Review TRENDS in Genetics Vol.23 No.11

New insights into the history of rice domestication

Michael J. Kovach, Megan T. Sweeney and Susan R. McCouch

Cornell University, Department of Plant Breeding and Genetics, Ithaca, NY 14853–1901, USA

The history of rice domestication has long been a subject of debate. Recently obtained genetic evidence provides new insights into this complex story. Genome-wide studies of variation demonstrate that the two varietal groups in *Oryza sativa* (*indica* and *japonica*) arose from genetically distinct gene pools within a common wild ancestor, *Oryza rufipogon*, suggesting multiple domestications of *O. sativa*. However, the evolutionary history of recently cloned domestication genes adds another layer of complexity to the domestication of rice. Although some alleles exist only within specific subpopulations, as would be expected if the domestications occurred independently, other major domestication alleles are common to all cultivated *O. sativa* varieties. Our current view of rice domestication supports multiple domestications coupled with limited introgression that transferred key domestication alleles between divergent rice gene pools.

From wild grass to staple crop

More than 10 000 years ago, ancient peoples began to gather and consume *Oryza rufipogon*, an unruly wild grass species that grew in the swamps and marshes throughout tropical and subtropical Asia (Box 1). Through a process of continuous selection for desirable features, these early farmers slowly transformed wild rice into *Oryza sativa*, which is now an essential staple crop for billions of people worldwide. Indeed, rice is now the primary source of food and livelihood for more than a third of the world's population, and is produced on every continent with arable land [1].

Domestication involves a series of profound genetic changes resulting from selection that make a wild species more amenable for cultivation and consumption by humans. It is widely recognized that domestication is not a single 'event', but rather a dynamic evolutionary process that occurs over time and, in some species, continues to this day [2]. The traits that distinguish modern rice varieties from their wild ancestor can range from subtle to dramatic (Figure 1). In addition to traits that resulted in major alterations of plant structure and/or reproductive physiology, humans have selected for characteristics that made rice grains more appealing as a food source, including grain size, shape, color, fragrance and amylose content.

Another feature of domestication traits is that they are generally quantitative in nature, meaning they are under the control of numerous genetic components that often interact in complex biochemical and regulatory pathways. For example, in rice, there are at least five grain shattering loci (see Glossary) [3–5], a minimum of six dormancy

Glossary

Abscission layer: a zone of cells at the interface of plant organs that breaks down during senescence, promoting the organ to fall off.
Accession: a sample of seed collected to represent a species, population or variety.
Admixed: an organism whose genome contains DNA inherited from different subpopulations as the result of outcrossing and recombination.
Apiculus: the tip of the lemma or palea; a small point at the tip of the seed hull.
Dormancy: a physiological period of quiescence during which a mature seed will not germinate.
F_{st}: a measure of population differentiation based on polymorphism data. It compares the genetic variability within and between populations. If two populations have completely divergent DNA sequences, their F_{st} value equals 1, indicating complete differentiation. If two populations have identical DNA sequences, their F_{st} value equals 0, indicating no differentiation. Therefore, when comparing two populations, the larger the F_{st} value (the closer to 1), the more differentiated those populations are from one another.
Glume: small bract found at the base of cereal grains.
Grain shattering: seed abscission; when ripe seed falls from the panicle before harvesting occurs.
Haplotype: a combination of alleles or sequence variants at multiple linked loci that are transmitted together.
Hybridization: the result of successful outcrossing and fertilization of two genetically distinct plants.
Introgression: the movement of a discrete portion of a genome from one genotype into another via meiotic recombination.
Isozymes: variants of an enzyme; proteins that differ in amino-acid sequence, but catalyze the same reaction.
Landrace: a primative variety consisting of a heterogeneous mixture of genotypes selected and maintained by early farmers.
Life history habit: the reproductive cycle of a plant; annual habit refers to plants that live, reproduce and die in one year or season, whereas perennial habit refers to plants whose life cycle spans more than one year.
Molecular clock: using DNA polymorphisms between the sequences of two organisms to deduce the amount of time that has elapsed since they diverged from a common ancestor.
Panicle: a compound raceme; the inflorescence of a rice plant.
Pericarp: the outermost layer of cells in the seed coat.
Photoperiodic response: a plant's ability to flower in response to increasing or decreasing day length.
Phylogenetic: referring to the evolutionary relationships between organisms.
Retrotransposon: a type of transposable element (mobile DNA) that moves through an RNA intermediate inserting new copies of itself throughout the genome of an organism.
RFLP marker: restriction fragment length polymorphism; a DNA marker that detects genetic differences between individuals based on the presence or absence of restriction enzyme sites.
SNP: single nucleotide polymorphism.
SSR marker: simple sequence repeat; a DNA marker that detects genetic differences between individuals based on the length of short repetitive sequences found throughout the genome.
Subpopulation: a population within a species that is genetically divergent from other populations within that species.
Tiller: a branch of the rice plant including roots, culms, leaves and panicle.
Thresh: the process of removing ripe grains from harvested panicles.
Transcription factor: a protein that affects the transcription and thus expression of specific target genes.

Corresponding author: McCouch, S.R. (SRM4@cornell.edu).
Available online 25 October 2007.

 0168-9525/$ – see front matter doi:10.1016/j.tig.2007.08.012

Box 1. The wild ancestor of *O. sativa*

The wild ancestors of our cultivated crop species not only represent valuable sources of genetic diversity for crop improvement, but also provide a window through which we can catch a glimpse of the evolutionary history of our domesticated species. Several lines of evidence suggest that *O. sativa* was domesticated from *O. rufipogon*, also known as Asian common wild rice. Numerous genetic studies using molecular markers or DNA sequence information demonstrate that *O. sativa* cultivars are more closely related to *O. rufipogon* than to any other *Oryza* species, with *O. rufipogon* accessions often grouping together with *O. sativa* in phylogenetic (see Glossary) analyses [19,30]. The geographic ranges of *O. rufipogon* and *O. sativa* overlap throughout Asia, and there are no major reproductive barriers between them, leading to a continuous series of intermediate admixed (see Glossary) genotypes [37]. These findings clearly demonstrate the close relationship between cultivated Asian rice and *O. rufipogon*.

O. rufipogon is a complex species that includes both annual (*Oryza nivara*) and perennial (*O. rufipogon*) forms. Some researchers distinguish between *O. rufipogon* and *O. nivara* due to their variation in life history habit [63,64]. Others consider them to be ecotypes of a single species because they exhibit continuous variation in nature for their life history habit, coupled with the fact that there is no major reproductive barrier or significant genetic differentiation between them [19,31,37,65], as demonstrated by the abundance of shared sequence polymorphisms, lack of fixed differences, and very low F_{st} values [19,42,66]. As such, we will treat these two species as a single large gene pool referred to as *O. rufipogon*.

Figure 1. The domestication transformation: From *O. rufipogon* to *O. sativa*. The wild ancestor of cultivated rice (*O. sativa*) is *O. rufipogon*, a diverse species that exists over broad geographic and ecological regions across Asia. During the domestication of rice, an entire suite of morphological and physiological traits were altered in response to human selection. Compared to its wild *O. rufipogon* ancestor, cultivated rice typically exhibits reduced grain shattering, reduced dormancy, loss of pigmentation in the hull and seed coat and a reduction in the rate of outcrossing. Modern rice varieties also display increased synchronization of tiller (see Glossary) development and panicle formation, more secondary panicle branches, increased grain number and weight and a modified photoperiodic response (see Glossary) [62]. **(a)** Panicle from *O. rufipogon*; **(b)** seeds from *O. rufipogon*; **(c)** panicles from *O. sativa*; and **(d)** seeds from *O. sativa*.

loci(see Glossary) [6,7], as many as 31 loci associated with the change from outcrossing to inbreeding [8,9], more than eight loci associated with grain weight [10] and at least five loci associated with grain color [11–13]. Despite the quantitative nature of domestication traits, there appear to be a few key discrete genetic loci responsible for the major shift from wild to cultivated forms [14]. In addition to these major quantitative trait loci (QTLs), numerous genetic modifiers, or minor QTLs, contribute to the range of phenotypic variation for a given trait. To date, all cloned domestication-related loci in rice have involved reproductive traits (Table 1).

In addition to the suite of traits that were altered during domestication, the evolutionary history of rice is not complete without considering the deep population structure within *O. sativa*. Ancient records from the Han dynasty in China recognized the existence of two distinguishable types of rice, which were referred to as Hsien and Keng [15,16]. We now refer to these genetically distinct varietal groups as *indica* and *japonica*, respectively. The two groups have several morphological and physiological differences despite being grown in overlapping geographical ranges today.

Table 1. Key domestication-related genes cloned in rice

Gene	Chr[a]	Trait affected	Functional mutation	Molecular function	Refs
Sh4	4	Grain shattering	SNP causing an amino-acid substitution	Myb3 transcriptional regulator	[48,49]
qSH1	1	Grain shattering	SNP in regulatory region	BEL1-type homeobox transcriptional regulator	[50]
Rc	7	Grain pericarp color	Deletion causing protein truncation	Basic helix–loop–helix transcriptional regulator	[11]
Waxy	6	'Sticky' (glutinous) grains	SNP in intron affecting mRNA splicing	Granule-bound starch synthase	[52–54]
GS3	3	Grain size/shape	SNP causing protein truncation	Cellular signaling protein with a VWFC module	[55]
[BAD2][b]	8	Grain fragrance/flavor	Deletion causing protein truncation	Betaine aldehyde dehydrogenase	[58]
[Gn1a][b]	1	Grain number	Several possible mutations	Cytokinin oxidase/dehydrogenase	[57]
[GW2][b]	2	Grain weight/width	Deletion causing protein truncation	RING-type protein with E3 ubiquitin ligase activity	[56]

[a]Chr, chromosome.
[b]Bracketed genes affect domestication-related traits, but there is no evidence at this time to determine whether these genes were part of the domestication process or whether they are novel mutations that have recently become the targets of selection.

Although *indica* and *japonica* represent the deepest genetic differentiation within *O. sativa*, five major subpopulations (see Glossary) are widely recognized (*indica, tropical japonica*, *temperate japonica*, *aus* and *aromatic*) [17,18] (Box 2). This population structure adds further complexity to our understanding of the domestication process that led to cultivated rice.

Modern genetics provides the necessary tools to understand the dynamics of the domestication process and to identify the key genes and alleles that were the targets of selection during rice domestication. Although only a limited number of genetic loci affecting key domestication traits in rice have been characterized at the gene level, an examination of the evolutionary history of these domestication genes provides novel insights into the movement of both cultivated rice and human populations across Asia. This review will critically examine the existing hypotheses used to describe the rice domestication process by integrating previous studies of genome-wide variation in rice with recent data on the evolutionary history of key domestication genes. The view that emerges is one that challenges us to rethink old paradigms as we continue to unravel the complex story of rice domestication.

Prevailing hypotheses regarding rice domestication

Several major hypotheses describe the process of rice domestication, each attempting to account for the evolution of the deeply differentiated subpopulations of *O. sativa*. First, several researchers have suggested that the *indica* varietal group was originally domesticated from *O. rufipogon* and the *japonica* group was derived later from *indica* [15,19,20]. The greater genetic diversity found within *indica* compared to *japonica* [17] and the close genetic relationship between *indica* and existing populations of both annual and perennial forms of *O. rufipogon* [21] provided evidence to support this hypothesis. The second theory is that the *indica–japonica* differentiation occurred as a result of adaptation to different ecological and geographical environments following a single domestication of *O. sativa* from *O. rufipogon* [22,23]. A third hypothesis, originally proposed by Chou [16], is that rice was domesticated independently at least twice from a pre-differentiated ancestral *O. rufipogon* gene pool [17,24–28].

An understanding of both the patterns of genome-wide variation in wild and cultivated rice and the number and origins of alleles associated with rice domestication genes

Box 2. Subpopulation structure in *O. sativa*

There are two genetically distinct varietal groups within *O. sativa*, referred to as *indica* and *japonica*. Traditionally, the two groups have been distinguished based on morphological characters, including grain shape, apiculus (see Glossary) hair length, leaf color or through biochemical assays for reaction to phenol and sensitivity to potassium chlorate [37,67]. There are also numerous reproductive barriers between the modern *indica* and *japonica* groups [60]. Yet the range of variation for any one of these phenotypic traits overlaps between the two groups, leading to confusion regarding the classification of particular genotypes [37].

Since publication of the first rice genetic map [68], DNA markers have been widely used to explore the genetic architecture of rice. RFLP markers (see Glossary) readily detect the differentiation between *indica* and *japonica* [69,70]. Isozymes, simple sequence repeat (SSR) markers (see Glossary), and single nucleotide polymorphisms (SNPs) provide a higher resolution of population structure. Using 169 SSR markers on a set of 234 diverse *O. sativa* genotypes, Garris *et al.* identified five subpopulations: *indica, aus, tropical japonica*, *temperate japonica* and *aromatic* [17] (Figure I). The same groups were identified using SNP markers derived from 111 randomly sequenced regions of the genome on a subset of 72 accessions [34]. These studies were both consistent with the original study by Glaszmann [18] using 15 isozyme markers on a larger set of 1700 diverse *O. sativa* genotypes. Of these five subpopulations, *indica* and *aus* belong to the *indica* varietal group whereas *tropical japonica*, *temperate japonica* and *aromatic* are closely related to the *japonica* varietal group. A high degree of differentiation exists between the groups, with pairwise estimates of F_{st} ranging from 0.20 (*tropical* versus *temperate japonica*) to 0.45 (*indica* versus *temperate japonica*) [17].

The results of these studies highlight the importance of utilizing molecular markers to assign genotypes to the appropriate subpopulation for genetic analyses. For example, the Basmati rices (known for their aroma and specialized grain quality) have traditionally been classified as members of the *indica* varietal group because of their long, slender grain shape. However, the Basmati rice genotypes form a genetically distinct subpopulation (*aromatic*) that is more closely related to *japonica* than *indica* (F_{st} for *tropical japonica–aromatic* is 0.23; F_{st} for *indica–aromatic* is 0.39) [17]. Similarly, the *aus* varieties have traditionally been referred to as *indica* for lack of distinguishing morphological differences, but they are genetically divergent from *indica* (F_{st} for *indica–aus* is 0.26) [17].

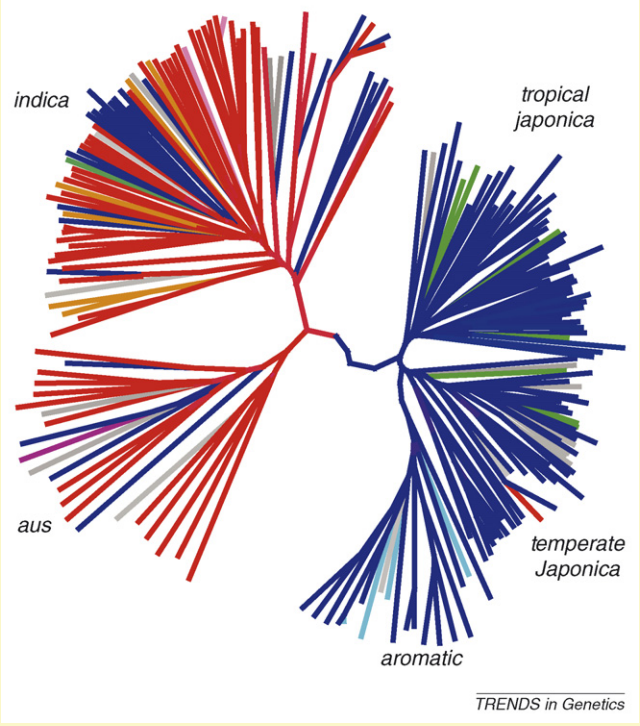

Figure I. Subpopulation structure in *O. sativa*. *O. sativa* is characterized by the presence of deep genetic differentiation. This unrooted phylogenetic tree was constructed from data using 169 nuclear SSR and two chloroplast markers on 234 landraces of *O. sativa* [17]. The branch structure reflects the phylogenetic relationships based on the nuclear SSR markers. The branch color corresponds to the chloroplast haplotype of each accession. This tree clearly illustrates the major division between the two varietal groups (*indica* and *japonica*), which are further subdivided into the five rice subpopulations: *indica, aus, tropical japonica*, *temperate japonica* and *aromatic*. Reproduced and modified, with permission, from Ref. [17].

will enable us to consider the validity of the proposed domestication hypotheses. According to the first hypothesis, if *japonica* were derived from *indica*, we would expect *japonica* genotypes to be more closely related to *indica* than to the *O. rufipogon* ancestor. Additionally, under this hypothesis we would expect to see a common set of domestication alleles in both groups with evidence that the alleles had originated in *indica* and were later introgressed into *japonica*. In the second scenario, if a single domestication from *O. rufipogon* were followed by post-domestication divergence between *indica* and *japonica,* we would expect sequence differences between *indica* and *japonica* to postdate the time of domestication. Again under this theory, a common set of domestication alleles would be shared between the two groups. According to the third hypothesis, if *indica* and *japonica* were domesticated independently from a pre-differentiated *O. rufipogon* ancestor, we would expect sequence differences between the two groups to predate the time of domestication. We would also expect there to be unique sets of *indica*- and *japonica*-specific domestication alleles originating from genetically distinct *O. rufipogon* populations. These group-specific alleles would each contribute to the same vital domestication phenotypes in their respective varietal groups. It is highly unlikely that an identical mutation would have arisen and been selected for in these two gene pools, had they been independently domesticated.

Evolutionary history of genome-wide variation in *O. rufipogon* and *O. sativa*

To address these hypotheses, various studies have examined *O. rufipogon* for evidence of differentiation into *indica* and *japonica* types. Genetic analysis of *O. rufipogon* chloroplasts revealed two distinct haplotypes (see Glossary), which have clear frequency differences between the *indica* and *japonica* groups of *O. sativa* [26,29]. Support for a differentiated *O. rufipogon* gene pool also comes from multiple studies using isozymes (see Glossary), restriction fragment length polymorphisms (RFLPs), single nucleotide polymorphisms (SNPs; see Glossary), transposable elements and the published genomic sequence, all demonstrating that *indica* and *japonica* accessions (see Glossary) are more closely related to certain *O. rufipogon* accessions than to each other [19,24,26,30–34].

The diverse geography and ecology of the broad region that is home to *O. rufipogon* (East, South and Southeast Asia) suggests that this ancestral species was broadly adapted to diverse climates and would have provided ample genetic diversity from which to have domesticated *O. sativa* on numerous occasions. Studies exploring the diversity within *O. rufipogon* have recently reported genetically identifiable population sub-structure with at least some hint of geographic associations [25,26,34,35]. The degree of genetic differentiation between populations of *O. rufipogon* increases with geographical distance, suggesting that geographical isolation played a major role in establishing the predifferentiated gene pools within *O. rufipogon* [35]. Also, the inbreeding habit that accompanied rice domestication and transformed the highly outcrossing *O. rufipogon* (outcrossing rate of 30–50%) [36,37] to the almost exclusively inbreeding *O. sativa* (outcrossing rate ~2%) [37,38] would have further promoted isolation between rice subpopulations.

A further source of evidence supporting multiple domestications comes from recent work estimating the divergence time between *indica* and *japonica*. The availability of complete genomic sequence for both *japonica* and *indica* [39–41] has allowed the estimation of the divergence time between these two groups. Two independent studies used a molecular clock approach (see Glossary) to calculate the *indica–japonica* divergence time to be around 0.4 million years ago (mya) [42,43]. Another study evaluating the patterns of retrotransposon insertion (see Glossary) in the two varietal groups estimated the divergence time to be at least 0.2 mya [44]. This discrepancy may be due to the fact that the level of divergence varies across the genome [45]. In any case, these molecular clock estimates are several orders of magnitude earlier than the first archaeological evidence for rice consumption by humans (Box 3), providing strong evidence that the divergence of the two cultivated rice gene pools predated rice domestication. Taken together, these studies contradict the first two hypotheses and support multiple domestications of *O. sativa* from a predifferentiated *O. rufipogon* ancestor.

Evolutionary history of cloned rice domestication genes

To determine whether the history of rice domestication alleles is consistent with multiple domestications, we will now examine several examples of cloned rice domestication genes whose evolutionary histories have recently been explored. If *indica* and *japonica* were independently domesticated as the patterns of genome-wide variation suggest, we would expect to see different domestication alleles present in the different varietal groups.

Grain shattering: Sh4

One of the most universal changes during the domestication of the cereal crops was a reduction in grain shattering. This change was essential for humans to efficiently harvest the crop, but rendered the plant almost entirely dependent on humans for dispersal [46]. In a cross between an *indica* cultivar and an annual form of *O. rufipogon*, Li *et al.* identified a major QTL, *shattering4* (*sh4*), which affects grain shattering [47]. The *Sh4* gene was subsequently cloned [48,49] and, although its exact function remains unknown, SH4 is thought to be a transcription factor (see Glossary) involved in cell wall degradation and/or establishment of the abscission layer (see Glossary) that releases the grain from the panicle (see Glossary) [48,49]. The functional nucleotide polymorphism in the gene was identified as an SNP causing a single amino acid change within the predicted DNA-binding domain of SH4.

The *sh4* mutation is found in all five *O. sativa* subpopulations, whereas the dominant *Sh4* allele exists in all wild, shattering accessions of *O. rufipogon* along with six other closely related species of *Oryza* [47]. Lin *et al.* recently examined *sh4* (which they refer to as *sha1*) in 96 *indica* and 112 *japonica* cultivars and found that they all contained the same functional SNP associated with the domesticated allele [49]. Additionally, *sh4* cDNA sequencing

Box 3. Effects of geography and climate on rice diversity

As early as the late 19th Century, De Candolle suggested that rice cultivation began in China, where the oldest historical records could be found [71]. By contrast, Vavilov designated India as the location of rice domestication [72]. Since then, countless studies have generally agreed that the birthplace of rice domestication lies within a broad arc extending from eastern China, through the foothills of the Himalayas in Vietnam, Thailand and Myanmar, to eastern India [20,37]. More recently, archaeological studies have strongly supported the Middle and lower Yangzte River Valley in China as the site where rice was first domesticated (Figure II). Some hypothesize that the rice remains discovered at these sites in China represent foraged wild forms instead of truly domesticated forms of rice [46]. Yet, there is mounting evidence to the contrary. A new study investigating an archaeological site in Lower Yangtze region found evidence of paddy rice cultivation, suggestion that the inhabitants of this region had already formed agricultural societies around rice by 7,700 years ago [73] Also, analysis of phytolith samples (silicon fossils of rice plant cells) from the Diaotonghuan site in Northern Jiangxi province (China) dated the earliest known samples of domesticated rice, based on a morphological indicator [double-peaked glume (see Glossary) cell phytoliths], between 9000–10 000 years ago [74].

Despite this extensive archaeological evidence supporting the Yangtze River Valley in China as a region where rice was first domesticated, the wild *O. rufipogon* populations in this region today exhibit very low genetic diversity. An examination of twelve wild-rice populations in China demonstrated that the lowest genetic diversity existed in the northern-most populations near the Yangtze River [35]. It is also well known that the *japonica* varietal group, believed to have undergone domestication in this region, has dramatically less genetic diversity than the *indica* group [17]. The reasons for the lower genetic diversity of *japonica* relative to *indica* can be explained by reviewing the history of the region. A major short-term climate change during a period known as the Younger Dryas resulted in a return to glacial-like conditions across Northern Asia from 11 500 to 13 000 years ago [75,76]. The colder climate would have eliminated a large portion of the *japonica*-like wild ancestors in the Yangtze River Valley. As a result, humans would have been forced to rely on a narrowing gene pool, motivating a more rapid movement toward domestication of rice in this region. By contrast, the warmer tropics of South and Southeast Asia would have maintained larger, more diverse populations of *indica*-like *O. rufipogon* that would have been foraged by humans for a longer time, resulting in a more gradual domestication process. Under this scenario, the domestication bottleneck would have been more abrupt and severe in the *japonica* gene pool than in the *indica* gene pool, consistent with the available genetic data [17].

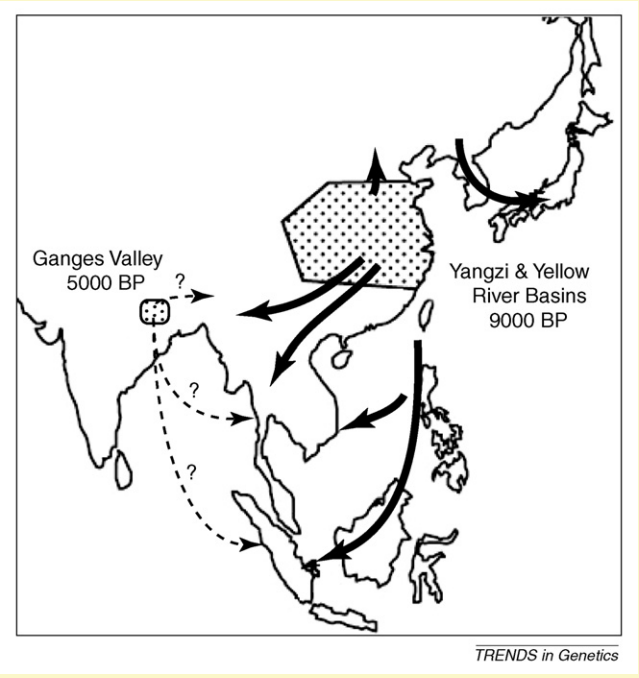

Figure II. The origin and dispersal of domesticated rice. The majority of archaeological evidence points to the Yangtze River Valley of China as the birthplace of rice cultivation by humans. This region is where the oldest archaeological remains of rice have been uncovered. Radiocarbon data from fossilized rice remains from over 100 sites along the length of the Yangtze River were analyzed with the oldest samples coming from the Middle Yangtze in Hubei and Hunan provinces, dating to around 11 500 years ago [77]. Other studies analyzing radiocarbon data from fossilized rice remains in this region report dates ranging from 8000 to 13 900 years ago [74,78]. The early rice domesticates are then thought to have moved north into Korea and Japan, and south and west into Southeast Asia [79,80]. Fuller postulates the Ganges Valley of India as a site of independent rice cultivation and, potentially, domestication [81]. The dates for the earliest rice cultivation in this region, based on archaeological evidence, are not as ancient as those in the Yangtze Valley of China. This region might represent the site of early cultivated forms of *indica*-like rice. The dotted boxes represent potential regions of rice domestication, the bold arrows denote paths of domesticated rice dispersal strongly supported by archaeological and linguistic evidence, and the dotted arrows with question marks denote other potential paths of domesticated rice dispersal. Abbreviation: BP, before present. Reproduced and modified from [79,80] with permission from Peter Bellwood.

revealed an identical allele in four *indica* and four *japonica* accessions. The authors suggest that these results indicate the non-shattering *sh4* allele arose and was selected before the differentiation of *indica* and *japonica* from a common *O. rufipogon* ancestor. Although this is a logical conclusion, an alternative hypothesis is that the domesticated *sh4* allele arose and was selected during the domestication of one of the *O. sativa* subpopulations and was subsequently introgressed into all rice genotypes. To clarify the origin and routes of global dispersal of the non-shattering *sh4* allele, it will be necessary to examine the haplotype structure around the *Sh4* gene in diverse populations of *O. rufipogon* and *O. sativa*.

Grain shattering: qSH1

In addition to the *Sh4* gene, a second major shattering QTL, *qSH1*, was recently identified from an intra-specific cross between *japonica* (cultivar Nipponbare) and *aus* (cultivar Kasalath) [50]. The functional mutation was a regulatory SNP 12 kilobases upstream of a BEL1-type homeobox transcription factor, which decreased expression of the transcription factor only at the provisional abscission layer, resulting in reduced shattering.

By examining the haplotypes around the *qSH1* gene in 118 rice lines including five *O. rufipogon* accessions, Konishi *et al.* deduced the subpopulation origin of this allele [50]. Two clear haplotype groups were found, corresponding to the *indica* and *japonica* varietal groups. The *japonica* haplotypes carrying the non-shattering allele were most closely related to shattering *O. rufipogon* accession W1943, which originated in China and grouped closely with *japonica* cultivars from a previous study [32]. This led to the conclusion that the causative mutation in the domesticated *qsh1* allele occurred in early domesticates of the *japonica* varietal group [50]. The non-shattering allele was prevalent in the *temperate japonica* subpopulation, but was absent in accessions from *tropical japonica* and *indica*, suggesting that this allele was not widely disseminated following its fixation within a portion of the *temperate japonica* gene pool. This distribution pattern is also

consistent with the finding that *temperate japonica* varieties tend to be more difficult to thresh (see Glossary) than varieties from other subpopulations, reflecting the presence of a non-shattering allele in addition to *sh4*.

Pericarp color: Rc

All wild *Oryza* species and some of the early *O. sativa* landraces (see Glossary) exhibit red pigmentation of the seed coat. However, virtually all modern rice cultivars lack red pigmentation and appear white or beige. Classical genetic analysis identified a single locus in rice, *Rc,* that conditioned the change from red to white pericarp (see Glossary) [13]. The *Rc* gene was recently cloned and shown to encode a basic helix–loop–helix (bHLH) transcription factor [11]. The *rc* allele differs from the wild-type allele by a 14-base pair (bp) deletion that results in truncation of the protein before the bHLH domain and causes the change from red to white pericarp [11,12].

To trace the origin and distribution of the *rc* allele, Sweeney *et al.* [51] evaluated 440 geographically and genetically diverse rice cultivars, representing landraces and modern varieties from all five of the subpopulations defined by Garris *et al.* [17] (Box 2). They determined that 98% (330/337) of rice accessions with white pericarp contained the 14-bp deletion, whereas the mutation was not found in any of the landrace varieties with red pericarp (n = 103). To determine the subpopulation origin of the 14-bp mutation, Sweeney *et al.* [51] used diverse red-grained varieties to identify polymorphisms in a 6.5-kilobase (kb) region around the *Rc* gene whose alleles were specific to either red *indica* or red *japonica*. They observed clear haplotype differentiation between *indica* and *japonica* across the region, making it possible to trace the origin of the *rc* allele conferring white pericarp. The white *rc* haplotype differed from the ancestral *japonica* haplotype by only the 14-bp deletion (Figure 2). This analysis suggests that the mutation associated with the change from red to white pericarp in 98% of the *O. sativa* varieties surveyed originated in a *japonica* ancestor [51]. This study presents clear genetic evidence for a domestication allele arising in the *japonica* varietal group and subsequently spreading across genetic, geographic and ethnic boundaries to become fixed in all cultivated rice.

Pericarp color: Rc-s

Whereas 98% of surveyed rice accessions with white pericarp had the same 14-bp deletion, two *aromatic* and five *aus* varieties had white pericarp but did not contain the *rc* allele. These seven varieties were sequenced to look for alternative mutations in the *Rc* gene that could explain the lack of pigment [51]. A comparison of the sequences from these seven varieties to nine control varieties with red pericarp (from all five subpopulations) identified a single SNP in close proximity to the where the 14-bp deletion was in all the other white accessions. This second mutation (*Rc-s*) introduces a premature stop codon that truncates the protein before the bHLH domain, similar to the *rc* mutation [51]. Interestingly, none of the plants contained a complete knockout of this gene, suggesting that the protein is critical to plant function independently of its ability to promote red grain color. All seven individuals

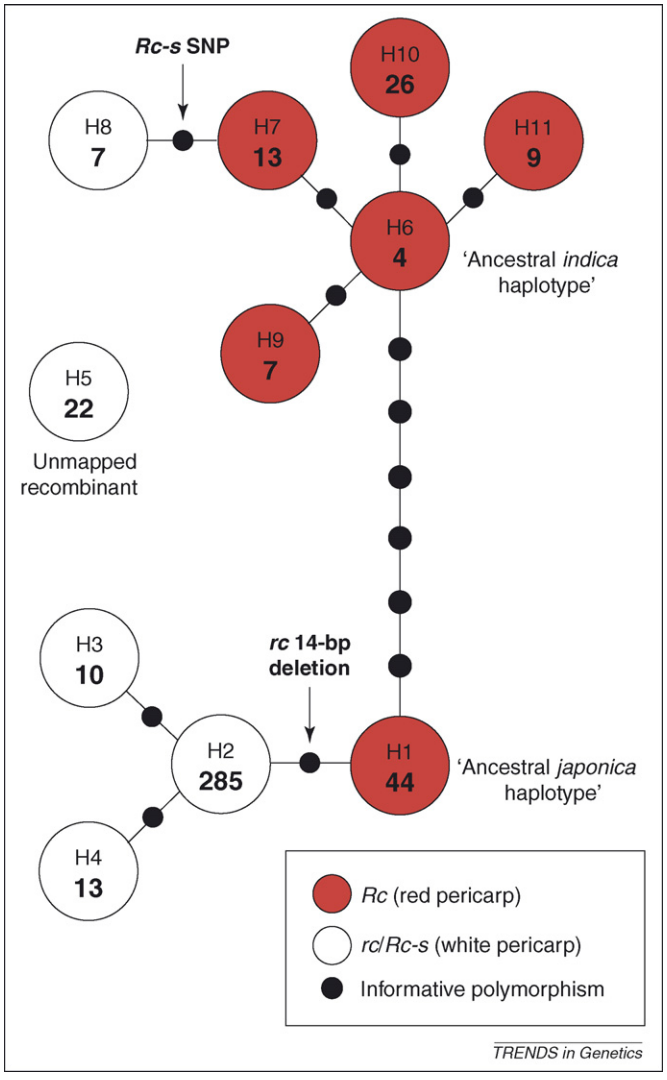

Figure 2. Haplotype network for the *Rc* gene (based on data from Ref. [51]). The *Rc* gene encodes a bHLH transcription factor that regulates the accumulation of red pigments in the rice pericarp. Whereas genotypes with a functional *Rc* have red pericarp, genotypes with either a 14-bp deletion (creating the *rc* allele) or a SNP (creating the *Rc-s* allele) have white pericarp [11]. To determine the origins and dispersal of these mutations in *Rc,* the sequences of 440 diverse rice cultivars were evaluated at this locus [51]. This figure uses the informative polymorphisms found around *Rc* to create a minimum spanning haplotype network. The open circles represent haplotypes H1–H11 and the numbers indicate the number of accessions in that haplotype. Circles with red shading indicate haplotypes with a functional *Rc* allele (red pericarp) whereas white shading signifies a nonfunctional *rc* or *Rc-s* allele (white pericarp). The two mutations in *Rc* resulting in white pericarp are denoted by arrows. Closed black circles represent the number of informative polymorphisms between haplotypes. The ancestral *japonica* (H1) and *indica* (H6) haplotypes are indicated. Haplotype H5 is an intragenic recombinant between two mapped haplotypes and, therefore, was not included in the network. This network illustrates that 98% of all white genotypes surveyed were derived from the ancestral *japonica* haplotype [51].

that contained the *Rc-s* allele had an identical haplotype, which differed from the ancestral (red pericarp) *aus* haplotype only at the functional SNP. Thus, it was concluded that the *Rc-s* allele, which was only found in *aus* and a small percentage of *aromatic* accessions, had its origins in the *aus* subpopulation [51]. This clearly demonstrates that different mutations in the same gene can arise in different subpopulations over the course of domestication. Both *rc* and *Rc-s* apparently became the targets of selection and are virtually indistinguishable at the phenotypic level. It is interesting to note that, unlike the *rc* and *sh4* alleles, but

similar to the non-shattering *qSH1* allele, the *Rc-s* allele was not widely disseminated during the domestication of *O. sativa*.

Amylose content: Wx

Glutinous or 'sticky' rice is used as a staple food in Southeast Asia and has long-standing culinary and cultural importance in China, Japan and Korea. *Japonica* rices are typically valued for their low amylose content, leading to a soft, sticky grain texture ideal for eating with chopsticks, whereas *indica* rices have a higher amylose content leading to firm grains that flake apart when cooked. The glutinous phenotype is the result of a defect in the *Waxy* (*Wx*) gene, which encodes a granule-bound starch synthase responsible for amylose biosynthesis in the grain [52]. There are two functional forms of *Wx* (non-glutinous), with a clear divergence between the two varietal groups in rice: Wx^a is predominant in *indica* and Wx^b is predominant in *japonica*. The Wx^a allele was fixed in all *O. rufipogon* accessions examined [53], indicating that the Wx^a allele is the wild type and Wx^b is the derived form, having been selected during the domestication of *japonica*.

Unlike the Wx^a allele, the Wx^b allele contains an SNP at the 5′ splice site of the first intron, causing incomplete post-transcriptional processing and lower amylose production in these genotypes [52,54]. A second mutation in either Wx^a or Wx^b creates the *wx* allele, which is recessive and fully glutinous. Glutinous rice strains can have either Wx^a- or Wx^b-derived *wx* alleles, but the Wx^b-derived *wx* allele is predominant, as it was found in 97% of 353 glutinous samples from both varietal groups collected from diverse regions across Asia [53]. This finding demonstrates that the majority of glutinous rice cultivars from both *indica* and *japonica* carry the *japonica*-derived allele, mirroring the situation described above for the *rc* mutation. Also, similar to the *Rc-s* and *qsh1* alleles, glutinous genotypes derived from the *indica* Wx^a allele were not widely disseminated.

Other domestication-related traits

The examples presented above include only the rice domestication genes that have been cloned and investigated in terms of their evolutionary origins. Several other genes that affect domestication-related traits have been cloned in rice, including *GS3* and *GW2*, which condition grain size and weight [55,56], *Gn1a*, associated with grain number [57], and *BAD2*, which conditions fragrance [58] (Table 1). Further examination of these genes will determine whether they were involved in the rice domestication process or have been the targets of more recent selection.

What we can interpret from the examples presented above is that several key domestication alleles had a single origin and subsequently became fixed across all rice subpopulations. In the case of white pericarp, the *rc* allele arose in the *japonica* gene pool and was then introgressed into the other subpopulations [51]. This story is repeated for the Wx^b-derived glutinous rice mutation (*wx*) that originated in the *japonica* group and was also introgressed and dispersed [53]. In the future, we might find that the *sh4* allele for non-shattering followed a similar evolutionary path. By contrast, several other domestication alleles were not widely disseminated across subpopulations. The *Rc-s* allele for white pericarp is largely restricted to the *aus* subpopulation, with some introgression (see Glossary) into the *aromatic* subpopulation [51]. The *qsh1* allele for non-shattering and the Wx^b non-glutinous allele both remain largely contained within the *temperate japonica* subpopulation [50,53].

Current observations regarding the history of rice domestication

The examples of rice domestication genes, coupled with the studies involving genome-wide variation in both *O. rufipogon* and *O. sativa*, cannot be reconciled with any of the three prevailing hypotheses about rice domestication. Available genetic data suggest that there were predifferentiated gene pools within the *O. rufipogon* ancestor and that these gene pools gave rise to the *indica* and *japonica* varietal groups. Yet, the presence of identical domestication alleles in both groups strongly suggests that these domestications were not entirely independent. Gene flow across the distinct gene pools of rice appears to have been crucial to the domestication process. Taken together, the current data supports the recently proposed 'combination model' for rice domestication [59]. In this model, the early *japonica* and *indica* cultivars were domesticated from divergent *O. rufipogon* populations and subsequent introgression of key domestication alleles between these early domesticates resulted in a common set of domestication alleles being fixed in all modern varieties. Although all of the evolutionary data from domestication genes thus far demonstrates gene flow from *japonica* to *indica*, only a limited number of rice domestication alleles have thus far been described. This situation is likely to change as the costs of high-resolution genotyping and resequencing continue to decline, and the development of new statistical approaches improves our ability to identify and characterize the targets of selection in diverse germplasm. Whether future genetic studies involving larger numbers of domestication genes will substantiate this observation of gene flow from *japonica* to *indica* or will identify ubiquitous domestication alleles arising in other subpopulations remains to be determined.

If we accept the 'combination model' for rice domestication, how can we reconcile the maintenance of deep genetic differentiation between rice varietal groups with the movement of domestication genes? A recent study involving SNP data on 20 accessions of *O. rufipogon* and 72 landrace and modern *O. sativa* cultivars demonstrated that *O. rufipogon* has the greatest level of heterozygosity, followed by the landraces, with very low levels observed in modern cultivars. These data are consistent with a gradual change from an outcrossing to an inbreeding habit during the domestication process; a change that coincided with the rise of the deeply differentiated subpopulations of *O. sativa* [34]. Numerous reproductive barriers help maintain this genetic differentiation in modern *O. sativa* [60], but cross-hybridization was likely to have been easier in early landrace varieties. Evidence of *indica* x *japonica* hybridization (see Glossary) comes from the recovery of early domesticates containing *indica*-like organellar genomes in

japonica nuclear backgrounds, and visa versa, throughout Himalayan region of Southeast Asia [25,61]. Therefore, it appears that higher outcrossing rates and fewer sterility barriers provided a natural corridor for domestication genes to be transferred between the two emerging varietal groups. Yet, despite evidence of substantial hybridization, the two *O. sativa* varietal groups remain largely isolated from each other at the genetic level. This apparent contradiction might be resolved by the recent discovery of well-defined regions of high and low divergence between *indica* and *japonica* in the rice nuclear genome [45]. Genomic regions of low diversity might be more permissive to recombination and, therefore, introgression between groups, whereas regions of high genetic diversity might reinforce the sterility barriers and help to maintain the independence of the *indica* and *japonica* gene pools [45]. Consistent with this hypothesis, Sweeney *et al.* recently documented that the domestication allele for white pericarp, *rc,* entered the *indica* gene pool as a small (<1 Mb) introgression from a *japonica* ancestor and that this region of introgression is characterized by an abrupt drop in F_{st} values (see Glossary) [51]. This indicates that the region of introgression can be classified as an island of low divergence in a background of well-differentiated *indica* or *japonica*-like genomes. Thus, observations at both the gene and the genome levels support the view that relatively small portions of genetic information were exchanged across gene pools during the process of rice domestication [45,51]. The introgressions containing key domestication loci would have been maintained by human selection. Limited amounts of introgression are consistent with both the deep genetic differentiation that distinguishes the *indica* and *japonica* varietal groups and the presence of common domestication alleles across *O. sativa*.

Concluding remarks – lessons for the future

The story of rice domestication is one of both genetic exchange and genetic containment. It involved episodes of hybridization and introgression between ancient gene pools against a backdrop of emerging inbreeding and sterility barriers that progressively restricted gene flow between subpopulations. These competing evolutionary forces, sculpted by human selection, gave rise to an array of interconnected, but well-differentiated, subpopulations of *O. sativa*. Although these subpopulations are intimately associated with their specific geographies, ecologies and cultures, the history of rice domestication is a dynamic rather than static story.

Today, these same forces of genetic exchange and containment are at work, and the tools of modern genetics provide new opportunities for the identification and the mobilization of genetic variation. High-yielding hybrid rice varieties take advantage of the traditional subpopulation structure to systematically create new combinations of genes by bringing together inbred parents from different subpopulations. The development of inter- and intraspecific introgression lines mimics the ancient road to domestication by allowing only selected segments of the donor genome to be introgressed into adapted genetic backgrounds. Genomics techniques offer a rapid way of identifying useful variation, which can then be exploited with transgenic technology. Yet, with all of these possibilities for genetic manipulation available to the modern plant breeder, the most valuable source of genetic variation today is still the wild germplasm itself. Learning how humans once domesticated crops from wild ancestors might provide more than a retrospective. As we seek new ways of harnessing genetic variation in the context of the 21st century, understanding the dynamics of crop domestication might offer important lessons for the future.

Acknowledgements

We acknowledge financial support from the Plant Genome Program of the National Science Foundation (Award Numbers #0606461 and #0110004) to S.R.M. We are grateful to Lois Swales for formatting and administrative help, and Jennifer Kimball-Pevo and Lisa Polewczak for photographic contributions. Also, many thanks to Laurent Sagart for numerous conversations concerning rice domestication.

References

1 Khush, G.S. (1997) Origin, dispersal, cultivation and variation of rice. *Plant Mol. Biol.* 35, 25–34
2 Gepts, P. (2004) Crop domestication as a long-term selection experiment. *Plant Breed. Rev.* 24, 1–44
3 Xiong, L.Z. *et al.* (1999) Identification of genetic factors controlling domestication-related traits of rice using an F_2 population of a cross between *Oryza sativa* and *O. rufipogon*. *Theor. Appl. Genet.* 98, 243–251
4 Kinoshita, T. (1998) Linkage mapping using mutant genes in rice. *Rice Genet. Nwsltr.* 15, 13–74
5 Gu, X.Y. *et al.* (2005) Genetic analysis of adaptive syndromes interrelated with seed dormancy in weedy rice (*Oryza sativa*). *Theor. Appl. Genet.* 110, 1108–1118
6 Gu, X.Y. *et al.* (2004) Multiple loci and epistases control genetic variation for seed dormancy in weedy rice (*Oryza sativa*). *Genetics* 166, 1503–1516
7 Gu, X.Y. *et al.* (2005) Phenotypic selection for dormancy introduced a set of adaptive haplotypes from weedy into cultivated rice. *Genetics* 171, 695–704
8 Miyata, M. *et al.* (2007) Marker-assisted selection and evaluation of the QTL for stigma exsertion under *japonica* rice genetic background. *Theor. Appl. Genet.* 114, 539–548
9 Uga, Y. *et al.* (2003) Mapping QTLs influencing rice floral morphology using recombinant inbred lines derived from a cross between *Oryza sativa* L. and *Oryza rufipogon* Griff. *Theor. Appl. Genet.* 107, 218–226
10 Thomson, M.J. *et al.* (2003) 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.* 107, 479–493
11 Sweeney, M.T. *et al.* (2006) Caught red-handed: *Rc* encodes a basic helix-loop-helix protein conditioning red pericarp in rice. *Plant Cell* 18, 283–294
12 Furukawa, T. *et al.* (2007) The *Rc* and *Rd* genes are involved in proanthocyanidin synthesis in rice pericarp. *Plant J.* 49, 91–102
13 Kato, S. and Ishikawa, J. (1921) On the inheritance of the pigment of red rice. *Jpn. J. Genet.* 1, 1–7
14 Doebley, J.F. *et al.* (2006) The molecular genetics of crop domestication. *Cell* 127, 1309–1321
15 Ting, Y. (1957) The origin and evolution of cultivated rice in China. *Acta Agron. Sinica* 8, 243–260
16 Chou, S.L. (1948) China is the place of origin of rice. *J. Rice Soc., China* 7, 53–54
17 Garris, A.J. *et al.* (2005) Genetic structure and diversity in *Oryza sativa* L. *Genetics* 169, 1631–1638
18 Glaszmann, J.C. (1987) Isozymes and classification of Asian rice varieties. *Theor. Appl. Genet.* 74, 21–30
19 Lu, B.R. *et al.* (2002) Genetic differentiation of wild relatives of rice as assessed by RFLP analysis. *Theor. Appl. Genet.* 106, 101–106
20 Chang, T.T. (1976) Origin, evolution, cultivation, dissemination, and diversification of Asian and African rices. *Euphytica* 25, 425–441
21 Morishima, H. and Gadrinab, L.U. (1987) Are the Asian common wild rices differentiated into the indica and japonica types? In *Crop*

Exploration and Utilization of Genetic Resources (Hsieh, S.C., ed.), pp. 11–20, Taichung District Agricultural Station
22 Wang, X. *et al.* (1984) Coordinated studies on genetic resources of rice in Yunnan. III. Glaborous hull rice (guangkedao) in Yunnan. *Res. Bull. Beijing Agric. Univ.* 10, 333–344
23 Oka, H.I. and Morishima, H. (1982) Phylogenetic differentiation of cultivated rice, potentiality of wild progenitors to evolve the *indica* and *japonica* types of rice cultivars. *Euphytica* 31, 41–50
24 Second, G. (1982) Origin of the genetic diversity of cultivated rice (*Oryza* spp.): study of the polymorphism scored at 40 isozyme loci. *Jpn. J. Genet.* 57, 25–57
25 Sun, Q. *et al.* (2002) Genetic differentiation for nuclear mitochondrial and chloroplast genomes in common wild rice (*O. rufipogon* Griff) and cultivated rice (*O. sativa* L). *Theor. Appl. Genet.* 104, 1335–1345
26 Londo, J.P. *et al.* (2006) Phylogeography of Asian wild rice, *Oryza rufipogon* reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proc. Natl. Acad. Sci. U. S. A.* 103, 9578–9583
27 Nakano, M. *et al.* (1992) Phylogenetic study of cultivated rice and its wild relatives by RFLP. *Rice Genet. Nwsltr.* 9, 132–134
28 Dally, A.M. and Second, G. (1990) Chloroplast DNA diversity in wild and cultivated species of rice genus *Oryza* section *Oryza* cladistic-mutation and genetic-distance analysis. *Theor. Appl. Genet.* 80, 209–222
29 Chen, W-B. *et al.* (1993) Distribution of deletion type in cpDNA of cultivated and wild rice. *Jpn. J. Genet.* 68, 597–603
30 Rakshit, S. *et al.* (2007) Large-scale DNA polymorphism study of *Oryza sativa* and *Oryza rufipogon* reveals the origin and divergence of Asian rice. *Theor. Appl. Genet.* 114, 731–743
31 Duan, S. *et al.* (2007) Phylogenetic analysis of AA-genome *Oryza* species (Poaceae) based on chloroplast, mitochondrial and nuclear DNA sequences. *Biochem. Genet.* 45, 113–129
32 Cheng, C. *et al.* (2003) Polyphyletic origin of cultivated rice: Based on the interspersion pattern of SINEs. *Mol. Biol. Evol.* 20, 67–75
33 Hu, H. *et al.* (2006) 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. Plant. Biol.* 48, 260–267
34 Caicedo, A. *et al.* (2007) Genome-wide patterns of nucleotide polymorphism in domesticated rice. *PLoS Genet*. DOI: 10.1371/journal.pgen.0030163.eor http://genetics.plosjournals.org/
35 Zhou, H.F. *et al.* (2003) Microsatellite analysis of genetic diversity and population genetic structure of a wild rice (*O. rufipogon* Griff) in China. *Theor. Appl. Genet.* 107, 332–339
36 Barbier, P. (1989) Genetic variation and ecotypic differentiation in the wild rice species *O. rufipogon*. II. Influence of the mating system and life history traits on the genetic structure of populations. *Jpn. J. Genet.* 63, 273–285
37 Oka, H.I. (1988) *Origin of Cultivated Rice,* Elsevier Science/Japan Scientific Societies Press
38 Messeguer, J. *et al.* (2004) A field study of pollen-mediated gene flow from Mediterranean GM rice to conventional rice and the red rice weed. *Mol. Breed.* 13, 103–112
39 Goff, S.A. *et al.* (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296, 92–100
40 Yu, J. *et al.* (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* 296, 79–92
41 International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. *Nature* 436, 793–800
42 Zhu, Q. and 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
43 Ma, J. and Bennetzen, J. (2004) Rapid recent growth and divergence of rice nuclear genomes. *Proc. Natl. Acad. Sci. U. S. A.* 101, 12404–12410
44 Vitte, C. *et al.* (2004) Genomic paleontology provides evidence for two distinct origins of Asian rice (*Oryza sativa* L.). *Mol. Genet. Genomics* 272, 504–511
45 Tang, T. *et al.* (2006) Genomic variation in rice: genesis of highly polymorphic linkage blocks during domestication. *PLoS Genet.* 2, e199
46 Fuller, D.Q. (2007) Contrasting patterns in crop domestication and domestication rates: recent archaeobotanical insights from the Old World. *Ann. Bot. (Lond.)*. DOI: 10.1093/aob/mcm048 http://aob.oxfordjournals.org/
47 Li, C. *et al.* (2006a) Genetic analysis of rice domestication syndrome with the wild annual species, *Oryza nivara*. *New Phytol.* 170, 185–193
48 Li, C. *et al.* (2006b) Rice domestication by reducing shattering. Science 311, 1936–1939
49 Lin, Z. *et al.* (2007) Origin of seed shattering in rice (*O. sativa* L.). *Planta* 226, 11–20
50 Konishi, S. *et al.* (2006) An SNP caused loss of seed shattering during rice domestication. *Science* 312, 1382–1396
51 Sweeney, M.T. *et al.* (2007) Global dissemination of a single mutation conferring white pericarp in rice. *PLoS Genet* 3, e133
52 Wang, Z.Y. *et al.* (1995) The amylose content in rice endosperm is related to the post transcriptional regulation of the *waxy* gene. *Plant J.* 7, 613–622
53 Yamanaka, S. *et al.* (2004) Identification of SNPs in the *waxy* gene among glutinous rice cultivars and their evolutionary significance during the domestication process of rice. *Theor. Appl. Genet.* 108, 1200–1204
54 Frances, H. *et al.* (1998) Use of alternative splice sites in granule-bound starch synthase mRNA from low-amylose rice varieties. *Plant Mol. Biol.* 38, 407–415
55 Fan, C. *et al.* (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
56 Song, X-J. *et al.* (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* 39, 623–630
57 Ashikari, M. *et al.* (2005) Cytokinin oxidase regulates rice grain production. *Science* 309, 741–745
58 Bradbury, L.M. *et al.* (2005) The gene for fragrance in rice. *Plant Biotechnol. J.* 3, 363–370
59 Sang, T. and Ge, S. (2007) The puzzle of rice domestication. *J. Integr. Plant Biol* 49, 760–768
60 Harushima, Y. *et al.* (2002) Diverse variation of reproductive barriers in three intraspecific rice crosses. *Genetics* 160, 313–322
61 Ishikawa, R. *et al.* (2006) Genetic erosion from modern varieties into traditional upland rice cultivars. *Genet. Res. Crop Evol.* 53, 245–252
62 Smartt, J. and Simmonds, N.W., eds (1995) *Evolution of Crop Plants*, Longman Group UK Limited
63 Yamanaka, S. *et al.* (2003) Dual origin of the cultivated rice based on molecular markers of newly collected annual and perennial strains of wild rice species, *Oryza nivara* and *O. rufipogon*. *Genet. Res. Crop Evol.* 50, 529–538
64 Li, C. *et al.* (2006c) Genetic analysis of rice domestication syndrome with the wild annual species, *Oryza nivara*. *New Phytol.* 170, 185–193
65 Morishima, H. (2001) Evolution and domestication of rice. *Rice Genet. Nwsltr.* IV, 63–78
66 Zhu, Q. *et al.* (2007) Multilocus analysis of nucleotide variation of *Oryza sativa* and its wild relatives: severe bottleneck during the domestication of rice. *Mol. Biol. Evol.* 24, 875–888
67 Kato, S. *et al.* (1928) On the affinity of rice varieties as shown by fertility of hybrid plants. *Bull. Sci. Fac. Agric. Kyushu Univ.* 3, 132–147
68 McCouch, S.R. *et al.* (1988) Molecular mapping of rice chromosomes. *Theor. Appl. Genet.* 76, 815–829
69 Wang, Z. and Tanksley, S. (1989) Restriction fragment length polymorphism in *Oryza sativa* L. *Genome* 32, 1113–1118
70 Zhang, Q.F. *et al.* (1992) Genetic diversity and differentiation of *indica* and *japonica* rice detected by RFLP analysis. *Theor. Appl. Genet.* 83, 495–499
71 De Candolle, A.P., ed. (1882), *Origins of Cultivated Species* (Reprint 1967), Hafner
72 Vavilov, N.I. (1926) Studies in the origin of cultivated plants. *Bull. Appl. Bot.* 16, 139–248
73 Zong, Y. *et al.* (2007) Fire and flood management of coastal swamp enabled first rice paddy cultivation in east China. *Nature* 449, 459–462
74 Zhao, Z. (1998) The Middle Yangtze region in China is one place where rice was domesticated: phytolith evidence from the Diaotonghuan Cave, *Northern Jiangxi. Antiquity* 72, 885–897
75 Lu, H. *et al.* (2002) Rice domestication and climatic change: phytolith evidence from East China. *Boreas* 31, 378–385
76 Higham, C. (2002) Languages and farming dispersals: austroasiatic languages and rice cultivation. In *Examining the Farming/Language Dispersal Hypothesis* (Bellwood, P. and Renfrew, C., eds), pp. 223–232, McDonald Institute for Archaeological Research

77 Normile, D. (1997) Archaeology: Yangtze seen as earliest rice site. *Science* 275, 309–310
78 Jiang, L. and Liu, L. (2006) New evidence for origins of sedetism and rice domestication in the Lower Yangzi River. *China. Antiquity* 80, 355–361
79 Diamond, J. and Bellwood, P. (2003) Farmers and their languages: the first expansions. *Science* 300, 597–603
80 Bellwood, P. (2005) Examining the language/farming dispersal hypothesis in the East Asian context. In *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics* (Sagart, L., Blench, R. and Sanchez-Mazas, A., eds), pp. 17–30, London, RoutledgeCurzon
81 Fuller, D. (2002) An agricultural perspective on Dravidian historical linguistics: archaeological crop packages, livestock and Dravidian crop vocabulary. In *Examining the Farming/language Dispersal Hypothesis* (Bellwood, P. and Renfrew, C., eds), pp. 191–213, McDonald Institute for Archaeological Research