergent alleles of two adjacent genes.** *Proc Natl Acad Sci U S A* 2008, **105**:18871-18876.
 20. Chen J, Ding J, Ouyang Y, Du H, Yang J, Cheng K, Zhao J, Qiu S, Zhang X, Yao J *et al.*: **A triallelic system of S5 is a major regulator of the reproductive barrier and compatibility of**

- indica-japonica* hybrids in rice. *Proc Natl Acad Sci U S A* 2008, **105**:11436-11441.
21. Du H, Ouyang Y, Zhang C, Zhang Q: **Complex evolution of S5, a major reproductive barrier regulator, in the cultivated rice *Oryza sativa* and its wild relative.** *New Phytol* 2011, **191**:275-287.
 22. Yang L, Zhao X, Cheng K, Du H, Ouyang Y, Chen J, Qiu S, Huang J, Jiang Y, Jiang L *et al.*: **A killer-protector system regulates both hybrid sterility and segregation distortion in rice.** *Science* 2012, **337**:1336-1340.
 23. Doebley J: **The genetics of maize evolution.** *Annu Rev Genet* 2004, **38**:37-59.
 24. Morrell PL, Clegg MT: **Genetic evidence for a second domestication of barley (*Hordeum vulgare*) east of the Fertile Crescent.** *Proc Natl Acad Sci U S A* 2007, **104**:3289-3294.
 25. Huang X, Kurata N, Wei X, Wang Z-X, Wang A, Zhao Q, Zhun Y, Liu K, Lu H, Li W *et al.*: **A map of rice genome variation reveals the origin of cultivated rice.** *Nature* 2012, **490**:497-501.
- This study re-sequenced 446 genomes of the wild progenitor of rice, *O. rufipogon*. In conjunction with the whole-genome sequencing data of 1083 rice accessions, the authors constructed a detailed map of rice genome variation and characterized sequence variation of chromosomal regions contained domestication-related genes. Particularly, by identifying and analyzing 55 regions of selective sweeps involved in rice domestication, this study suggested that *japonica* rice was first domesticated from the wild populations in the middle Pearl River regions in southern China, and that *indica* rice was subsequently domesticated through introgressing of domestication genes from *japonica* into wild populations toward southeastern and southern Asia.
26. Onishi K, Horiuchi Y, Ishigoh-Oka N, Takagi K, Ichikawa N, Maruoka M, Sano Y: **A QTL cluster for plant architecture and its ecological significance in Asian wild rice.** *Breed Sci* 2007, **57**:7-16.
 27. Grillo MA, Li C, Fowlkes AM, Briggeman TM, Zhou A, Schemske DW, Sang T: **Genetic architecture for the adaptive origin of annual wild rice, *Oryza nivara*.** *Evolution* 2009, **63**:870-883.
 28. Zheng X-M, Ge S: **Ecological divergence in the presence of gene flow in two closely related *Oryza* species (*Oryza rufipogon* and *O. nivara*).** *Mol Ecol* 2010, **19**:2439-2454.
 29. Thurber CS, Reagon M, Gross BL, Olsen KM, Jia Y, Caicedo AL: **Molecular evolution of shattering loci in U.S. weedy rice.** *Mol Ecol* 2010, **19**:3271-3284.
 30. Sun J, Qian Q, Ma D-R, Xu Z-J, Liu D, Du H-B, Chen W-F: **Introgression and selection shaping the genome and adaptive loci of weedy rice in northern China.** *New Phytol* 2012. doi:10.1111/nph.12012.
 31. Li C, Zhou A, Sang T: **Genetic analysis of rice domestication syndrome with the wild annual species, *Oryza nivara*.** *New Phytol* 2006, **170**:185-194.
 32. Oka HI: *Origin of Cultivated Rice*. Japan Scientific Society Press; 1988.
 33. Doebley JF: **George Beadle's other hypothesis: one-gene, one-trait.** *Genetics* 2001, **158**:487-493.
 34. Jin J, Huang W, Gao J-P, Yang J, Shi M, Zhu M-Z, Luo D, Lin H-X: **Genetic control of rice plant architecture under domestication.** *Nat Genet* 2008, **40**:1365-1369.
 35. Tan L, Li X, Liu F, Sun X, Li C, Zhu Z, Fu Y, Cai H, Wang X, Xie D, Sun C: **Control of a key transition from prostrate to erect growth in rice domestication.** *Nat Genet* 2008, **40**:1360-1364.
 36. Zhu B, Si L, Wang Z, Zhou Y, Zhu J, Shangguan Y, Lv D, Fan D, Li C, Lin HQ *et al.*: **Genetic control of a transition from black to straw-white seed hull in rice domestication.** *Plant Physiol* 2001, **155**:1301-1311.
 37. Yu Y, Tang T, Qian Q, Wang Y, Yan M, Zeng D, Han B, Wu C-I, Shi S, Li J: **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* 2008, **20**:2946-2959.
 38. Sweeney MT, Thomson MJ, Pfeil BE, McCouch S: **Caught red-handed: Rc encodes a basic helix-loop-helix protein conditioning red pericarp in rice.** *Plant Cell* 2006, **18**:283-294.
 39. Furukawa T, Maekawa M, Oki T, Suda I, Lida S, Shimada H, Takamura I, Kadowaki K: **The Rc and Rd genes are involved in proanthocyanidin synthesis in rice pericarp.** *Plant J* 2007, **49**:91-102.
 40. Sweeney MT, Thomson MJ, Cho YG, Park YJ, Williamson SH, Bustamante CD, McCouch SR: **Global dissemination of a single mutation conferring white pericarp in rice.** *PLoS Genet* 2007, **3**:e133.
 41. Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X, Zhang Q: **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* 2006, **112**:1164-1171.
 42. Shomura A, Izawa T, Ebana K, Ebitani T, Kanegae H, Konishi S, Yano M: **Deletion in a gene associated with grain size increased yields during rice domestication.** *Nat Genet* 2008, **40**:1023-1028.
 43. Takano-Kai N, Jiang H, Kubo T, Sweeney M, Matsumoto T, Kanamori H, Padhukasahasram B, Bustamante C, Yoshimura A, Doi K *et al.*: **Evolutionary history of GS3, a gene conferring grain length in rice.** *Genetics* 2009, **182**:1323-1334.
 44. Izawa T, Konishi S, Shomura A, Yano M: **DNA changes tell us about rice domestication.** *Curr Opin Plant Biol* 2009, **12**:185-192.
 45. Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M: **An SNP caused loss of seed shattering during rice domestication.** *Science* 2006, **312**:1392-1396.
 46. Asano K, Yamasaki M, Takuno S, Miura K, Katagiri S, Ito T, Doi K, Wu J, Ebana K, Matsumoto T: **Artificial selection for a green revolution gene during japonica rice domestication.** *Proc Natl Acad Sci U S A* 2011, **108**:11034-11039.
- Using the backcross inbred lines between *japonica* and *indica* (Kasalath) cultivars, this study identified two functional nucleotide polymorphisms (FNPs) at the *SD1* locus (known as the 'green revolution' gene), which were responsible for reduced gibberellin biosynthetic activity and consequently shorter culms. Population genetic analysis of cultivars and the wild ancestor (*O. rufipogon*) revealed that *japonica* landraces showed a dramatic reduction in nucleotide diversity around *SD1* relative to the *indica* landraces and *O. rufipogon*. These results indicated that *SD1* was subjected to artificial selection during *japonica* domestication, suggesting that the 'green revolution gene' was targeted by artificial selection leading to rice domestication.
47. Sasaki A, Ashikari M, Ueguchi-Tanaka M, Itoh H, Nishimura A, Swapan D, Ishiyama K, Saito T, Kobayashi M, Khush GS *et al.*: **A mutant gibberellin-synthesis gene in rice.** *Nature* 2002, **416**:701-702.
 48. Paterson AH, Li Z-K: **Paleo-green revolution for rice.** *Proc Natl Acad Sci U S A* 2011, **108**:10931-10932.
 49. Kovach MJ, McCouch SR: **Leveraging natural diversity: back through the bottleneck.** *Curr Opin Plant Biol* 2008, **11**:193-200.
 50. Glaszmann JC, Kilan B, Upadhyaya HD, Varshney RK: **Accessing genetic diversity for crop improvement.** *Curr Opin Plant Biol* 2010, **13**:167-173.
 51. Takeda S, Matsuoka M: **Genetic approaches to crop improvement: responding to environmental and populatin changes.** *Nat Rev Genet* 2008, **9**:444-457.
 52. Huang XH, Feng Q, Qian Q, Zhao Q, Wang L, Wang AH, Guan JP, Fan DL, Weng QJ, Huang T *et al.*: **High-throughput genotyping by whole-genome resequencing.** *Genome Res* 2009, **19**:1068-1076.
 53. Wang L, Wang AH, Huang XH, Zhao Q, Dong GJ, Qian Q, Sang T, Han B: **Mapping 49 quantitative trait loci at high resolution through sequencing-based genotyping of rice recombination inbred lines.** *Theor Appl Genet* 2011, **122**:327-340.
- The study demonstrated that QTL analysis supported by dense SNP markers obtained from next-generation sequencing could provide a resolution fine enough to allow direct identification of candidate genes. This approach has the potential to considerably speed up the process of map-based QTL cloning.

8 Genome studies and molecular genetics

54. Zhang Q: **Strategies for developing green super rice.** *Proc Natl Acad Sci U S A* 2007, **104**:16402-16409.
55. Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z *et al.*: **Genome-wide association studies of 14 agronomic traits in rice landraces.** *Nat Genet* 2010, **42**:961-967.
56. Huang X, Zhao Y, Wei X, Wang A, Zhao Q, Li W, Guo Y, Deng L, Zhu C, Fan D *et al.*: **Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm.** *Nat Genet* 2012, **44**:32-39.
- The study showed that genome maps supported by dense SNPs generated by next-generation sequencing are powerful tool for genome-wide association study of important agronomical traits with *indica* and *japonica* cultivars sampled and analyzed together. The resolution could reach the gene level.
57. Zhao K, Tung C-W, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islan MR, Reynolds A, Mezey J *et al.*: **Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*.** *Nat Commun* 2011, **2**:467.
- The study demonstrated that the genome-wide associate study of samples representing the world rice diversity was a promising approach for identifying genes potentially valuable for rice improvement.
58. Zhu Q, Zheng X, Luo J, Gaut BS, Ge S: **Multilocus analysis of nucleotide variation of *Oryza sativa* and its wild relatives: severe bottleneck during domestication of rice.** *Mol Biol Evol* 2007, **24**:857-888.
59. Hamblin MT, Buckler ES, Jannink JL: **Population genetics of genomics-based crop improvement methods.** *Trends Genet* 2011, **27**:98-106.
60. Ross-Ibarra J, Morrell P, Gaut B: **Plant domestication, a unique opportunity to identify the genetic basis of adaptation.** *Proc Natl Acad Sci U S A* 2007, **104**:8643-8648.
61. Xu X, Liu X, Ge S, Jensen JD, Hu F, Li X, Dong Y, Gutenkunst RN, Fang L, Huang L *et al.*: **Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes.** *Nat Biotechnol* 2012, **30**:105-111.
- On the basis of 6.5 million high-quality single nucleotide polymorphisms obtained from genome re-sequencing, the analysis recognized four phylogenetic coherent groups, each of the two cultivars, *indica* and *japonica*, and each of the two wild species, *O. rufipogon* and *O. nivara*. Further analyses indicated that *indica* was more closely related to *O. nivara* whereas *japonica* was closer to *O. rufipogon*, favoring the hypothesis of independent domestication. The authors identified thousands of genes with significantly lower diversity in cultivated but not wild rice, representing candidate regions targeted by artificial selection during rice domestication.