ELSEVIER

Leveraging natural diversity: back through the bottleneck
MJ Kovach and SR McCouch

Plant breeders have long recognized the existence of useful genetic variation in the wild ancestors of our domesticated crop species. In cultivated rice (*Oryza sativa*), crosses between high-yielding elite cultivars and low-yielding wild accessions often give rise to superior offspring, with wild alleles conferring increased performance in the context of the elite cultivar genetic background. Because the breeding value of wild germplasm cannot be determined by examining the performance of wild accessions, a phylogenetic approach is recommended to determine which interspecific combinations are most likely to be useful in a breeding program. As we deepen our understanding of how genetic diversity is partitioned within and between cultivated and wild gene pools of *Oryza*, breeders will have increased power to make predictions about the most efficient strategies for utilizing wild germplasm for rice improvement.

Address
Department of Plant Breeding and Genetics, 162 Emerson Hall, Cornell University, Ithaca, NY 14853, USA

Corresponding author: McCouch, SR (srm4@cornell.edu)

Introduction
Approximately 10,000 years ago, Neolithic hunter–gatherers throughout Asia began to collect wild rice and impose unconscious selection, marking the beginning of a complex history of rice domestication. As people slowly tamed wild rice and learned to mold their societies around the requirements of rice production, they created the world's most enduring monoculture, and in turn became entirely dependent on this cereal for their daily sustenance. The global dependency on cultivated rice (*Oryza sativa*) has continued to strengthen, as now nearly half the world's population relies on rice as a staple food [1]. But how did the unruly wild rice encountered by early human societies become transformed into the domesticated, high-yielding varieties of rice we have today? This change was made possible by the existence of natural genetic variation, which humans harnessed by selecting

for favorable traits. Modern breeding practices continue to follow in the footsteps of our ancestors as we seek new sources of genetic variation for rice improvement. This review will outline how breeders are making use of the natural diversity in *O. sativa* and discuss the usefulness of wild rice germplasm for rice improvement. A framework for generating the maximum amount of useful genetic novelty in an efficient and predictive manner will be presented, along with the challenges associated with this strategy.

Genetic bottlenecks and the evolution of population substructure in *O. sativa*
During the initial process of rice domestication, key traits such as diminished grain shattering (see Glossary) and less persistent grain dormancy (see Glossary) were strongly selected for by humans [2]. Preferential propagation of individuals possessing these valuable traits would have created a primary domestication bottleneck. During this genetic bottleneck, many undesirable alleles from the wild ancestor, along with some potentially beneficial alleles, were not carried through to early domesticates (landraces), resulting in a narrowing of the domesticated rice gene pool [3]. Modern plant breeding continues to constrain the genetic diversity of cultivated rice by selecting for optimal performance under a highly managed set of agricultural conditions [4,5]. Therefore, modern cultivated rice is estimated to retain only approximately 10–20% of the genetic diversity present in its wild rice ancestor, *O. rufipogon* [6,7•].

Two genetically distinct groups within *O. sativa*, *indica* and *japonica*, have been recognized since ancient times [reviewed in [8•]]. These two varietal groups (sometimes referred to as subspecies) are believed to have been domesticated from geographically overlapping, yet genetically divergent populations of *O. rufipogon* (Figure 1), a concept that is extensively supported by molecular evidence [7•, 9–14,15•,16,17•]. Further, sequence comparisons of chloroplast, mitochondrial and nuclear genomes suggest that the divergence of the *indica* and *japonica* gene pools predates the earliest archaeological evidence for rice domestication by 50–100,000 years [18,19,20•,21].

The two major varietal groups in *O. sativa* are further differentiated into five distinct subpopulations that can be clearly diagnosed using isozyme, simple sequence repeat (SSR), chloroplast, and/or single nucleotide polymorphism (SNP) markers [7•,22,23••]. Based on these studies, the *japonica* varietal group can be divided into the *temperate japonica*, *tropical japonica*, and *aromatic/basmati* subpopulations while the *indica* varietal group contains

the *indica* and *aus* subpopulations. These subpopulations are well differentiated from each other, as indicated by pairwise F_{ST} values ranging from 0.2 to 0.42 [23••], and they also differ in effective population size as a result of the proximity, duration and severity of the population bottlenecks experienced by each [7•,23••] (Figure 2).

The process of rice domestication was accompanied by a gradual shift from the primarily out-crossing wild ancestor, *O. rufipogon*, to the primarily inbreeding domesticated species, *O. sativa*. This shift in mating system is largely responsible for the partitioning of genetic variation into the present-day subpopulations of *O. sativa*. Despite this shift in mating system, there is evidence of gene flow among the early *O. sativa* domesticates and between them and their sympatric wild relatives, infusing alleles that contributed in different ways to the diversity of each of the cultivated subpopulations [8•]. Several in-depth studies demonstrate that key domestication alleles are shared between the *indica* and *japonica* varietal groups [24–26], suggesting that rice domestication involved multiple genetic bottlenecks, coupled with episodes of hybridization and introgression between early rice domesticates from divergent gene pools [27•,28•] (Figure 1).

Implications of subpopulation structure on rice breeding

The deep genetic differentiation among *O. sativa* subpopulations has several major implications for rice breeders. First, it is accompanied by intraspecific sterility barriers and reproductive incompatibilities that make it difficult to recover a full array of viable recombinant offspring when crosses are made between the *indica* and *japonica* varietal groups [8•,29]. As a result, rice breeders have historically focused on crosses between genotypes within a varietal group (i.e. *temperate japonica × tropical japonica*; *indica × indica*) [30–32].

A more provocative implication of the subpopulation structure in rice is that it provides rice breeders with a suite of naturally occurring, highly divergent gene pools that can appropriately be considered 'heterotic' or 'combinability' groups. In the traditional sense, heterotic groups are populations of a species that are sufficiently divergent so that when crosses are made between the groups, a significant amount of heterosis (hybrid vigor) may be observed in the F_1 generation [33]. In rice, several

Figure 1
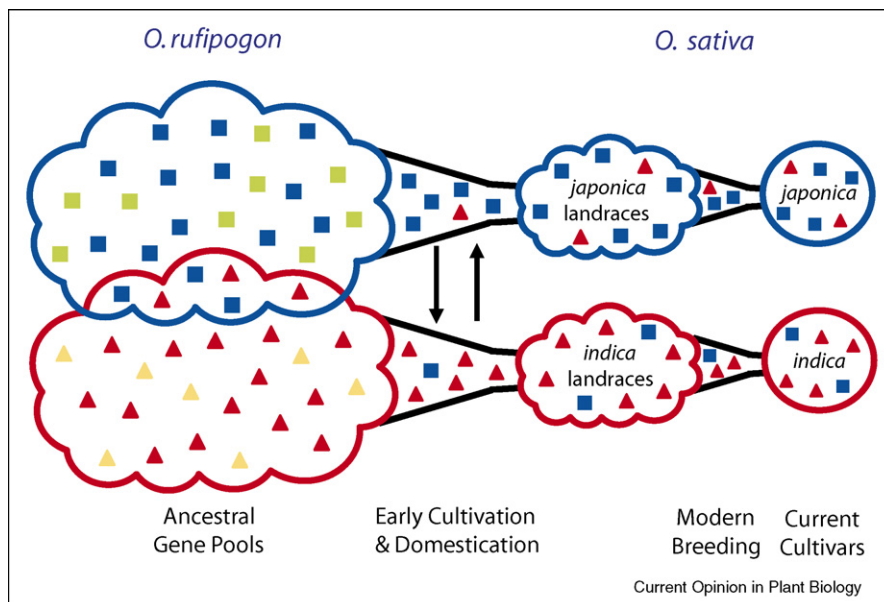
The complex domestication process of *O. sativa*. In contrast to the linear domestication bottleneck model proposed by Tanksley and McCouch [3], the domestication process in *O. sativa* was considerably more complex. Phylogenetic, molecular, and archaeological evidence support the concept that diverse ancestral *O. rufipogon* populations existed over a broad geographical range across Asia and that multiple *O. rufipogon* populations gave rise to at least two primary domesticated varietal groups, namely the *indica* and *japonica* gene pools of domesticated rice. The cloud shapes represent the ancient gene pools of *O. rufipogon* and *O. sativa* that gave rise to modern cultivars. The triangle (*indica*-specific) and square (*japonica*-specific) shapes represent alleles; some of which (red/blue colors) were carried through the domestication bottleneck to modern cultivars, while others (yellow/green colors) were left behind in the wild species. Gene flow between early *indica* and *japonica* domesticates is depicted by arrows between the gene pools.

Figure 2
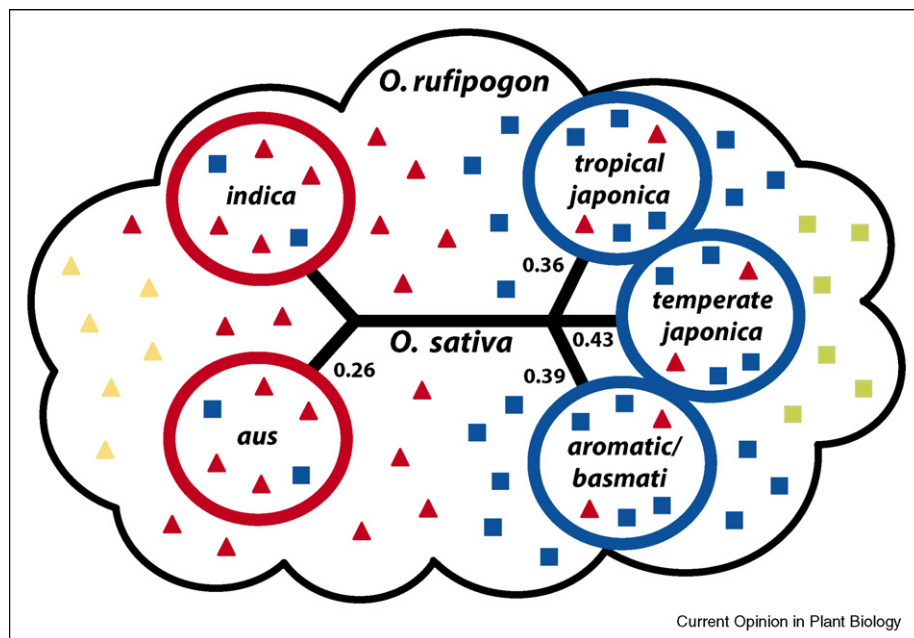
Subpopulation structure of *O. sativa*. In-depth genetic analysis has revealed that the two varietal groups in *O. sativa* (*indica* and *japonica*) are further subdivided into five distinct subpopulations: *indica* and *aus* (in the *indica* varietal group) and *temperate japonica*, *tropical japonica*, and *aromatic/basmati* (in the *japonica* varietal group) [7•,21,23••]. Circles representing the five *O. sativa* subpopulations are colored to indicate their relationship to the two varietal groups (*indica* = red, *japonica* = blue); domesticated subpopulations are superimposed over the large and diverse *O. rufipogon* ancestral gene pool where *indica*- alleles are represented by triangles and *japonica*-alleles are represented by squares; allele distribution within the *O. rufipogon* cloud indicates that some wild genotypes are more closely related to certain *O. sativa* genotypes than to each other. The five groups are highly differentiated from each other, as evidenced by the large F_{ST} values [23••]. Pairwise F_{ST} values relative to the *indica* subpopulation are indicated along the branches of the tree.

decades of work in China have demonstrated that F_1 hybrids derived from crosses between divergent rice subpopulations are generally more productive than F_1 hybrids derived from crosses between closely related cultivars, as long as the sterility barriers are carefully managed [34,35].

Divergent rice populations have also been proposed to function as combinability groups for exploiting transgressive variation during the development of superior inbred varieties [36••]. Transgressive variation is a phenomenon that is recognized by the appearance of individuals in the progeny of a cross that exceed the performance of the better parent. It is generally explained by the fact that most lines contain a distribution of both positive and negative alleles that contribute to an intermediate phenotype. When genetically divergent parents are crossed, recombination in the offspring gives rise to segregants (individual progeny) that are more extreme than either parent because they contain higher frequencies of either favorable or unfavorable alleles. The exploitation of this phenomenon has guided the improvement of inbred varieties where divergent germplasm resources are used to generate transgressive variation, which breeders can then fix in the elite backgrounds of interest [36••,37,38•].

Wild germplasm as a resource for capturing positive transgressive segregation

Rice breeders today face the formidable challenge of achieving the pest resistance, stress tolerance, yield, and quality improvements that will be necessary to keep pace with rising global food requirements. The probability of success in this endeavor depends to a great extent on our ability to make use of novel sources of genetic variation. One way to do this is to explore the largely untapped reservoir of allelic diversity that remains hidden within existing populations of early landraces and wild relatives. New technology makes it possible to readily identify wild alleles that were left behind by ancient farmers and to selectively harness those that enhance performance when introduced into our highly productive modern varieties.

Landraces of *O. sativa* are genetic intermediates between wild ancestors and modern, elite cultivars. Having been selected for alleles and adaptive gene complexes that are favorable to humans, they represent a rich pool of genetic diversity that is readily accessible to modern rice breeders. Why then should we look to the poor performing, low-yielding rice ancestors for the novelty necessary for rice improvement? One reason is that all elite cultivars are

Figure 3
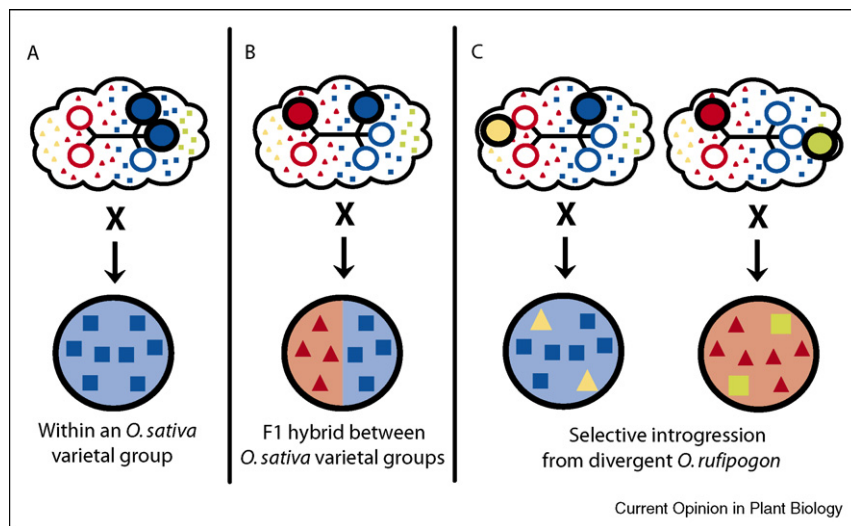
Rice breeding options: How to generate novelty? Traditional rice breeding has generated elite cultivars derived from crosses between genetically similar germplasm, such as between members of the same varietal group (panel **A**). While this avoids potential problems with reproductive barriers and quality issues, little genetic novelty is available for enhancing the performance of the cultivar. By contrast, F_1 hybrids between genetically divergent groups, such as between *indica* and *japonica* parents, bring together alleles that contribute to heterosis (panel **B**). A third option is to selectively introgress genes from genetically divergent germplasm (i.e. from *O. rufipogon*) into elite *O. sativa* cultivars, creating introgression lines that exhibit positive transgressive variation (panel **C**). It is of interest to determine whether carefully crafted introgression lines make it possible for inbred varieties to equal or outperform F_1 hybrids. The small filled circles with a bold outline at the top of the panel indicate the germplasm pools being crossed (corresponding to the subpopulations indicated in Figure 2 or wild germplasm pools). The large circles at the bottom of each panel represent individual genotypes resulting from a particular cross; the red triangles and blue squares represent *indica*-like and *japonica*-like alleles, respectively; yellow triangles and green squares represent alleles from *O. rufipogon* that were left behind during the domestication bottleneck, but are re-introduced into inbred elite lines through selective backcrossing. Panel **(C)** shows two introgression lines, where divergent *O. rufipogon* alleles were introduced into either a *japonica* (blue) or an *indica* (red) genetic background.

the result of selections from landrace varieties that have themselves been through the primary domestication bottleneck. Modern varieties are therefore expected to share a higher proportion of alleles with landraces than with wild accessions. Thus, the probability of generating novel genetic variation from crosses between elite varieties and wild germplasm is greater than would be expected from crosses to landrace materials. In addition, crosses between elite cultivars and wild germplasm generally present fewer reproductive barriers than do crosses between *indica* and *japonica* cultivars [8,29,36**].

Numerous studies report improvements in performance because of the introgression of valuable genes from wild germplasm into elite rice cultivars. Historically, breeders identified phenotypes such as disease resistance or male-sterility in a wild rice species and then introduced the trait through backcross breeding (reviewed in [39]). More recently, the use of advanced backcross quantitative trait locus (QTL) analysis and near isogenic lines (NILs) have made it technically and economically feasible to identify and selectively introgress genes or QTL that confer superior performance in the genetic background of an elite cultivar but that have no observable phenotype in the wild donor (Figure 3). Despite its inferior yield and

agronomic performance, *O. rufipogon* has been the source of beneficial alleles for diverse quantitative traits including grain size, grain weight [40,41], grain yield [36**,38*,41–47,48*,49*,50], grain quality [51], cold tolerance [52], aluminum tolerance [53], and flowering time [54]. Yield and grain quality enhancing alleles have also been identified from *O. glaberrima* [55–57] and *O. glumaepatula* [58,59].

In cases where genes are introgressed from genetically divergent, low-performing wild or weedy donors, the alleles of interest are associated with positive transgressive variation in elite genetic backgrounds (Figure 4). This phenomenon has been demonstrated through interspecific crosses in many crop species [60–67], highlighting the potential usefulness of exploring exotic germplasm sources for the improvement of a wide array of domesticated crop species.

Making the most of transgressive variation in rice

One method of delivering superior varieties to farmers is through F_1 hybrid technology. In China, the superior performance of hybrid versus inbred rice varieties has resulted in the expansion of hybrid production to approxi-

Figure 4
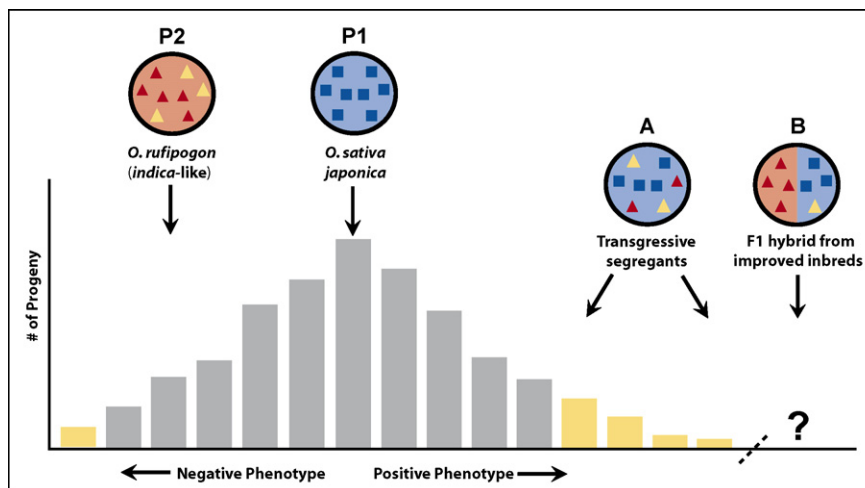
Transgressive segregation. Transgressive segregation is observed in crosses between genetically divergent genotypes where the progeny exceed the performance of the parents. This figure displays the phenotypic distribution of the progeny of a hypothetical cross between an elite *japonica* cultivar (P1) and an *indica*-like *O. rufipogon* accession (P2). The average performance of P1 and P2 are indicated by arrows. Since the two parents do not share many of the same alleles, there is a high probability that some of the progeny **(A)** will possess novel combinations of alleles that confer a performance advantage. Selective introgression of beneficial alleles from exotic germplasm sources offers a way to expand the gene pool of modern cultivars without disrupting many of the gene complexes that contribute to the quality and adaptation of elite cultivars. Divergent introgression lines could then be used as parents to create F_1 hybrids **(B)** that aim to maximize the heterotic potential of *O. sativa*.

mately 50% of the total rice production [35]. For the last 80 years, hybrid technology has been largely responsible for the steady increase in maize yields in the United States (~1% per year) and more recently for sorghum as well [68–70]. In light of these successes, and of the industry's interest in the hybrid model, it is tempting to assume that hybrids will drive increases in productivity for most crops in the future. However, the relative costs and benefits of hybrid versus inbred variety development in inbreeding versus out-crossing species, and in high-value versus low-value crops, suggest that the answer is not entirely clear.

The perfect, cleistogamous flowers (see Glossary) of rice make it difficult and costly to reliably obtain out-crossed F_1 seed, even with the development of both two- and three-line male sterility systems [71]. In addition, genetic evidence suggests that overdominance does not appear to be the major cause of heterosis in *O. sativa*, making it possible to capture complimentary alleles in inbred varieties. Further, because inbreeding crops have expunged most deleterious recessive alleles over the course of evolution, the heterozygosity provided by F_1 hybrids is not theoretically necessary to achieve superior performance [72,73].

Thus, we propose that rice breeders can capture a large portion of heterosis in inbred varieties and that wild germplasm represents an underutilized source of novel alleles. To approach this systematically, genetic diversity within *O. rufipogon* must be characterized and the relation-ships between subpopulations of *O. sativa* and *O. rufipogon* must be defined. Several studies have reported ecological and/or geographical population substructure in *O. rufipogon*. While the relationship between the wild and culti-vated subpopulations is complicated by the substantial and well-documented gene flow between them [74,75], there are subpopulations of *O. rufipogon* that cluster nearer to some subpopulations of *O. sativa* than others [13,14,15•,16,17•] (Figure 2). We can, therefore, use a phylogenetic approach to select wild genotypes that are genetically divergent from target elite cultivars to use as parents. This strategy will help to maximize the prob-ability of creating useful transgressive segregation from which to select superior phenotypes (Figure 3). By intro-gressing a few, selected chromosomal segments (QTLs) from genetically divergent wild donors, breeders aim to move existing elite rice cultivars 'up the fitness land-scape' [76], fixing positive transgressive segregants through repeated backcrossing and selfing [36••]. Once these 'wild QTLs' are fixed in improved inbred varieties, they may also be useful to hybrid breeders who can take advantage of them to create a new generation of superior hybrids [77] (Figure 4).

Conclusions—future challenges and needs

The immediate wild ancestor of rice, *O. rufipogon*, is known to contain alleles that confer valuable transgres-sive variation when introgressed into elite cultivars of *O. sativa*, but at this time there is no predictive model that will tell us, *a priori*, where to look for the valuable wild alleles. The challenge before us is therefore to integrate

information from both whole genome SNP assays and targeted gene-mapping studies as a step toward more efficient utilization of wild relatives for rice improvement. New technologies can now be used to resequence entire genomes and to define regions that are highly divergent between gene pools or regions that are shared. This will allow us to identify genomic segments that are common by descent in both *indica* and *japonica* but divergent in *O. rufipogon* (and may correspond to domestication loci), as well as regions of admixture between populations. Chromosomal regions of interest can be introgressed into a suite of elite cultivars to determine whether estimates of divergence are predictive of positive transgressive variation following hybridization. As we gain knowledge about the genes, functional nucleotide polymorphisms and pathways underlying positive transgressive variation, we will be able to make and test predictions about how specific genes or alleles will interact with each other in a given genetic backgound. We will also be able to examine the relationship between SNP frequencies, genome wide patterns of diversity and patterns of linkage disequilibrium to make and test predictions about which of the many wild or exotic accessions combine best with specific elite materials. This information will lay the foundation for 'reverse genetics' models that allow us to more efficiently utilize the wealth of natural variation that resides on the other side of the domestication bottleneck.

Acknowledgments

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. International Rice Genome Sequencing Project: **The map-based sequence of the rice genome**. *Nature* 2005, **436**:793-800.

2. Chang TT: **Rice**. In *Evolution of Crop Plants,* edn 2. Edited by Smartt J, Simmonds NW:Longman Scientific and Technical; 1995

3. Tanksley SD, McCouch SR: **Seed banks and molecular maps: unlocking genetic potential from the wild**. *Science* 1997, **277**:1063-1066.

4. Ladinsky G: **Founder effect in crop–plant evolution**. *Econ Bot* 1985, **39**:191-198.

5. McCouch SR: **Diversifying selection in plant breeding**. *PLoS Biol* 2004, **2**:e347.

6. Zhu Q, Zheng X, Luo J, Gaut B, Ge S: **Multilocus analysis of nucleotide variation of Oryza sativa and its wild relatives: severe bottleneck during the domestication of rice**. *Mol Biol Evol* 2007, **24**:875-888.

7. Caicedo AL, Williamson SH, Hernandez RD, Boyko A,
•   Fledel-Alon A, York TL, Polato NR, Olsen KM, Nielsen R, McCouch SR et al.: **Genome-wide patterns of nucleotide polymorphism in domesticated rice**. *PLoS Genet* 2007, **3**:e163.

This study examined SNP markers developed from 111 randomly sequenced regions of the genome (totaling approximately 54 kb) on 72 *O. sativa* and 21 *O. rufipogon* accessions. Consistent with previous reports, they identified five subpopulations within *O. sativa* and two populations within *O. rufipogon.* They found a genome-wide excess of high frequency derived SNPs, a pattern that has not been found in other domesticated crop species. Several demographic models were developed to explain these patterns of polymorphism. Their results support a bottleneck model that incorporates selective sweeps, or alternatively, a model that includes subdivision and gene flow.

8. Oka HI: In   *Origin of Cultivated Rice.* Edited by Oka HI. Tokyo:
•   Elsevier Science/Japan Scientific Societies Press; 1988.

This book provides an overview of basic rice biology, genetics, and evolution and offers an excellent starting point for a wide range of plants scientists. The subject is approached from many angles, including ethnobotanical, agronomic, genetic, cytogenetic, physiological, molecular, etc. as it focuses on *O. sativa*, *O. glaberrima* and their respective wild ancestors.

9. Tang J, Xia H, Cao M, Zhang X, Zeng W, Hu S, Tong W, Wang J, Wang J, Yu J et al.: **A comparison of rice chloroplast genomes**. *Plant Physiol* 2004, **135**:314-420.

10. Tian X, Zheng J, Hu S, Yu J: **The rice mitochondrial genomes and their variations**. *Plant Physiol* 2006, **140**:401-410.

11. Zhang H, Sun J, Wang M, Liao D, Zeng Y, Shen S, Yu P, Mu P, Wang X, Li Z: **Genetic structure and phylogeography of rice landraces in Yunnan, China, revealed by SSR**. *Genome* 2006, **50**:72-83.

12. Chen W-B, Nakamura I, Sato Y-I, Nakai H: **Distribution of deletion type in cpDNA of cultivated and wild rice**. *Jpn J Genet* 1993, **68**:597-603.

13. Lu BR, Zheng KL, Qian HR, Zhuang JY: **Genetic differentiation of wild relatives of rice as assessed by RFLP analysis**. *Theor Appl Genet* 2002, **106**:101-106.

14. Cheng CY, Motohashi R, Tsuchimoto S, Fukuta Y, Ohtsubo H, Ohtsubo E: **Polyphyletic origin of cultivated rice: Based on the interspersion pattern of SINEs**. *Mol Biol Evol* 2003, **20**:67-75.

15. Londo J, Chiang Y, Hung K, Chiang T, Schaal B: **Phylogeography**
•   **of Asian wild rice, Oryza rufipogon reveals multiple independent domestications of cultivated rice Oryza sativa**. *Proc Natl Acad Sci* 2006, **103**:9578-9583.
Using a phylogeographic approach, this study examines a total of 2.2 kb of sequence variation across two nuclear and one chloroplast gene region under different levels of selection in 203 *O. sativa* cultivars and 161 *O. rufipogon* accessions to investigate the number and origin of rice domestication events. Their results support at least two independent domestications of *O. sativa* from geographically distinct *O. rufipogon* populations, corresponding to the *indica* and *japonica* varietal groups.

16. Hu H, Mu J, Zhange H, Tao Y, Han B: **Differentiation of a miniature inverted transposable element (MITE) system in Asian rice cultivars and its inference for a diphyletic origin of two sub-species of Asian cultivated rice**. *J Integ Pl Biol* 2006, **48**:260-267.

17. Rakshit S, Rakshit A, Masumura H, Takahashi Y, Hasegawa Y,
•   Ito A, Ishii T, Miyashita N, Terauchi R: **Large-scale DNA polymorphism study of Oryza sativa and Oryza rufipogon reveals the origin and divergence of Asian rice**. *Theor Appl Genet* 2007, **114**:731-743.
A polymorphism survey across ∼26 kb of sequence in 30 diverse *O. sativa* and *O. rufipogon* accessions was used to perform a phylogenetic analysis. The grouping of *O. rufipogon* accessions as sister groups to *O. sativa* accessions lends further support to the hypothesis of multiple rice domestications. The nucleotide diversity of *indica* genotypes was an order of magnitude greater than their *japonica* counterparts and *japonica* had a larger number of derived and fixed polymorphisms, which is likely because of a much larger effective population size in *indica* than in *japonica*. Linkage disequilibrium was found to decay within 5 kb in *O. rufipogon*, but extended to over 50 kb in *indica*.

18. Ma J, Bennetzen J: **Rapid recent growth and divergence of rice nuclear genomes**. *Proc Natl Acad Sci* 2004, **101**:12404-12410.

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

20. 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.
The investigation of nucleotide diversity data from 10 unlinked nuclear loci (totaling approximately 8 kb) in 30 *O. sativa* and 30 *O. rufipogon/O. nivara* accessions reveals that *O. rufipogon* possesses similar levels of nuclear variation to wild relatives of other crops, but cultivated rice retains only 10–20% of the diversity of its wild ancestor. The *indica* varietal group was found to possess considerably more nucleotide diversity than *japonica*. While linkage disequilibrium decayed at around 400 bp in *O. rufipogon*, it extended across the entire sequenced region in their *O. sativa* samples. The data from this study also supports *O. rufipogon* and *O. nivara* as two ecotypes of the same species.

21. Tang T, Lu J, Huang J, He J, McCouch SR, Shen Y, Kai Z, Purugganan MD, Shi S, Wu C: **Genomic variation in rice: Genesis of highly polymorphic linkage blocks during domestication**. *PLoS Genetics* 2006, **2(11)**:e199.

22. Glaszmann JC: **Isozymes and classification of Asian rice varieties**. *Theor Appl Genet* 1987, **74**:21-30.

23. Garris AJ, Tai TH, Coburn JR, Kresovich S, McCouch S: **Genetic**
•• **structure and diversity in Oryza sativa L.**. *Genetics* 2005, **169**:1631-1638.
Using 169 nuclear SSR and two chloroplast markers on a set of 234 diverse *O. sativa* landrace varieties, five distinct subpopulations were revealed with an average FST value of 0.37. The indica and aus sub-populations were more closely related to the indica varietal group while the tropical japonica, temperate japonica, and aromatic subpopu-lations were part of the japonica varietal group. These subpopulations correspond well to those first described by Glaszmann (1987) using 17 isozyme loci and have been confirmed by Caicedo et al. (2007) using SNP analysis.

24. Yamanaka S, Nakamura I, Watanabe KN, Sato YI: **Identification of SNPs in the waxy gene among glutinous rice cultivars and their evolutionary significance during the domestication process of rice**. *Theor Appl Genet* 2004, **108**:1200-1204.

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

26. Sweeney M, Thomson M, 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(8)**:e133.

27. Kovach MJ, Sweeney M, McCouch SR: **New insights into**
• **the history of rice domestication**. *Trends Genet* 2007, **23(11)**:578-587.
This paper reviews the evolutionary histories of cloned domestication-related genes in rice and concludes that several key domestication alleles arose and were fixed in the *japonica* gene pool and subsequently spread into *indica* and throughout the rice world. This pattern suggests there was substantial gene flow between divergent subpopulations during the early domestication process, resulting in a common set of domestication alleles across *O. sativa*. How gene flow can be reconciled with the maintenance of several deeply differentiated subpopulations in *O. sativa* is also discussed.

28. Sang T, Ge S: **The puzzle of rice domestication**. *J Integ Pl Biol*
• 2007, **49**:760-768.
The *sh4* allele in rice is responsible for a major reduction in grain shattering, and this allele is fixed across all of *O. sativa*. In order to reconcile the data supporting multiple rice domestications with the presence of this common domestication gene, this review proposes both a 'snow-balling model' and a 'combination model' to describe the process of rice domestication. In the 'snow-balling model', key domes-tication alleles were fixed in a small *O. rufipogon* population, from which an early cultivar would have been selected and subsequently hybridized with local rice populations, giving rise to both *indica* and *japonica* genotypes fixed for the domestication alleles. The 'combination model' allows for divergent wild rice populations being slowly domesticated, with gene flow between early domesticates resulting in a common set of domestication alleles in both *indica* and *japonica*.

29. Harushima Y, Nakagahara M, Yano M, Sasaki T, Kurata N: **Diverse variation of reproductive barriers in three intraspecific rice crosses**. *Genetics* 2002, **160**:313-322.

30. Sano Y: **Constraints in using wild relatives in breeding: lack of basic knowledge on crop gene pools**. In *International Crop Science*, vol. I. Edited by Buxton DR: Crop Science Society of America; 1993:437-443.

31. Ni J, Colowit PM, Mackill DJ: **Evaluation of genetic diversity in rice subspecies using microsatellite markers**. *Crop Sci* 2002, **42**:601-607.

32. Lu H, Redus MA, Coburn JR, Rutger JN, McCouch SR, Tai TH: **Population structure and breeding patterns of 145 US rice cultivars based on SSR marker analysis**. *Crop Sci* 2004, **45**:66-76.

33. Allard RW: **Principles of Plant Breeding**. New York: John Wiley and Sons Inc.; 1960.

34. Li J, Yuan L: **Hybrid rice: genetics, breeding, and seed production**. *Pl Breed Rev* 2000, **17**:150-158.

35. Cheng S-H, Zhuang J-Y, Fan Y-Y, Du J-H, Cao L-Y: **Progress in research and development on hybrid rice: a superdomesticate in China**. *Ann Bot* 2007, 1-8, doi:10.1093/aob/mcm121.

36. McCouch S, Sweeney M, Li J, Jiang H, Thomson M,
•• Septiningsih E, Edwards J, Moncada P, Xiao J, Garris A *et al.*: **Through the genetic bottleneck: O. rufipogon as a source of trait-enhancing alleles for O. sativa** *Euphytica* 2007, **154**:317-339.
Several parallel advanced backcross QTL studies over the past decade were performed using a single accession of *O. rufipogon* (IRGC 105491) as the donor parent in combination with elite cultivars from around the world as recurrent parents. This paper summarizes the results of this research, including the identification of QTL and cloning of genes under-lying several 'wild' QTL. These studies provide empirical evidence demonstrating the value of wild germplasm for improving yield and other agronomic traits in elite rice cultivars despite the inferior performance of the wild accessions. The authors also provide insights about strategies for utilizing wild germplasm in a more systematic manner.

37. Li ZK, Fu BY, Gao YM, Xu JL, Ali J, Lafitte HR, Jiang YZ, Rey JD, Vijayakumar CHM, Maghirang R *et al.*: **Genome-wide introgression lines and their use in genetic and molecular dissection of complex phenotypes in rice (Oryza sativa L.)** *Plant Mol Biol* 2005, **59**:33-52.

38. Tian F, Li D, Fu Q, Zhu Z, Fu Y, Wang X, Sun C: **Construction of**
• **introgression lines carrying wild rice (Oryza rufipogon Griff.) segments in cultivated rice (Oryza sativa L.) background and characterization of introgressed segments associated with yield-related traits** *Theor Appl Genet* 2006, **112**:570-580.
A set of 159 introgression lines (ILs) was developed between the recurrent parent Guichao 2 (*indica*) and a japonica-like *O. rufipogon* accession from eastern China. The 159 ILs captured 67.5% of the *O. rufipogon* genome, with only 2.2% donor genome per line. Three yield-related traits were measured at two sites, and 35% (6/17) QTL had positive alleles coming from the *O. rufipogon* parent.

39. Brar DS, Khush GS: **Alien introgression in rice**. *Pl Mol Biol* 1997, **35**:35-47.

40. Li J, Thomson MJ, McCouch SR: **Fine mapping of a grain-weight quantitative trait locus in the pericentromeric region of rice chromosome 3**. *Genetics* 2004, **168**:2187-2195.

41. Xie X, Song MH, Jin F, Suh JP, McCouch SR, Ahn SN: **Fine mapping of a grain weight quantitative trait locus on rice chromosome 8 using near isogenic lines derived from a cross between Oryza sativa and O. rufipogon** *Genetics* 2006, **113**:885-894.

42. Xiao JH, Grandillo S, Ahn SN, McCouch SR, Tanksley SD, Li JM, Yuan LP: **Genes from wild rice improve yield**. *Nature* 1996, **384**:223-224.

43. Xiao JH, Li JM, Grandillo S, Ahn SN, Yuan LP, Tanksley SD, McCouch SR: **Identification of trait-improving quantitative trait loci alleles from a wild rice relative Oryza rufipogon**. *Genetics* 1998, **150**:899-909.

44. Moncada M, Martínez C, Tohme J, Guimaraes E, Chatel M, Borrero J, Gauch H, McCouch S: **Quantitative trait loci for yield and yield components in an Oryza sativa × Oryza rufipogon BC2F2 population evaluated in an upland environment**. *Theor Appl Genet* 2001, **102**:41-52.

45. Li D, Sun C, Fu Y, Li C, Zhu Z, Chen L, Cai H, Wang X: **Identification and mapping of genes for improving yield from Chinese common wild rice (O. rufipogon Griff.) using advanced backcross QTL analysis** *Chin Sci Bull* 2002, **18**:1533-1537.

46. Septiningsih EM, Prasetiyono J, Lubis E, Tai TH, Tjubaryat T, Moeljopawiro S, McCouch SR: **Identification of quantitative trait loci for yield and yield components in an advanced backcross population derived from the Oryza sativa variety IR64 and the wild relative O. rufipogon** *Theor Appl Genet* 2003, **107**:1419-1432.

47. Thomson MJ, Tai TH, McClung AM, Hinga ME, Lobos KB, Xu Y, Martinez C, McCouch SR: **Mapping quantitative trait loci for yield, yield components, and morphological traits in an advanced backcross population between Oryza rufipogon and the Oryza sativa cultivar Jefferson**. *Theor Appl Genet* 2003, **107**:479-493.

48. Marri PR, Sarla N, Reddy LV, Siddiq EA: **Identification and**
• **mapping of yield and yield related QTLs from an Indian accession of Oryza rufipogon**. *BMC Genet UK* 2005, **6**:33 doi: 10.1186/1471-2156-1186-1133.
An Indian accession of *O. rufipogon* was used in an advanced backcross QTL analysis to identify QTL related to yield and yield components. Thirty-nine QTL were found for the 13 agronomic traits measured, with 74% of the *O. rufipogon* alleles having a positive effect in the *indica* genetic background.

49. He G, Luo X, Tian F, Li K, Zhu Z, Su W, Qian X, Fu Y, Wang X, Sun C
• *et al.*: **Haplotype variation in structure and expression of a gene cluster associated with a quantitative trait locus for improved yield in rice**. *Genome Res* 2006, **16**:618-626.
Near isogenic lines were used to fine map a previously reported QTL for yield in rice to a ~100 kb region on chromosome 2. A leucine-rich repeat kinase gene cluster was found to have extensive allelic variation in both structure and expression, suggesting this cluster is a good candidate for the yield QTL. Interestingly, the recurrent parent for this study was an *indica* accession while the *O. rufipogon* accession contributing the positive effect on yield was a *japonica*-like sample collected from eastern China.

50. Tian F, Zhu Z, Zhang B, Tan L, Fu Y, Wang X, Sun C: **Fine mapping of a quantitative trait locus for grain number per panicle from wild rice (Oryza rufipogon Griff.)** *Theor Appl Genet* 2006, **113**:619-629.

51. Septiningsih EM, Trijatmiko KR, Moeljopawiro S, McCouch SR: **Identification of quantitative trait loci for grain quality in an advanced backcross population derived from the Oryza sativa variety IR64 and the wild relative O. rufipogon** *Theor Appl Genet* 2003, **107**:1433-1441.

52. Liu F, Sun C, Tan L, Fu Y, Li D, Wang X: **Identification and mapping of quantitative trait loci controlling cold-tolerance of Chinese common wild rice (O. rufipogon Griff.) at booting to flowering stages** *Chin Sci Bull* 2003, **48**:2068-2071.

53. Nguyen BD, Brar DS, Bui BC, Nguyen TV, Pham LN, Nguyen HT: **Identification and mapping of the QTL for aluminum tolerance introgressed from the new source, Oryza rufipogon Griff., into indica rice (Oryza sativa L.)** *Theor Appl Genet* 2003, **106**:583-593.

54. Thomson MJ, Edwards JD, Septiningsih EM, Harrington S, McCouch SR: **Substitution mapping of dth1.1,a flowering time QTL associated with transgressive variation in rice, reveals a cluster of QTLs** *Genetics* 2006, **172**:2501-2514.

55. Aluko G, Martinez C, Tohme J, Castano C, Bergman C, Oard J: **QTL mapping of grain quality traits from the interspecific cross Oryza sativa × O. glaberrima** *Theor Appl Genet* 2004, **109**:630-639.

56. Li J, Jinhua Xiao, Silvana Grandillo, Longying Jiang, Yizhen Wan, Qiyun Deng, Longping Yuan, Susan McCouch: **QTL detection for rice grain quality traits using an interspecific backcross population derived from cultivated Asian (O. sativa L.) and African (O. glaberrima S.) rice** *Genome* 2004, **47**:697-704.

57. Sarla N, Swamy BP: **Oryza glaberrima: a source for the improvement of Oryza sativa**. *Curr Sci* 2005, **89**:955-963.

58. Brondani C, Rangel N, Brondani V, Ferreira E: **QTL mapping and introgression of yield-related traits from Oryza glumaepatula to cultivated rice (Oryza sativa) using microsatellite markers**. *Theor Appl Genet* 2002, **104**:1192-1203.

59. Rangel PHN, Brondani C, Rangel PN, Brondani RPV, Zimmermann FJP: **Development of rice lines with gene introgression from the wild Oryza glumaepatula by the AB-QTL methodology**. *Crop Breed Appl Biotech* 2005, **5**:10-19.

60. Rick CM: **High soluble-solids content in large-fruited tomato lines derived from a wild green-fruited species**. *Hilgarida* 1974, **42**:493-510.

61. Cox DJ, Frey KJ: **Improving cultivated oats (A vena sativa L.) with alleles for vegetative growth index from A. sterilis L.** *Theor Appl Genet* 1984, **68**:239-245.

62. deVicente MC, Tanksley SD: **QTL analysis of transgressive segregation in an interspecific tomato cross**. *Genetics* 1993, **134**:585-596.

63. Eshed Y, Zamir D: **An introgression line population of Lycopersicon pennellii in the cultivated tomato enables the identification and fine mapping of yield-associated QTL**. *Genetics* 1995, **141**:1147-1162.

64. Rieseberg LH, Archer MA, Wayne RK: **Transgressive segregation, adaptation and speciation**. *Heredity* 1999, **83**:363-372.

65. Rao GU, Ben Chaim A, Borovsky Y, Paran I: **Mapping of yield-related QTLs in pepper in an interspecific cross of Capisicum annum and C. frutescens** *Theor Appl Genet* 2003, **106**:1457-1466.

66. Lacape JM, Nguyen TB, Courtois B, Belot JL, Giband M, Gourlot JP, Gawryziak G, Roques S, Hau B: **QTL analysis of cotton fiber quality using multiple Gossypium hirsutum × Gossypium barbadense backcross generations**. *Crop Sci* 2005, **45**:123-140.

67. Liu B, Fujita T, Yan Z-H, Sakamoto S, Xu D, Abe J: **QTL mapping of domestication-related traits in soybean (Glycine max)**. *Ann Bot* 2007:1-12 doi: 10.1093/annbot/mcm149.

68. Duvick DN, Cassman KG: **Post-green revolution trends in yield potential of temperate maize in the North-Central United States**. *Crop Sci* 1999, **39**:1622-1630.

69. Duvick DN: **Feeding people and protecting natural resources**. In *The Genetics and Exploitation of Heterosis in Crops.* Edited by Coors JG, Pandey S: Crop Science Society of America and Soil Science Society of America; 1999:19-29.

70. Duvick DN: **The contribution of breeding to yield advances in maize (Zea mays L.)** *Advan Agron* 2005, **86**:83-145.

71. Virmani SS, Kumar I: **Advances in hybrid rice research and development in the tropics**. In *Proceedings of the 4th International Symposium on Hybrid Rice June 2004, Los Banos (Philippines).* Edited by Virmani SS, Mao CX, Hardy B: *Proceedings of the 4th International Symposium on Hybrid Rice June 2004, Los Banos (Philippines)International Rice Research Institute: 2002:10-20*.

72. Xiao J, Li J, Yuan L, Tanksley SD: **Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers**. *Genetics* 1995, **140**:745-754.

73. Hua JP, Xing YZ, Xu CG, Sun XL, Yu SB, Zhang Q: **Genetic dissection of an elite rice hybrid revealed that heterozygotes are not always advantageous for performance**. *Genetics* 2002, **162**:1885-1895.

74. Sun C, Wang X, Yoshimura A, Doi K: **Genetic differentiation for nuclear mitochondrial and chloroplast genomes in common wild rice (O. rufipogon Griff) and cultivated rice (O. sativa L)** *Theor Appl Genet* 2002, **104**:1335-1345.

75. Ishikawa R, Yamanaka S, Fukuta Y, Chitrakon S, Bounphanousay C, Kanyavong K, Tang L, Nakamura I, Sato T, Sato YI: **Genetic eroison from modern varieties into traditional upland rice cultivars**. *Genet Res Crop Evol* 2006, **53**:245-252.

76. Cooper M, Podlich DW, Smith OS: **Gene-to-phenotype models and complex trait genetics**. *Aus J Ag Res* 2005, **56**:895-918.

77. Ma G, Yuan L: **Hybrid rice achievements and development in China**. In *Proceedings of the 4th International Symposium on Hybrid Rice June 2004, Los Banos (Philippines).* Edited by Virmani SS, Mao CX, Hardy B: *Proceedings of the 4th International Symposium on Hybrid Rice June 2004, Los Banos (Philippines)International Rice Research Institute: 2002:247-257*.