A Rice Genetic Improvement Boom by Next-generation Sequencing

Xiangchun Zhou†, Xufeng Bai† and Yongzhong Xing*

National Key Laboratory of Crop Genetic Improvement and National Center of Plant Gene Research, Huazhong Agricultural University, Wuhan, China.

*Correspondence: yzxing@mail.hzau.edu.cn
†These authors contributed equally

https://doi.org/10.21775/cimb.027.109

Abstract

Rice (*Oryza sativa*. L) is a staple food crop for people worldwide, and a key goal has been to increase its grain yield. An increasing population that relies on a decreasing level of farmland has rendered the traditional method for the isolation and use of genetic loci in rice breeding unsatisfactory. Recently, the rapid development in next-generation sequencing (NGS) has boosted the number of genome sequences for hundreds to thousands of rice varieties. A MutMap strategy and bulk segregation analysis (BSA) has been developed to directly identify candidate genes based on NGS. The genome-wide association analysis (GWAS) has become a commonly used approach towards identifying the genetic loci and candidate genes for several traits that are closely associated with grain yield. The Multi-parent Advanced Generation Inter-Cross population (MAGIC) is introduced here to discuss potential applications for mapping QTLs for rice varietal development. These strategies broaden the capacity of functional gene identification and its application as a complementary method to insert mutants that comprise T-DNA and transposons. High-throughput SNP analysis platforms, such as the SNP array, provide novel strategies for genomic-assisted selections (GAS) for rice genetic improvements. Moreover, accurate genome sequence information enables genome editing for the utilization of key recessive genes that control important agronomic traits. This review summarizes how NGS accelerates rice genetic improvements through the identification and utilization of key functional genes that regulate agronomic traits.

Introduction

The world population will reach 9 billion in 2050, and a 70–100% increase in food production will be required to meet the demand if the population continues to increase at the current rate (Godfray *et al.*, 2010). Rice is a vitally important food crop worldwide (Xing and Zhang, 2010). The limited area of farmland, the occurrence of extreme climates, and emerging biotic and abiotic stresses have driven the need for rice genetic improvements. Additional knowledge of functional genes for agronomic traits is expected to inform the breeding design for rice. Next-generation sequencing (NGS) provides numerous genome resources by re-sequencing diverse accessions, including landraces and varieties. Millions of single nucleotide polymorphisms (SNPs) are available, and they facilitate the genome-wide association study (GWAS). Recently, NGS has enabled rice genetic improvements by identifying hundreds of key functional genes or by establishing associations between SNPs and agronomic traits.

The available methods for identification of functional genes

There are over 40,000 putative genes in rice (Feng *et al.*, 2002; Goff *et al.*, 2002; Sasaki *et al.*, 2002; Yu *et al.*, 2002). Most genes are quantitative trait loci (QTL), each with a minor effect. Currently, a small group of genes that control agronomic traits, such as the yield traits and heading date (which are all controlled by multiple QTLs), have been identified by a forward genetics strategy. In the past two decades, thousands of QTLs have been located (www.gramene.org/qtl), but only approximately 100 of the QTLs were cloned by map-based cloning. Map-based cloning is time-consuming and laborious. Therefore, several alternative approaches have been developed to efficiently discover functional genes, including gene knockdown and gene knockout.

Antisense and RNA interference (RNAi) technologies (Chuang and Meyerowitz, 2000) decrease the expression of target genes, which can lead to phenotypic changes. This method enables a direct mutation of a target gene. Direct gene knockout is accomplished using engineering techniques, such as Zinc finger nuclease (ZFN), transcription activator-like effector nuclease (TALEN), and clustered regularly interspaced short palindromic repeats-associated protein 9 (CRISPR-Cas9). Random gene knockout can cause functional mutations, which are randomly produced by a transposon insertion, a T-DNA insertion or treatments with physical factors and chemical mutagens.

Mutant library for functional genomics in rice

T-DNA insertion library

The isolation of genes using T-DNA mutants based on reverse genetics is common, and a series of T-DNA tagging populations have been developed. For example, approximately 100,000 fertile rice T-DNA lines have been tagged in Korea (Jeon *et al.*, 2000; Jeong *et al.*, 2002), and over 30,000 T-DNA insertion lines have been resulted (Wu *et al.*, 2003). Several other groups have reported T-DNA insertion lines in the genetic background of japonica rice (Chen *et al.*, 2003; Sallaud *et al.*, 2003; Sha *et al.*, 2004; Yin and Wang, 2000). No matter how large the T-DNA library is, it does not ensure that all genes will be inserted because the T-DNA insertion preferentially occurs in gene-rich regions (An *et al.*, 2005; An *et al.*, 2003; Barakat *et al.*, 2000; Chen *et al.*, 2003; Wu *et al.*, 2003). Moreover, T-DNA insertions are non-uniformly distributed on each chromosome; they are higher at the distal ends and lower near the centromeres. Additionally, several regions show extreme peaks and valleys of insertion frequency, which suggests hot and cold spots for T-DNA integration (Jeong *et al.*, 2006). Applying this strategy to *indica* rice has been particularly difficult because the transformation efficiency is not high enough to produce the thousands of requisite transgenic plants. Therefore, other mutant library types are needed to complement the T-DNA library vacancy.

Transposon-derived library

A two-element *Activator/Dissociation* (*Ac/Ds*) transposon system has been extended to rice from maize (Enoki *et al.*, 1999; Greco *et al.*, 2001, 2003; Kolesnik *et al.*, 2004). The generation of a *Ds* transposon population has been achieved using a regeneration procedure that involves culturing seed-derived calli that carry *Ac* and non-autonomous *Ds* elements. *Ds* insertions are randomly distributed throughout the genome, but a hot spot for *Ds* insertions has been identified on chromosome 7 within a 40-kb region (Kolesnik *et al.*, 2004). However, its insertion preference and unstable insertion across different generations have also been reported (An *et al.*, 2005; Greco *et al.*, 2003).

Tos17 is a retrotransposon in rice. Although its copy number is low in plants (one to five, depending on the cultivar), it can increase up to 30 copies during tissue culture (Hirochika *et al.*, 1996). A previous report describes the generation of 47,196 *Tos17*-induced insertion mutants in rice (Miyao *et al.*, 2003). Although insertion targets are distributed throughout the chromosomes, they tend to cluster, and 76% of the clusters are located in genic regions. Moreover, *Tos* populations can only play minor roles in genome-wide mutagenesis because the number of original donor sites is limited (Kim *et al.*, 2004).

EMS-induced mutant library

The establishment of a large cohort of insertion mutants will accelerate investigations of rice gene functions via reverse genetics approaches.

However, the preferred mutagenesis locations of the *Ac*/*Ds*-, *Tos17*- and T-DNA-derived mutant libraries make them incapable of covering all genes. Therefore, physical and chemical mutagenesis strategies are strong candidates to complement their shortcomings. Fast neutrons, gamma rays, and diepoxybutane induce deletions or translocations; ethyl methanesulfonate (EMS) always induces point mutations. IR64, an *indica* rice, possesses multiple favourable agronomic characteristics (e.g. wide adaptability, high yield potential, tolerance to multiple diseases and pests, and a suitable quality for eating) and is an ideal genotype for identifying mutational phenotype changes (Wu *et al*., 2005). Approximately 60,000 IR64 mutants have been generated using chemicals [diepoxybutane (DEB) and EMS] and irradiation (fast neutron and gamma ray). Targeting Induced Local Lesions in Genomes (TILLING) is an efficient way to identify point mutations (Comai and Henikoff, 2006). However, TILLING has largely been avoided as a first option towards identifying mutations because the procedure, which includes the specific gene primer design, PCR application, digestion with a crude celery juice extract that contains the CELI nuclease, and electrophoresis with a denaturing polyacrylamide gel, is lengthy.

Next-generation sequencing-based MutMap strategy for gene mutation

MutMap strategies

It is easy to obtain a large number of mutants using the transposon technique and chemical and physical mutagens, but it is difficult to target the mutation that is responsible for the phenotype change. The rapid development of NGS has illuminated and improved the ability to identify nucleotide variations at the genome level. MutMap has been used for gene isolation in rice (Abe *et al*., 2012). The MutMap strategy frequently involves multiple steps (Fig. 7.1). A rice elite cultivar with

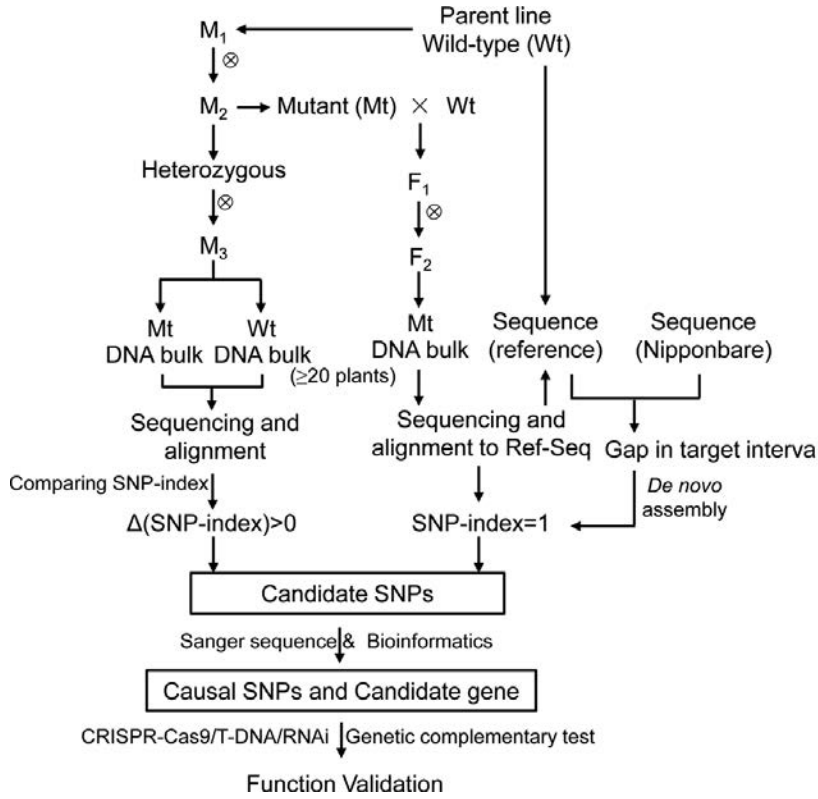

Figure 7.1 The combined scheme of the different strategies of MutMap.

a reference genome sequence, such as Hitomebore from Northern Japan, has been mutagenized with the EMS mutagen to yield the M_1 generation, which self-pollinated to yield M_2; the M_2 segregation population yielded the recessive homozygous mutant with an altered agronomic trait, such as heading date, plant height, panicle architecture or grain size. The mutant was crossed with the Hitomebore wild-type to produce the heterozygous F_1, and the F_2 progeny were developed by self-pollination of F_1. The F_2 segregation population comprised over 100 plants. The DNA from 20 mutants was equally bulked and subjected to whole genome sequencing. The depth of the sequence was typically above 10-fold coverage of the rice genome (370 Mb). These mutant sequences were aligned to the reference sequence of the wild-type plants. The SNP-index was calculated, and the candidate SNPs were identified with a SNP-index = 1. The causal SNPs and the candidate genes were determined by further analysis of the Sanger sequence and with bioinformatics. Finally, the candidate gene was verified by overexpression, RNAi or knockout. *OsCOA1* and *OsRR22*, which are two genes that control pale green leaf and salinity tolerance, were identified by the MutMap strategy (Table 7.1) and validated with mutants and transgenic plants (Abe *et al.*, 2012; Takagi *et al.*, 2015).

The mutants with early development lethality or sterility cannot be crossed with the wild-type; this makes them unsuitable for MutMap applications. MutMap+ has been developed by a versatile extension of MutMap to address this problem (Fekih *et al.*, 2013). MutMap+ selects a heterozygous plant from the M_2 generation to give rise to the M_3 segregating population by selfing, which contains the wild-type and homozygous mutants. The DNA bulks (≥ 20 plants) for the wild-type and mutants have been constructed for whole-genome sequencing. The sequences from both bulks were aligned to the reference sequence of the cultivar that was used for mutagenesis. SNP-index plots were calculated for both the mutant and wild-type M_3 bulks. The two SNP-index plots were compared to identify the region with SNP-index = 1 that was specific to the mutant bulk. In other words, the Δ (SNP-index) plot that was obtained by subtracting the wild-type bulk SNP-index value from that of the mutant bulk was a positive value [Δ (SNP-index) > 0]. The candidate SNPs were likely harboured in the genomic region of the peak value of the Δ (SNP-index). The causal SNPs and the candidate genes were then determined by further analysis of the Sanger sequence and with bioinformatics. Finally, the candidate gene was verified by overexpression, RNAi or gene knockout (Fig. 7.1). *OsNAP6*, which encodes a SufBD protein that regulates dwarfism and pale green leaf, has been identified by the MutMap+ strategy, and its function has been validated by RNAi (Takagi *et al.*, 2015). A gene (Os08g0139100) for albino seeding, which is incapable of growing to maturity, has also been isolated with the MutMap+ method (Takagi *et al.*, 2015). In specific cases where the re-sequenced cultivar/line displays a significant structural variation from the reference genome (Nipponbare), mutations in genome regions that are missing from the reference genome (gaps) cannot be identified by simple alignment. A method called 'MutMap-Gap' has been developed (Takagi *et al.*, 2013b), which involves delineating a candidate region that harbours a mutation of interest using the MutMap method, followed by de novo assembly, alignment, and identification of the mutation within the genome gaps. The candidate SNPs are further identified based on SNP-index = 1, and the causal SNPs and candidate gene are determined by the analysis of the Sanger sequence and with bioinformatics. Finally, they are validated with mutants and/or transgenic plants (Fig. 7.1). *HIT7*, which encodes a NBS-LRR type R protein that is resistant to rice blast fungus, has been isolated using this method (Takagi *et al.*, 2013b).

Options for further analysis of MutMap detection

MutMap is an efficient approach for isolating mutated genes, but the prerequisite for this strategy's use is the easy identification of distinct phenotypic types in a F_2 population that is derived from a cross between a wild-type and mutant. Specifically, a single Mendelian segregation ratio of 3:1 should be observed before MutMap is used to target the mutation. There is always a large candidate region when the MutMap strategy is used for gene identification. For example, the candidate regions of the five genes that were isolated by the MutMap strategy ranged from 2.1 to 5.8 Mb; the number of SNPs in the target region ranged from two to ten (Table 7.1). The SNPs of the coding

Table 7.1 The summary information of the cloned genes using the different MutMap strategies

Mutants	Sequence depth	Candidate region (Mb)	SNPs	Causal SNP	Validation	Gene	Function annotation	Strategy	Reference
Pale green leaf	>12×	2.1	7	Leu to Phe	Mutant, RNAi and complementary test	*OsCAO1* (Os10g0567400)	Chlorophyll a oxygenase	MutMap	Abe *et al*., 2012
Salinity tolerance	20×	5.8	2	Premature stop codon	Tos17 mutant	*OsRR22* (Os06g0183100)	B-type response regulator	MutMap	Takagi *et al*., 2015
Dwarfism and pale green leaf	Mt:19.4×Wt:11.7×	5.2	8	Ala to Thr	RNAi	*OsNAP6* (Os01g0127300)	SufBD protein	MutMap+	Fekih *et al*., 2013
Albino seeding	Mt:12×Wt:11.4×	4.3	10	Premature stop codon		Os08g0139100	Homologues of chloroplast precursor differentiation and greening	MutMap+	
Blast fungus resistance	Mt1:6.6×	4.1	4	Premature stop codon	Association analysis	*HIT7* (Os09t0327600–01)	NBS-LRR type R protein	MutMap-Gap	Takagi *et al*., 2013b
	Mt2:7.7×	5.8	10	Splicing error					

region induced a premature stop codon, amino acid change and splicing error that frequently caused varied phenotypes and were commonly considered the most likely causal SNPs (Table 7.1). The corresponding gene that carries the causal SNP may be the candidate gene. This has been further validated by another method, which produced mutants or transgenic plants (Fig. 7.1). Notably, the regulatory region is also important for gene function, and mutations in the regulatory region may result in a mutant phenotype. Attention should be directed towards its candidate. Moreover, when the candidate SNPs are located on other genes without the annotation of function (unknown protein), it is difficult to predict the causal SNP and candidate gene. All the genes that harbour the candidate SNPs should be verified with other resource mutants and transgenic plants. Alternatively, a large F_2 population should be used to isolate the mutation by map-based cloning.

Genome-wide quantitative trait loci mapping

Most agronomic traits are quantitatively inherited in rice, and the natural variations in a segregating population frequently exhibit continuous extremes in high and low phenotypic values. These natural variations provide us with the opportunity to breed super varieties. Generating favourable variations from this diversity is the most important task towards genetic improvements. NGS has more advantages towards QTL mapping than traditional molecular marker techniques. It is ideally suited for a natural collection or for a bi- or multi-parental population.

Genome-wide association study

The GWAS is a routinely used method for identifying the association between markers and traits in humans and plants. Recently, several NGS projects have been widely applied to rice genomic studies, which have provided millions of SNPs. Because of their abundance and uniform distributions throughout the genome, SNPs are becoming the most desirable molecular markers for QTL scanning. Thus, GWAS has been regularly applied for different traits in rice. The agronomic traits have been investigated by several research groups using GWAS. A mixed linear model (MLM) has been introduced as an improved method to simultaneously account for the population structure and unequal relatedness among individuals, which should decrease false-positive associations (Zhang *et al.*, 2010). A report by Huang *et al.* (2010) describes the identification of approximately 3.6 million SNPs through the sequencing of 517 rice landraces; it also details the construction of a high-density haplotype map of the rice genome. GWAS was performed for 14 agronomic traits in a population of the *Oryza sativa indica* subspecies. The association population was extended to a larger and more diverse sample of 950 worldwide rice varieties, and GWAS was performed for 11 grain-related traits (Huang *et al.*, 2011). GWAS was performed for agronomic traits and metabolites by other groups (Bai *et al.*, 2016; Chen *et al.*, 2014b; Crowell *et al.*, 2016; Magwa *et al.*, 2016a,b; Zhao *et al.*, 2011). Huang *et al.* (2015) developed an integrated genomic approach using the SNP data that was obtained by NGS to construct a genome map for 1495 elite hybrid rice varieties and their inbred parental lines. In total, 38 agronomic traits were investigated, and 130 associated loci were identified. Based on the in-depth analyses of the effects of the heterozygous genotypes of these loci, they found that high-yielding hybrid varieties had several superior alleles, but most parental inbred lines had a small number. Therefore, they concluded that the accumulation of numerous rare superior alleles with positive dominance was an important contributor to the heterotic phenomena. Huang *et al.* (2016) reported the genomic sequencing of 10,074 F_2 lines from 17 representative hybrid rice crosses and constructed a dense genotype map with 347,803 recombination events using the NGS-based genotyping method. They classified modern hybrid rice varieties into three groups, which represented different hybrid breeding systems, and obtained support from several loci for a partial dominance of the heterozygous locus for yield-related traits. These results address the genomic architecture of heterosis and rice hybrid breeding. Han *et al.* (2016) performed an association analysis of 11 known flowering genes at the SNP and haplotype levels. The haplotype-level association analysis detected 7 of the 11 known genes, including the only four detected at the SNP level. Notably, the haplotype-level association strongly increased the power of QTL mapping. Moreover,

the haplotype-level association mapping suggested favourable haplotypes for agronomic traits. However, genome-wide haplotype-level association mapping continues to be a difficult procedure.

The resolution of GWAS is dependent on the SNP local linkage disequilibrium (LD). The average LD decay in cultivated rice has been extended from 100 kb to 200 kb (Huang *et al.*, 2010; Mather *et al.*, 2007; McNally *et al.*, 2009). Although it is difficult to map a QTL to a single gene, this resolution can provide valuable information for marker-selected breeding.

Recently, a Multi-parent Advanced Generation Inter-Cross population (MAGIC) was introduced to discuss potential applications for mapping QTLs and for rice varietal development. MAGIC that is derived from a cross between multiple parents has a genotypic diversity and reduced linkage drag (Leung *et al.*, 2015). There is a clear genealogy among the lines and between the generations, which can effectively avoid the population structure. Several MAGIC populations that were derived from diverse elite parental lines have been used to detect QTLs for different traits in rice. The MAGIC genetic map was constructed using the SNP markers based on NGS and/or SNP array (Bandillo *et al.*, 2013; Meng *et al.*, 2016a,b). A GWAS with these MAGIC populations had a higher detection power than with the assembled populations/germplasm collection and had a higher resolution than the bi-parental populations (Meng *et al.*, 2016a). Moreover, the Hidden Markov Model (HMM) was used for a QTL analysis in MAGIC, which led to a low false positive rate and a low bias of the estimated QTL effect sizes. An implementation of the approach is available as an R package (Verbyla *et al.*, 2014). Currently, GWAS is primarily performed based on the SNPs. Here, we strongly suggest the use of bin-based GWAS for QTL mapping; this may improve the precision of the QTL location and reduce the false positive associations compared with those of the SNP-based GWAS.

Quick QTL mapping via bulk segregation analysis

F_2 and/or BC_1F_2 are generally used for QTL mapping. To further isolate the candidate genes that underlie the QTLs that are mapped in the F_2, RIL or BC_1F_2 primary populations, advanced populations, such as near isogenic lines (NILs), are required. However, five to six cropping seasons are necessary to develop the advanced population, which is a lengthy time period. Fortunately, the BSA strategy (Takagi *et al.*, 2013a) has become more vigorous towards identifying QTLs with the NGS application in rice, particularly for major QTLs. Two DNA bulks are constructed by selecting extremely different phenotypic groups from the mapping populations. The QTL region is quickly mapped with an in-depth comparative analysis of the NGS-based sequence. Specifically, if the allele frequency in one region is very different between these two bulks, the region is most likely to contribute to the phenotypic difference. Applications of this strategy will quickly provide valuable information for breeders and will shorten the required times for major QTL isolations in rice.

Genomic selection for breeding

Genetic diagnosis

The selection of parental lines is a key step in the genetic research and breeding processes of rice (Chen *et al.*, 2013). Breeders often choose varieties with diverse genetic backgrounds for their breeding programs. For example, to screen QTLs for different agronomic traits, a population developed from the varieties with dramatically different backgrounds will enhance the mapping power and will map numerous QTLs. Phylogenetic analysis is useful with this breeding protocol. Rice varieties can be clustered into several subgroups based on the SNP analysis. The distance in the neighbour-joining tree indicates the genetic similarities among these varieties. The cluster information can be used to guide the parental line selections for genetic studies and rice breeding, such as in the construction of MAGIC populations for genetic studies. Parent founders with higher diversity may provide more opportunities for recombination, which enables efficient and precise explorations of multiple desirable traits (Bandillo *et al.*, 2013).

On the other hand, for hybrid breeding, we can perform genetic diagnosis through the SNP analysis to select the parental lines that have superior alleles to control yield-related traits, such as the plant height, heading date, panicle or grain shape. Exploitation of hybrid vigour (heterosis) has been accomplished in rice production. Several

yield-related genes, such as *sd1*, *Ghd7*, *Ghd8*, *gs3*, *dep1*, *ipa1*, and so on (Table 7.2), have been identified as being critical for hybrids and as having partial heterotic effects (Huang *et al.*, 2015, 2016). Huang *et al.* (2016) found that genes for yield-related traits were enriched for positive dominance effects, which indicated that superior alleles could be optimally utilized by complementing parents. Furthermore, most of these superior alleles were contributed by the male parents. These findings will be helpful towards rational genetic designs for producing rice hybrids. For example, using SNP analysis, we can choose a more suitable male parent with the desired alleles for *sd1*, *gs3*, *tac1* and *Ghd8* and a more suitable female parent with superior alleles for *nal1*, *ipa1*, and *dep1*. With these favourable alleles, the expected hybrid would exhibit a high-yield potential with an ideal plant architecture.

Genomic-assisted selection

Several NGS projects have been widely applied to rice genomic studies (Huang *et al.*, 2010, 2011, 2012; McNally *et al.*, 2009; Xu *et al.*, 2011), which have generated millions of SNPs. Benefiting from NGS, high-throughput SNP arrays have been developed to genotype numerous samples in a short timeframe by identifying thousands of SNPs that are distributed in the whole genome. Various SNP array platforms based on different principles are available in rice. An Affymetrix 44K SNP array has been developed and successfully used in GWAS (Famoso *et al.*, 2011; McCouch *et al.*, 2010; McNally *et al.*, 2009; Zhao *et al.*, 2011). However, most SNPs on the array are derived from the *Oryza*SNP project with 20 varieties (McNally *et al.*, 2009), which may not adequately represent the rice species. Various SNP chips based on the

Table 7.2 Collection of key genes controlling yield-related traits

Gene	Chromosome	Trait	Yield-related phenotype	Heterosis	Reference
*sd1**	1	PH	High YD, reduced PH	+	Spielmeyer *et al*., 2002
hd3a	6	HD	High YD, delayed HD	+	Kojima *et al*., 2002
*Hd1***	6	HD	High YD, delayed HD	+	Yano *et al*., 2000
Ghd7	7	HD	High YD, delayed HD	+	Xue *et al*., 2008
Ghd8	8	HD	High YD, delayed HD	+	Yan *et al*., 2011
Ghd7.1	7	HD	High YD, delayed HD		Yan *et al*., 2013
gs3	3	GL and KGW	Increased GL and KGW	+	Fan *et al*., 2006
GS5	5	GW and KGW	Increased GW and KGW		Li *et al*., 2011
GW2	2	GW and KGW	Increased GW and KGW	+	Song *et al*., 2007
GW5	5	GW and KGW	Increased GW and KGW	+	Weng *et al*., 2008
TGW6	6	KGW	Increased YD and KGW	+	Ishimaru *et al*., 2013
gn1a	1	TGN	Increased YD and KGW	+	Ashikari *et al*., 2005
osgi	1	GN	Increased YD and GN		Izawa *et al*., 2011
dep1	9	GN	Increased YD, SPP and GN	+	Huang *et al*., 2009
nal1	4	GN	Increased YD, SPP and GN	+	Fujita *et al*., 2013
MOC1	6	Tillering	Increased YD and tillers		Li *et al*., 2003
LAX1	1	Tillering	Increased YD and tillers	+	Komatsu *et al*., 2001
htd1/ osccd7	4	Tillering	Increased YD and tillers	+	Kulkarni *et al*., 2014
ipa1/ osspl14	8	Ideal architecture	Increased YD, an ideal architecture	+	Miura *et al*., 2010
tac1	9	Tiller angle	increased YD, a compact plant status	+	Yu *et al*., 2007

*Gene in lower case means the negative regulation of rice yield; **Gene in capital format means the positive controlling. +, means that partial dominance or overdominance effect has been reported within the gene (Huang *et al*., 2015, 2016). YD, yield; PH, plant height; HD, heading date; GL, grain length; GW, grain width; KGW, kilo-grain weight; TGN, total grain number; GN, grain number; SPP, spikelets per panicle.

Illumina GoldenGate SNP Chip have also been developed for genetic analysis (Chen *et al.*, 2011; Yamamoto *et al.*, 2010; Zhao *et al.*, 2010). SNP arrays that are specific for rice genotyping have recently been designed based on the Illumina Infinium technology; they include RICE6K (Yu *et al.*, 2014) and RiceSNP50 (Chen *et al.*, 2014). RICE6K is based on the low-coverage genome sequences (~1×) of the 520 rice germplasm collections (Huang *et al.*, 2010). It contains 5556 SNP sites and accommodates 80 functional markers that cover 40 rice genes controlling traits such as grain yield, grain quality, heading date, hybrid fertility and stress resistance (Yu *et al.*, 2014). RiceSNP50, a new high-density array, was developed after the 520 accessions (~1×) (Huang *et al.*, 2010) and 281 accessions were sequenced (2.5–15×) (Chen *et al.*, 2014). RiceSNP50 has been synthesized with 51,478 SNPs that are evenly distributed along the whole rice genome. These SNP arrays have become useful for rice genetic research and breeding (Chen *et al.*, 2014; Hu *et al.*, 2013; Tan *et al.*, 2013; Yu *et al.*, 2014).

Marker-assisted selection (MAS) is a valuable tool for improving target traits with DNA markers that are linked to the genes/QTLs underlying the traits (Xu and Crouch, 2008). A promising MAS approach, known as genomic-assisted selection (GAS), has been widely used for genetic improvements in rice (Heffner *et al.*, 2009; Nakaya and Isobe, 2012). GAS is a high-throughput and efficient approach that uses genome-wide SNP markers to improve the agronomic traits; GAS completes the foreground and background examination in a single SNP array assay. During the breeding process with backcross (BC), breeders aim to introduce favourable alleles from a donor parent into the genome of a desired variety without changing the genetic background. The progeny that carry the target genes would then be selected during the early generations of the BC populations (positive selection), while the genetic background of the progeny that carry the donor alleles of interest would be screened to minimize the donor background in advanced BC populations, such as BC_4F_1. Rice SNP arrays, such as RICE6K and RiceSNP50, are able to achieve the goals of foreground and background selection. Closely linked low-density SSR markers or functional markers are typically used for positive selection. For GS3 (Fan *et al.*, 2006, 2009; Mao *et al.*, 2010), a major gene that controls grain length, a SNP between C and A in the second exon induces the differentiation of the rice grain length. A total of 180 varieties with genetic diversity have been used for association analysis with the C/A mutation and grain length; they have confirmed the causal mutation to be a SNP (Fan *et al.*, 2009). Based on the causal SNP, a CAPS marker, SF28, has been developed, and it works well in the MAS of the rice grain length (Fan *et al.*, 2009; Lu *et al.*, 2013; Wang *et al.*, 2011). We can also use the RICE6K SNP array for this selection because it accommodates 80 functional probes for 40 genes, including *GS3*. For an advanced BC population, the individuals that possess the target alleles are required for the genetic background selection. For example, 29 rice individuals with introgressions of *Pi1* (Hua *et al.*, 2012) and *Pi2* (Zhou *et al.*, 2006) into Kongyu 131, a japonica cultivar that is popular in north-east China, were examined by RICE6K (Yu *et al.*, 2014). The results showed that several improved lines with minimal donor parental backgrounds could be selected, but the target genomic regions also contained large, dragged fragments from the donor parent. These results suggest that closely linked markers or gene functional markers should be used for selection and should be performed on early generations. As another example, the RiceSNP50 array was used in rice BC breeding to assist a selection of the fragment with *Pigm(t)* (Deng *et al.*, 2006) from Gumei 4 (a rice variety that possesses broad-spectrum resistance) into an elite parental line, R608. The results showed that two out of four BC_4F_1 individuals were identified as possessing the *Pigm(t)* allele, and one improved line that contained one additional fragment from the donor parent was selected for further breeding (Chen *et al.*, 2014). These application examples indicate that the high efficiencies of Rice SNP arrays make them applicable to rice GAS breeding programmes.

Genome editing

Genome editing is a precise approach for targeting genetic manipulation and is becoming a popular tool for rice genetic improvements. Genome editing relies on accurate genome sequence information to target DNA fragments for mutations, which can result in a DNA insertion, deletion, replacement or reverse replication. ZFN,

TALEN and CRISPR-Cas9 are the three preferred genome-editing tools (Barrangou, 2012; Bibikova *et al.*, 2003; Bogdanove and Voytas, 2011; Pennisi, 2013).

In rice, several important genes that negatively control agronomic traits have been targeted through genome editing (Table 7.3). In 2012, TALEN was used to improve bacterial blight resistance by mutating the *OsSWEET14* promoter region, which harbours the cis-elements that interact with pathogens (Li *et al.*, 2012). This mutation decreased the binding effect between the promoter and the effectors that were secreted by the bacterial blight pathogen. The following year, the genes that showed negative effects on the rice architecture, yield, and quality were targeted by TALEN and CRISPR (Shan *et al.*, 2013a,b), with the aim of improving the rice yield and quality. The examples of genome editing applications to rice indicate that the induction-specific mutations in the negative regulatory genes can rapidly produce desirable agronomic traits. There are numerous, well-performing lines that contain one or more undesirable traits for the rice breeding process. There are also dozens of accessible genes that can be edited to improve these traits. Heading date is a vital trait for rice diversification and domestication (Izawa, 2007; Meyer and Purugganan, 2013). The molecular regulatory network for flowering is becoming clearer and more complete. Several important flowering genes have been extensively studied (Xue *et al.*, 2008; Yan *et al.*, 2011; Yano *et al.*, 2000); they provide flexibility towards manipulating the heading date by providing different gene combinations for selection. For example, we previously observed that different allele-combinations of the major heading date genes [*Ghd7* (Xue *et al.*, 2008), *Ghd8* (Yan *et al.*, 2011) and *Hd1* (Yano *et al.*, 2000)] largely defined eco-geographical adaptation in rice cultivars (Zhang *et al.*, 2015). Therefore, we suppose that by targeting *Ghd7*, *Ghd8*, or *Hd1*, we can rapidly design the flowering time of the desired varieties to adapt to different ecological regions based on the combinations of these genes. On the other hand, modifications of the *R* genes, such as *Xa5* (Iyer-Pascuzzi *et al.*, 2008; Iyer and McCouch, 2004) and *Xa13* (Antony *et al.*, 2010; Chu *et al.*, 2006; Yuan *et al.*, 2009), would cause the *xa5* and *xa13* recessive genes to function particularly well. In the rice bacterial blight system, the dominant allele, *Xa5*, is necessary and sufficient for pathogen susceptibility, but *xa5* is a completely recessive gene that contributes to bacterial blight resistance (Iyer-Pascuzzi *et al.*, 2008). Hence, we believe that through genome editing, we can efficiently obtain improved lines with the *xa5xa5* or *xa13xa13* homozygous alleles, which would enhance bacterial blight resistance.

Gene replacements and insertions have recently been achieved in rice (Li *et al.*, 2016). Using a pair of single guide RNAs (sgRNAs) that targeted adjacent introns and a donor DNA

Table 7.3 Application examples of genome editing in rice genetic improvements

Mutation types	Genes	Aims	Tool	Repair	Reference
Gene disruption	*OsSWEET14*	Bacterial blight resistance	TALEN/CRISPR	NHEJ	Li *et al*., 2012
	OsBADH2	Rice with fragrance	TALEN/CRISPR	NHEJ	Shan *et al*., 2013a,b, 2015
	Waxy	Lower levels of amylose	CRISPR	NHEJ	Ma *et al*., 2015
	ROC5	Rolling leaves rice	CRISPR	NHEJ	Feng *et al*., 2013
	SD1	Semi-dwarfing breeding	TALEN	NHEJ	Shan *et al*., 2013a
	DEP1	Erect panicle breeding	TALEN	NHEJ	Li *et al*., 2016; Shan *et al*., 2013a
	Gn1a	High yield	TALEN	NHEJ	Li *et al*., 2016; Shan *et al*., 2013a
	GS3	Bacterial blight resistance	CRISPR	NHEJ	Li *et al*., 2016
	IPA1	Bacterial blight resistance	CRISPR	NHEJ	Li *et al*., 2016
Gene replacement	*EPSPS*	Glyphosate resistance	CRISPR	NHEJ	Li *et al*., 2016
Gene insertion	*EPSPS*	Glyphosate resistance	CRISPR	NHEJ	Li *et al*., 2016

template with the same sgRNAs sites, gene replacements were achieved for an endogenous gene, 5-enolpyruvylshikimate-3-phosphate synthase (*EPSPS*), in rice at a frequency of 2.0%. Moreover, the gene insertion frequency was 2.2% with this intron targeting method. Genetically modified plants with the *OsEPSPS* gene substitution were glyphosate resistant. These achievements for target insertions in rice genes may assist breeders with the introduction useful gene clusters as breeding modules, which could accelerate genetic improvements for the utilization of multiple genes.

The promise of genome editing in rice is its fast and accurate modifications/introductions of ideal agronomic traits for genetic improvements. However, there are always critical limitations, such as off-target effects. Procedures to minimize the off-target effects have been developed and include the double nickase system (Mali *et al.*, 2013) and an inactivated Cas9 (dCas9) that is fused to FokI (Guilinger *et al.*, 2014). Nevertheless, off-target editing can still occur. Another limitation is food safety. Whether the improved lines with the targeted DNA modifications/introductions can be described as genetically modified organisms (GMOs) remains a critical question (Hartung and Schiemann, 2014; Lusser and Davies, 2013; Pauwels *et al.*, 2014). Araki and Ishii (2015) suggest that the genome editing events that result in small insertions/deletions through the non-homologous end joining (NHEJ) repair pathway should not be regulated as GMOs, while genome editing that induces a foreign template by homologous recombination should be considered as a GMO. This issue remains controversial (Schaart *et al.*, 2016; Wolt *et al.*, 2016).

Omics integration will push rice genetic improvement forward

Several NGS-based omics systems have been established, and high-throughput omics measurements are routinely explored in rice, including genomics, transcriptomics, and epigenomics. Additionally, proteomics (Chen *et al.*, 2016; Das *et al.*, 2016), metabolomics (Chen *et al.*, 2014) and phenomics (McCouch *et al.*, 2016; Yang *et al.*, 2014, 2015) have been established for high-throughput measurements of phenotypic values, including biochemical, physiological and agronomic traits. Determining the association between the genome variation [SNP, Insertion/deletion (InDel) and copy number variation (CNV)] and trait performance is the crucial step towards estimating genetic effects for rice breeding. GWAS for metabolites has detected more loci with higher resolution than agronomical traits, and 36 candidate genes have been identified, five of which are functionally validated in rice (Chen *et al.*, 2014). However, the relationship between the metabolites and agronomical traits is unclear for most cases. Comparison of the GWAS for agronomics with that for the metabolites, such as assimilates and phytohormones, would address uncertainties by providing compensatory or supporting information for one dataset. False positives are less likely when multiple sources of evidence point to the same gene or pathway. A complete biological model is likely to be discovered only when omics at different levels are jointly considered in an analysis. Integration of these omics datasets could categorize them into two clades – a meta-dimensional analysis and a multistaged analysis (Ritchie *et al.*, 2015). Meta-dimensional analysis is a method where different scales of data are simultaneously combined to identify a complex. Meta-dimensional models involve multiple variables from different data types, but the models for a multistaged analysis are constructed using two different data scales at a time, in a stepwise or linear manner. The main objective of this method is to divide the analysis into multiple steps, where the associations are initially determined between the different data types and subsequently between the data type and trait phenotype of interest (Ritchie *et al.*, 2015). Developing powerful statistical methods will bridge the gap between our ability to generate vast amounts of data and our biological understanding to uncover the genetic basis that underlies the large datasets. A systems genomics approach can provide a more thorough and informative interrogation of genotype-to-phenotype associations.

Acknowledgements

This work was partially supported by natural science foundation of China (31571751) and the Natural Science Foundation of Hubei province, China (2015CFA006).

References

Abe, A., Kosugi, S., Yoshida, K., Natsume, S., Takagi, H., Kanzaki, H., Matsumura, H., Yoshida, K., Mitsuoka, C., Tamiru, M., *et al.* (2012). Genome sequencing reveals agronomically important loci in rice using MutMap. Nat. Biotechnol. *30*, 174–178. http://dx.doi.org/10.1038/nbt.2095

An, G., Lee, S., Kim, S.H., and Kim, S.R. (2005). Molecular genetics using T-DNA in rice. Plant Cell Physiol. *46*, 14–22. https://dx.doi.org/10.1093/pcp/pci502

An, S., Park, S., Jeong, D.H., Lee, D.Y., Kang, H.G., Yu, J.H., Hur, J., Kim, S.R., Kim, Y.H., Lee, M., *et al.* (2003). Generation and analysis of end sequence database for T-DNA tagging lines in rice. Plant Physiol. *133*, 2040–2047. http://dx.doi.org/10.1104/pp.103.030478

Antony, G., Zhou, J., Huang, S., Li, T., Liu, B., White, F., and Yang, B. (2010). Rice *xa13* recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene Os-*11N3*. Plant Cell *22*, 3864–3876. http://dx.doi.org/10.1105/tpc.110.078964

Araki, M., and Ishii, T. (2015). Towards social acceptance of plant breeding by genome editing. Trends Plant Sci. *20*, 145–149. http://dx.doi.org/10.1016/j.tplants.2015.01.010

Ashikari, M., Sakakibara, H., Lin, S., Yamamoto, T., Takashi, T., Nishimura, A., Angeles, E.R., Qian, Q., Kitano, H., and Matsuoka, M. (2005). *Cytokinin oxidase* regulates rice grain production. Science *309*, 741–745. http://dx.doi.org/10.1126/science.1113373

Bai, X.F., Zhao, H., Huang, Y., Xie, W.B., Han, Z.M., Zhang, B., Guo, Z.L., Yang, L., Dong, H.J., Xue, W.Y., *et al.* (2016). Genome-wide association analysis reveals different genetic control in panicle architecture between *indica* and *japonica* Rice. Plant Genome *9*, 2. https://dx.doi.org/10.3835/plantgenome2015.11.0115

Bandillo, N., Raghavan, C., Muyco, P.A., Sevilla, M.A., Lobina, I.T., Dilla-Ermita, C.J., Tung, C.W., McCouch, S., Thomson, M., Mauleon, R., *et al.* (2013). Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding. Rice *6*, 11. http://dx.doi.org/10.1186/1939-8433-6-11

Barakat, A., Gallois, P., Raynal, M., Mestre-Ortega, D., Sallaud, C., Guiderdoni, E., Delseny, M., and Bernardi, G. (2000). The distribution of T-DNA in the genomes of transgenic *Arabidopsis* and rice. FEBS Lett. *471*, 161–164. http://dx.doi.org/10.1016/S0014-5793(00)01393-4

Barrangou, R. (2012). RNA-mediated programmable DNA cleavage. Nat. Biotechnol. *30*, 836–838. http://dx.doi.org/10.1038/nbt.2357

Bibikova, M., Beumer, K., Trautman, J.K., and Carroll, D. (2003). Enhancing gene targeting with designed zinc finger nucleases. Science *300*, 764. http://dx.doi.org/10.1126/science.1079512

Bogdanove, A.J., and Voytas, D.F. (2011). TAL effectors: customizable proteins for DNA targeting. Science *333*, 1843–1846. http://dx.doi.org/10.1126/science.1204094

Chen, H., He, H., Zhou, F., Yu, H., and Deng, X.W. (2013). Development of genomics-based genotyping platforms and their applications in rice breeding. Curr. Opin. Plant Biol. *16*, 247–254. http://dx.doi.org/10.1016/j.pbi.2013.04.002

Chen, H., He, H., Zou, Y., Chen, W., Yu, R., Liu, X., Yang, Y., Gao, Y.M., Xu, J.L., Fan, L.M., *et al.* (2011). Development and application of a set of breeder-friendly SNP markers for genetic analyses and molecular breeding of rice (*Oryza sativa* L.). Theor. Appl. Genet. *123*, 869–879. http://dx.doi.org/10.1007/s00122-011-1633-5

Chen, H., Xie, W., He, H., Yu, H., Chen, W., Li, J., Yu, R., Yao, Y., Zhang, W., He, Y., *et al.* (2014). A high-density SNP genotyping array for rice biology and molecular breeding. Mol. Plant *7*, 541–553. http://dx.doi.org/10.1093/mp/sst135

Chen, L., Huang, Y., Xu, M., Cheng, Z., Zhang, D., and Zheng, J. (2016). iTRAQ-based quantitative proteomics analysis of black rice grain development reveals metabolic pathways associated with anthocyanin biosynthesis. PLOS ONE *11*, e0159238. http://dx.doi.org/10.1371/journal.pone.0159238

Chen, S., Jin, W., Wang, M., Zhang, F., Zhou, J., Jia, Q., Wu, Y., Liu, F., and Wu, P. (2003). Distribution and characterization of over 1000 T-DNA tags in rice genome. Plant J. *36*, 105–113. http://dx.doi.org/10.1046/j.1365-313X.2003.01860.x

Chen, W., Gao, Y., Xie, W., Gong, L., Lu, K., Wang, W., Li, Y., Liu, X., Zhang, H., Dong, H., *et al.* (2014). Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. Nat. Genet. *46*, 714–721. http://dx.doi.org/10.1038/ng.3007

Chu, Z., Fu, B., Yang, H., Xu, C., Li, Z., Sanchez, A., Park, Y.J., Bennetzen, J.L., Zhang, Q., and Wang, S. (2006). Targeting *xa13*, a recessive gene for bacterial blight resistance in rice. Theor. Appl. Genet. *112*, 455–461. http://dx.doi.org/10.1007/s00122-005-0145-6

Chuang, C.F., and Meyerowitz, E.M. (2000). Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. U.S.A. *97*, 4985–4990. http://dx.doi.org/10.1073/pnas.060034297

Comai, L., and Henikoff, S. (2006). TILLING: practical single-nucleotide mutation discovery. Plant J. *45*, 684–694. http://dx.doi.org/10.1111/j.1365-313X.2006.02670.x

Crowell, S., Korniliev, P., Falcão, A., Ismail, A., Gregorio, G., Mezey, J., and McCouch, S. (2016). Genome-wide association and high-resolution phenotyping link *Oryza sativa* panicle traits to numerous trait-specific QTL clusters. Nat. Commun. *7*, 10527. http://dx.doi.org/10.1038/ncomms10527

Das, K., Panda, B.B., Sekhar, S., Kariali, E., Mohapatra, P.K., and Shaw, B.P. (2016). Comparative proteomics of the superior and inferior spikelets at the early grain filling stage in rice cultivars contrast for panicle compactness and ethylene evolution. J. Plant Physiol. *202*, 65–74. http://dx.doi.org/10.1016/j.jplph.2016.07.008

Deng, Y., Zhu, X., Shen, Y., and He, Z. (2006). Genetic characterization and fine mapping of the blast resistance locus *Pigm*(*t*) tightly linked to *Pi2* and *Pi9* in a broad-spectrum resistant Chinese variety. Theor. Appl. Genet. *113*, 705–713. http://dx.doi.org/10.1007/s00122-006-0338-7

Enoki, H., Izawa, T., Kawahara, M., Komatsu, M., Koh, S., Kyozuka, J., and Shimamoto, K. (1999). Ac as a tool for

the functional genomics of rice. Plant J. *19*, 605–613. http://dx.doi.org/10.1046/j.1365-313X.1999.00549.x

Famoso, A.N., Zhao, K., Clark, R.T., Tung, C.W., Wright, M.H., Bustamante, C., Kochian, L.V., and McCouch, S.R. (2011). Genetic architecture of aluminum tolerance in rice (*Oryza sativa*) determined through genome-wide association analysis and QTL mapping. PLOS Genet. *7*, e1002221. http://dx.doi.org/10.1371/journal.pgen.1002221

Fan, C., Xing, Y., Mao, H., Lu, T., Han, B., Xu, C., Li, X., and Zhang, Q. (2006). *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. Theor. Appl. Genet. *112*, 1164–1171. http://dx.doi.org/10.1007/s00122-006-0218-1

Fan, C., Yu, S., Wang, C., and Xing, Y. (2009). A causal C-A mutation in the second exon of *GS3* highly associated with rice grain length and validated as a functional marker. Theor. Appl. Genet. *118*, 465–472. http://dx.doi.org/10.1007/s00122-008-0913-1

Fekih, R., Takagi, H., Tamiru, M., Abe, A., Natsume, S., Yaegashi, H., Sharma, S., Sharma, S., Kanzaki, H., Matsumura, H., *et al.* (2013). MutMap+: genetic mapping and mutant identification without crossing in rice. PLOS ONE *8*, e68529. http://dx.doi.org/10.1371/journal.pone.0068529

Feng, Q., Zhang, Y., Hao, P., Wang, S., Fu, G., Huang, Y., Li, Y., Zhu, J., Liu, Y., Hu, X., *et al.* (2002). Sequence and analysis of rice chromosome 4. Nature *420*, 316–320. http://dx.doi.org/10.1038/nature01183

Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D.L., Wei, P., Cao, F., Zhu, S., Zhang, F., Mao, Y., *et al.* (2013). Efficient genome editing in plants using a CRISPR/Cas system. Cell Res. *23*, 1229–1232. http://dx.doi.org/10.1038/cr.2013.114

Fujita, D., Trijatmiko, K.R., Tagle, A.G., Sapasap, M.V., Koide, Y., Sasaki, K., Tsakirpaloglou, N., Gannaban, R.B., Nishimura, T., Yanagihara, S., *et al.* (2013). *NAL1* allele from a rice landrace greatly increases yield in modern *indica* cultivars. Proc. Natl. Acad. Sci. U.S.A. *110*, 20431–20436. http://dx.doi.org/10.1073/pnas.1310790110

Godfray, H.C., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M., and Toulmin, C. (2010). Food security: the challenge of feeding 9 billion people. Science *327*, 812–818. http://dx.doi.org/10.1126/science.1185383

Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., *et al.* (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science *296*, 92–100. http://dx.doi.org/10.1126/science.1068275

Greco, R., Ouwerkerk, P.B., De Kam, R.J., Sallaud, C., Favalli, C., Colombo, L., Guiderdoni, E., Meijer, A.H., Hoge Dagger, J.H., and Pereira, A. (2003). Transpositional behaviour of an *Ac*/*Ds* system for reverse genetics in rice. Theor. Appl. Genet. *108*, 10–24. http://dx.doi.org/10.1007/s00122-003-1416-8

Greco, R., Ouwerkerk, P.B., Taal, A.J., Favalli, C., Beguiristain, T., Puigdomènech, P., Colombo, L., Hoge, J.H., and Pereira, A. (2001). Early and multiple *Ac* transpositions in rice suitable for efficient insertional mutagenesis. Plant Mol. Biol. *46*, 215–227. http://dx.doi.org/10.1023/A:1010607318694

Guilinger, J.P., Thompson, D.B., and Liu, D.R. (2014). Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nat. Biotechnol. *32*, 577–582. http://dx.doi.org/10.1038/nbt.2909

Han, Z., Zhang, B., Zhao, H., Ayaad, M., and Xing, Y. (2016). Genome-wide association studies reveal that diverse heading date genes respond to short and long day lengths between *indica* and *japonica* rice. Front. Plant Sci. *7*, 1270. http://dx.doi.org/10.3389/fpls.2016.01270

Hartung, F., and Schiemann, J. (2014). Precise plant breeding using new genome editing techniques: opportunities, safety and regulation in the EU. Plant J. *78*, 742–752. http://dx.doi.org/10.1111/tpj.12413

Heffner, E.L., Sorrells, M.E., and Jannink, J.L. (2009). Genomic Selection for Crop Improvement. Crop Sci. *49*, 1–12. http://dx.doi.org/10.2135/cropsci2008.08.0512

Hirochika, H., Sugimoto, K., Otsuki, Y., Tsugawa, H., and Kanda, M. (1996). Retrotransposons of rice involved in mutations induced by tissue culture. Proc. Natl. Acad. Sci. U.S.A. *93*, 7783–7788. http://dx.doi.org/10.1073/pnas.93.15.7783

Hu, W., Wen, M., Han, Z., Tan, C., and Xing, Y. (2013). Scanning QTLs for grain shape using a whole genome SNP array in rice. J. Plant Biochem. Physiol. *1*, 104. http://dx.doi.org/10.4172/2329-9029.1000104

Hua, L., Wu, J., Chen, C., Wu, W., He, X., Lin, F., Wang, L., Ashikawa, I., Matsumoto, T., Wang, L., *et al.* (2012). The isolation of *Pi1*, an allele at the *Pik* locus which confers broad spectrum resistance to rice blast. Theor. Appl. Genet. *125*, 1047–1055. http://dx.doi.org/10.1007/s00122-012-1894-7

Huang, X., Kurata, N., Wei, X., Wang, Z.X., Wang, A., Zhao, Q., Zhao, Y., Liu, K., Lu, H., Li, W., *et al.* (2012). A map of rice genome variation reveals the origin of cultivated rice. Nature *490*, 497–501. http://dx.doi.org/10.1038/nature11532

Huang, X., Qian, Q., Liu, Z., Sun, H., He, S., Luo, D., Xia, G., Chu, C., Li, J., and Fu, X. (2009). Natural variation at the *DEP1* locus enhances grain yield in rice. Nat. Genet. *41*, 494–497. http://dx.doi.org/10.1038/ng.352

Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., Li, C., Zhu, C., Lu, T., Zhang, Z., *et al.* (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. Nat. Genet. *42*, 961–967. http://dx.doi.org/10.1038/ng.695

Huang, X., Yang, S., Gong, J., Zhao, Q., Feng, Q., Zhan, Q., Zhao, Y., Li, W., Cheng, B., Xia, J., *et al.* (2016). Genomic architecture of heterosis for yield traits in rice. Nature *537*, 629–633. http://dx.doi.org/10.1038/nature19760

Huang, X., Yang, S., Gong, J., Zhao, Y., Feng, Q., Gong, H., Li, W., Zhan, Q., Cheng, B., Xia, J., *et al.* (2015). Genomic analysis of hybrid rice varieties reveals numerous superior alleles that contribute to heterosis. Nat. Commun. *6*, 6258. http://dx.doi.org/10.1038/ncomms7258

Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., Li, W., Guo, Y., Deng, L., Zhu, C., *et al.* (2011). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. Nat. Genet. *44*, 32–39. http://dx.doi.org/10.1038/ng.1018

Ishimaru, K., Hirotsu, N., Madoka, Y., Murakami, N., Hara, N., Onodera, H., Kashiwagi, T., Ujiie, K., Shimizu, B., Onishi, A., *et al.* (2013). Loss of function of the IAA-glucose hydrolase gene *TGW6* enhances rice grain weight and increases yield. Nat. Genet. *45*, 707–711. http://dx.doi.org/10.1038/ng.2612

Iyer-Pascuzzi, A.S., Jiang, H., Huang, L., and McCouch, S.R. (2008). Genetic and functional characterization of the rice bacterial blight disease resistance gene *xa5*. Phytopathology *98*, 289–295. http://dx.doi.org/10.1094/PHYTO-98-3-0289

Iyer, A.S., and McCouch, S.R. (2004). The rice bacterial blight resistance gene *xa5* encodes a novel form of disease resistance. Mol. Plant Microbe Interact. *17*, 1348–1354. http://dx.doi.org/10.1094/MPMI.2004.17.12.1348

Izawa, T. (2007). Adaptation of flowering-time by natural and artificial selection in *Arabidopsis* and rice. J. Exp. Bot. *58*, 3091–3097. http://dx.doi.org/10.1093/jxb/erm159

Izawa, T., Mihara, M., Suzuki, Y., Gupta, M., Itoh, H., Nagano, A.J., Motoyama, R., Sawada, Y., Yano, M., Hirai, M.Y., *et al.* (2011). Os-*GIGANTEA* confers robust diurnal rhythms on the global transcriptome of rice in the field. Plant Cell *23*, 1741–1755. http://dx.doi.org/10.1105/tpc.111.083238

Jeon, J.S., Lee, S., Jung, K.H., Jun, S.H., Jeong, D.H., Lee, J., Kim, C., Jang, S., Yang, K., Nam, J., *et al.* (2000). T-DNA insertional mutagenesis for functional genomics in rice. Plant J. *22*, 561–570. http://dx.doi.org/10.1046/j.1365-313x.2000.00767.x

Jeong, D.H., An, S., Kang, H.G., Moon, S., Han, J.J., Park, S., Lee, H.S., An, K., and An, G. (2002). T-DNA insertional mutagenesis for activation tagging in rice. Plant Physiol. *130*, 1636–1644. http://dx.doi.org/10.1104/pp.014357

Jeong, D.H., An, S., Park, S., Kang, H.G., Park, G.G., Kim, S.R., Sim, J., Kim, Y.O., Kim, M.K., Kim, S.R., *et al.* (2006). Generation of a flanking sequence-tag database for activation-tagging lines in *japonica* rice. Plant J. *45*, 123–132. http://dx.doi.org/10.1111/j.1365-313X.2005.02610.x

Kim, C.M., Piao, H.L., Park, S.J., Chon, N.S., Je, B.I., Sun, B., Park, S.H., Park, J.Y., Lee, E.J., Kim, M.J., *et al.* (2004). Rapid, large-scale generation of *Ds* transposant lines and analysis of the *Ds* insertion sites in rice. Plant J. *39*, 252–263. http://dx.doi.org/10.1111/j.1365-313X.2004.02116.x

Kojima, S., Takahashi, Y., Kobayashi, Y., Monna, L., Sasaki, T., Araki, T., and Yano, M. (2002). *Hd3a*, a rice ortholog of the *Arabidopsis FT* gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. Plant Cell Physiol. *43*, 1096–1105. http://dx.doi.org/10.1093/pcp/pcf156

Kolesnik, T., Szeverenyi, I., Bachmann, D., Kumar, C.S., Jiang, S., Ramamoorthy, R., Cai, M., Ma, Z.G., Sundaresan, V., and Ramachandran, S. (2004). Establishing an efficient *Ac*/*Ds* tagging system in rice: large-scale analysis of *Ds* flanking sequences. Plant J. *37*, 301–314. http://dx.doi.org/10.1046/j.1365-313X.2003.01948.x

Komatsu, M., Maekawa, M., Shimamoto, K., and Kyozuka, J. (2001). The *LAX1* and *FRIZZY PANICLE 2* genes determine the inflorescence architecture of rice by controlling rachis-branch and spikelet development. Dev. Biol. *231*, 364–373. http://dx.doi.org/10.1006/dbio.2000.9988

Kulkarni, K.P., Vishwakarma, C., Sahoo, S.P., Lima, J.M., Nath, M., Dokku, P., Gacche, R.N., Mohapatra, T., Robin, S., Sarla, N., *et al.* (2014). A substitution mutation in *OsCCD7* cosegregates with dwarf and increased tillering phenotype in rice. J. Genet. *93*, 389–401. http://dx.doi.org/10.1007/s12041-014-0389-5

Leung, H., Raghavan, C., Zhou, B., Oliva, R., Choi, I.R., Lacorte, V., Jubay, M.L., Cruz, C.V., Gregorio, G., Singh, R.K., *et al.* (2015). Allele mining and enhanced genetic recombination for rice breeding. Rice *8*, 34. http://dx.doi.org/10.1186/s12284-015-0069-y

Li, J., Meng, X., Zong, Y., Chen, K., Zhang, H., Liu, J., Li, J., and Gao, C. (2016). Gene replacements and insertions in rice by intron targeting using CRISPR-Cas9. Nat. Plants *2*, 16139. http://dx.doi.org/10.1038/nplants.2016.139

Li, M., Li, X., Zhou, Z., Wu, P., Fang, M., Pan, X., Lin, Q., Luo, W., Wu, G., and Li, H. (2016). Reassessment of the Four Yield-related Genes *Gn1a*, *DEP1*, *GS3*, and *IPA1* in Rice Using a CRISPR/Cas9 System. Front. Plant Sci. *7*, 377. http://dx.doi.org/10.3389/fpls.2016.00377

Li, T., Liu, B., Spalding, M.H., Weeks, D.P., and Yang, B. (2012). High-efficiency TALEN-based gene editing produces disease-resistant rice. Nat. Biotechnol. *30*, 390–392. http://dx.doi.org/10.1038/nbt.2199

Li, X., Qian, Q., Fu, Z., Wang, Y., Xiong, G., Zeng, D., Wang, X., Liu, X., Teng, S., Hiroshi, F., *et al.* (2003). Control of tillering in rice. Nature *422*, 618–621. http://dx.doi.org/10.1038/nature01518

Li, Y., Fan, C., Xing, Y., Jiang, Y., Luo, L., Sun, L., Shao, D., Xu, C., Li, X., Xiao, J., *et al.* (2011). Natural variation in *GS5* plays an important role in regulating grain size and yield in rice. Nat. Genet. *43*, 1266–1269. http://dx.doi.org/10.1038/ng.977

Lu, L., Shao, D., Qiu, X., Sun, L., Yan, W., Zhou, X., Yang, L., He, Y., Yu, S., and Xing, Y. (2013). Natural variation and artificial selection in four genes determine grain shape in rice. New Phytol. *200*, 1269–1280. http://dx.doi.org/10.1111/nph.12430

Lusser, M., and Davies, H.V. (2013). Comparative regulatory approaches for groups of new plant breeding techniques. New Biotechnol. *30*, 437–446. http://dx.doi.org/10.1016/j.nbt.2013.02.004

Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., Wang, B., Yang, Z., Li, H., Lin, Y., *et al.* (2015). A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot Plants. Mol. Plant *8*, 1274–1284. http://dx.doi.org/10.1016/j.molp.2015.04.007

Magwa, R.A., Zhao, H., and Xing, Y. (2016a). Genome-wide association mapping revealed a diverse genetic basis of seed dormancy across subpopulations in rice (*Oryza sativa* L.). BMC Genet. *17*, 28. http://dx.doi.org/10.1186/s12863-016-0340-2

Magwa, R.A., Zhao, H., Yao, W., Xie, W., Yang, L., Xing, Y., and Bai, X. (2016b). Genomewide association analysis for awn length linked to the seed shattering gene *qSH1* in rice. J. Genet. *95*, 639–646. http://dx.doi.org/10.1007/s12041-016-0679-1

Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. (2013). CAS9 transcriptional activators for target specificity

screening and paired nickases for cooperative genome engineering. Nat. Biotechnol. *31*, 833–838. http://dx.doi.org/10.1038/nbt.2675

Mao, H., Sun, S., Yao, J., Wang, C., Yu, S., Xu, C., Li, X., and Zhang, Q. (2010). Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. Proc. Natl. Acad. Sci. U.S.A. *107*, 19579–19584. http://dx.doi.org/10.1073/pnas.1014419107

Mather, K.A., Caicedo, A.L., Polato, N.R., Olsen, K.M., McCouch, S., and Purugganan, M.D. (2007). The extent of linkage disequilibrium in rice (*Oryza sativa* L.). Genetics *177*, 2223–2232. http://dx.doi.org/10.1534/genetics.107.079616

McCouch, S.R., Wright, M.H., Tung, C.W., Maron, L.G., McNally, K.L., Fitzgerald, M., Singh, N., DeClerck, G., Agosto-Perez, F., Korniliev, P., *et al.* (2016). Open access resources for genome-wide association mapping in rice. Nat. Commun. *7*, 10532. http://dx.doi.org/10.1038/ncomms10532

McCouch, S.R., Zhao, K.Y., Wright, M., Tung, C.W., Ebana, K., Thomson, M., Reynolds, A., Wang, D., DeClerck, G., Ali, M.L., *et al.* (2010). Development of genome-wide SNP assays for rice. Breeding Science *60*, 524–535. http://dx.doi.org/10.1270/jsbbs.60.524

McNally, K.L., Childs, K.L., Bohnert, R., Davidson, R.M., Zhao, K., Ulat, V.J., Zeller, G., Clark, R.M., Hoen, D.R., Bureau, T.E., *et al.* (2009). Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. Proc. Natl. Acad. Sci. U.S.A. *106*, 12273–12278. http://dx.doi.org/10.1073/pnas.0900992106

Meng, L.J., Guo, L.B., Ponce, K., Zhao, X.Q., and Ye, G.Y. (2016a). Characterization of Three Indica Rice Multiparent Advanced Generation Intercross (MAGIC) Populations for Quantitative Trait Loci Identification. Plant Genome *9*, 0. http://dx.doi.org/10.3835/plantgenome2015.10.0109

Meng, L.J., Zhao, X.Q., Ponce, K., Ye, G.Y., and Leung, H. (2016b). QTL mapping for agronomic traits using multi-parent advanced generation inter-cross (MAGIC) populations derived from diverse elite *indica* rice lines. Field Crops Res. *189*, 19–42. http://dx.doi.org/10.1016/j.fcr.2016.02.004

Meyer, R.S., and Purugganan, M.D. (2013). Evolution of crop species: genetics of domestication and diversification. Nat. Rev. Genet. *14*, 840–852. http://dx.doi.org/10.1038/nrg3605

Miura, K., Ikeda, M., Matsubara, A., Song, X.J., Ito, M., Asano, K., Matsuoka, M., Kitano, H., and Ashikari, M. (2010). *OsSPL14* promotes panicle branching and higher grain productivity in rice. Nat. Genet. *42*, 545–549. http://dx.doi.org/10.1038/ng.592

Miyao, A., Tanaka, K., Murata, K., Sawaki, H., Takeda, S., Abe, K., Shinozuka, Y., Onosato, K., and Hirochika, H. (2003). Target site specificity of the *Tos17* retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome. Plant Cell *15*, 1771–1780. http://dx.doi.org/10.1105/tpc.012559

Nakaya, A., and Isobe, S.N. (2012). Will genomic selection be a practical method for plant breeding? Ann. Bot. *110*, 1303–1316. http://dx.doi.org/10.1093/aob/mcs109

Pauwels, K., Podevin, N., Breyer, D., Carroll, D., and Herman, P. (2014). Engineering nucleases for gene targeting: safety and regulatory considerations. New Biotechnol. *31*, 18–27. http://dx.doi.org/10.1016/j.nbt.2013.07.001

Pennisi, E. (2013). The CRISPR craze. Science *341*, 833–836. http://dx.doi.org/10.1126/science.341.6148.833

Ritchie, M.D., Holzinger, E.R., Li, R., Pendergrass, S.A., and Kim, D. (2015). Methods of integrating data to uncover genotype-phenotype interactions. Nat. Rev. Genet. *16*, 85–97. http://dx.doi.org/10.1038/nrg3868

Sallaud, C., Meynard, D., van Boxtel, J., Gay, C., Bès, M., Brizard, J.P., Larmande, P., Ortega, D., Raynal, M., Portefaix, M., *et al.* (2003). Highly efficient production and characterization of T-DNA plants for rice (*Oryza sativa* L.) functional genomics. Theor. Appl. Genet. *106*, 1396–1408. http://dx.doi.org/10.1007/s00122-002-1184-x

Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T., Katayose, Y., Wu, J., Niimura, Y., Cheng, Z., Nagamura, Y., *et al.* (2002). The genome sequence and structure of rice chromosome 1. Nature *420*, 312–316. http://dx.doi.org/10.1038/nature01184

Schaart, J.G., van de Wiel, C.C., Lotz, L.A., and Smulders, M.J. (2016). Opportunities for Products of New Plant Breeding Techniques. Trends Plant Sci. *21*, 438–449. http://dx.doi.org/10.1016/j.tplants.2015.11.006

Sha, Y., Li, S., Pei, Z., Luo, L., Tian, Y., and He, C. (2004). Generation and flanking sequence analysis of a rice T-DNA tagged population. Theor. Appl. Genet. *108*, 306–314. http://dx.doi.org/10.1007/s00122-003-1423-9

Shan, Q., Wang, Y., Chen, K., Liang, Z., Li, J., Zhang, Y., Zhang, K., Liu, J., Voytas, D.F., Zheng, X., *et al.* (2013a). Rapid and efficient gene modification in rice and *Brachypodium* using TALENs. Mol. Plant *6*, 1365–1368. http://dx.doi.org/10.1093/mp/sss162

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K., Liu, J., Xi, J.J., Qiu, J.L., *et al.* (2013b). Targeted genome modification of crop plants using a CRISPR-Cas system. Nat. Biotechnol. *31*, 686–688. http://dx.doi.org/10.1038/nbt.2650

Shan, Q., Zhang, Y., Chen, K., Zhang, K., and Gao, C. (2015). Creation of fragrant rice by targeted knockout of the *OsBADH2* gene using TALEN technology. Plant Biotechnol. J. *13*, 791–800. http://dx.doi.org/10.1111/pbi.12312

Song, X.J., Huang, W., Shi, M., Zhu, M.Z., and Lin, H.X. (2007). A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat. Genet. *39*, 623–630. http://dx.doi.org/10.1038/ng2014

Spielmeyer, W., Ellis, M.H., and Chandler, P.M. (2002). Semidwarf (*sd-1*), "green revolution" rice, contains a defective gibberellin 20-oxidase gene. Proc. Natl. Acad. Sci. U.S.A. *99*, 9043–9048. http://dx.doi.org/10.1073/pnas.132266399

Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., Uemura, A., Utsushi, H., Tamiru, M., Takuno, S., *et al.* (2013a). QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. Plant J. *74*, 174–183. http://dx.doi.org/10.1111/tpj.12105

Takagi, H., Tamiru, M., Abe, A., Yoshida, K., Uemura, A., Yaegashi, H., Obara, T., Oikawa, K., Utsushi, H., Kanzaki, E., *et al.* (2015). MutMap accelerates breeding of a salt-tolerant rice cultivar. Nat. Biotechnol. *33*, 445–449. http://dx.doi.org/10.1038/nbt.3188

Takagi, H., Uemura, A., Yaegashi, H., Tamiru, M., Abe, A., Mitsuoka, C., Utsushi, H., Natsume, S., Kanzaki, H., Matsumura, H., *et al.* (2013b). MutMap-Gap: whole-genome resequencing of mutant F2 progeny bulk combined with de novo assembly of gap regions identifies the rice blast resistance gene *Pii*. New Phytol. *200*, 276–283. http://dx.doi.org/10.1111/nph.12369

Tan, C., Han, Z., Yu, H., Zhan, W., Xie, W., Chen, X., Zhao, H., Zhou, F., and Xing, Y. (2013). QTL scanning for rice yield using a whole genome SNP array. J. Genet. Genomics *40*, 629–638. http://dx.doi.org/10.1016/j.jgg.2013.06.009

Verbyla, A.P., George, A.W., Cavanagh, C.R., and Verbyla, K.L. (2014). Whole-genome QTL analysis for MAGIC. Theor. Appl. Genet. *127*, 1753–1770. http://dx.doi.org/10.1007/s00122-014-2337-4

Wang, C., Chen, S., and Yu, S. (2011). Functional markers developed from multiple loci in *GS3* for fine marker-assisted selection of grain length in rice. Theor. Appl. Genet. *122*, 905–913. http://dx.doi.org/10.1007/s00122-010-1497-0

Weng, J., Gu, S., Wan, X., Gao, H., Guo, T., Su, N., Lei, C., Zhang, X., Cheng, Z., Guo, X., *et al.* (2008). Isolation and initial characterization of *GW5*, a major QTL associated with rice grain width and weight. Cell Res. *18*, 1199–1209. http://dx.doi.org/10.1038/cr.2008.307

Wolt, J.D., Wang, K., and Yang, B. (2016). The regulatory status of genome-edited crops. Plant Biotechnol. J. *14*, 510–518. http://dx.doi.org/10.1111/pbi.12444

Wu, C., Li, X., Yuan, W., Chen, G., Kilian, A., Li, J., Xu, C., Li, X., Zhou, D.X., Wang, S., *et al.* (2003). Development of enhancer trap lines for functional analysis of the rice genome. Plant J. *35*, 418–427. http://dx.doi.org/10.1046/j.1365-313X.2003.01808.x

Wu, J.L., Wu, C., Lei, C., Baraoidan, M., Bordeos, A., Madamba, M.R., Ramos-Pamplona, M., Mauleon, R., Portugal, A., Ulat, V.J., *et al.* (2005). Chemical- and irradiation-induced mutants of *indica* rice IR64 for forward and reverse genetics. Plant Mol. Biol. *59*, 85–97. http://dx.doi.org/10.1007/s11103-004-5112-0

Xing, Y., and Zhang, Q. (2010). Genetic and molecular bases of rice yield. Annu. Rev. Plant Biol. *61*, 421–442. http://dx.doi.org/10.1146/annurev-arplant-042809-112209

Xu, X., Liu, X., Ge, S., Jensen, J.D., Hu, F., Li, X., Dong, Y., Gutenkunst, R.N., Fang, L., Huang, L., *et al.* (2011). Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. Nat. Biotechnol. *30*, 105–111. http://dx.doi.org/10.1038/nbt.2050

Xu, Y.B., and Crouch, J.H. (2008). Marker-assisted selection in plant breeding: From publications to practice. Crop Sci. *48*, 391–407. http://dx.doi.org/10.2135/cropsci2007.04.0191

Xue, W., Xing, Y., Weng, X., Zhao, Y., Tang, W., Wang, L., Zhou, H., Yu, S., Xu, C., Li, X., *et al.* (2008). Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. Nat. Genet. *40*, 761–767. http://dx.doi.org/10.1038/ng.143

Yamamoto, T., Nagasaki, H., Yonemaru, J., Ebana, K., Nakajima, M., Shibaya, T., and Yano, M. (2010). Fine definition of the pedigree haplotypes of closely related rice cultivars by means of genome-wide discovery of single-nucleotide polymorphisms. BMC Genomics *11*, 267. http://dx.doi.org/10.1186/1471-2164-11-267

Yan, W., Liu, H., Zhou, X., Li, Q., Zhang, J., Lu, L., Liu, T., Liu, H., Zhang, C., Zhang, Z., *et al.* (2013). Natural variation in *Ghd7.1* plays an important role in grain yield and adaptation in rice. Cell Res. *23*, 969–971. http://dx.doi.org/10.1038/cr.2013.43

Yan, W.H., Wang, P., Chen, H.X., Zhou, H.J., Li, Q.P., Wang, C.R., Ding, Z.H., Zhang, Y.S., Yu, S.B., Xing, Y.Z., *et al.* (2011). A major QTL, *Ghd8*, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. Mol. Plant *4*, 319–330. http://dx.doi.org/10.1093/mp/ssq070

Yang, W., Guo, Z., Huang, C., Duan, L., Chen, G., Jiang, N., Fang, W., Feng, H., Xie, W., Lian, X., *et al.* (2014). Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. Nat. Commun. *5*, 5087. http://dx.doi.org/10.1038/ncomms6087

Yang, W., Guo, Z., Huang, C., Wang, K., Jiang, N., Feng, H., Chen, G., Liu, Q., and Xiong, L. (2015). Genome-wide association study of rice (*Oryza sativa* L.) leaf traits with a high-throughput leaf scorer. J. Exp. Bot. *66*, 5605–5615. http://dx.doi.org/10.1093/jxb/erv100

Yano, M., Katayose, Y., Ashikari, M., Yamanouchi, U., Monna, L., Fuse, T., Baba, T., Yamamoto, K., Umehara, Y., Nagamura, Y., *et al.* (2000). *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. Plant Cell *12*, 2473–2484. http://dx.doi.org/10.1105/tpc.12.12.2473

Yin, Z., and Wang, G. (2000). Evidence of multiple complex patterns of T-DNA integration into the rice genome. Theor. Appl. Genet. *100*, 461–470. http://dx.doi.org/10.1007/s001220050060

Yu, B., Lin, Z., Li, H., Li, X., Li, J., Wang, Y., Zhang, X., Zhu, Z., Zhai, W., Wang, X., *et al.* (2007). *TAC1*, a major quantitative trait locus controlling tiller angle in rice. Plant J. *52*, 891–898. http://dx.doi.org/10.1111/j.1365-313X.2007.03284.x

Yu, H., Xie, W., Li, J., Zhou, F., and Zhang, Q. (2014). A whole-genome SNP array (RICE6K) for genomic breeding in rice. Plant Biotechnol. J. *12*, 28–37. http://dx.doi.org/10.1111/pbi.12113

Yu, J., Hu, S., Wang, J., Wong, G.K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., *et al.* (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science *296*, 79–92. http://dx.doi.org/10.1126/science.1068037

Yuan, M., Chu, Z., Li, X., Xu, C., and Wang, S. (2009). Pathogen-induced expressional loss of function is the key factor in race-specific bacterial resistance conferred by a recessive *R* gene *xa13* in rice. Plant Cell Physiol. *50*, 947–955. http://dx.doi.org/10.1093/pcp/pcp046

Zhang, J., Zhou, X., Yan, W., Zhang, Z., Lu, L., Han, Z., Zhao, H., Liu, H., Song, P., Hu, Y., *et al.* (2015). Combinations of the *Ghd7*, *Ghd8* and *Hd1* genes largely define the ecogeographical adaptation and yield potential of

cultivated rice. New Phytol. *208*, 1056–1066. http://dx.doi.org/10.1111/nph.13538

Zhang, Z., Ersoz, E., Lai, C.Q., Todhunter, R.J., Tiwari, H.K., Gore, M.A., Bradbury, P.J., Yu, J., Arnett, D.K., Ordovas, J.M., *et al.* (2010). Mixed linear model approach adapted for genome-wide association studies. Nat. Genet. *42*, 355–360. http://dx.doi.org/10.1038/ng.546

Zhao, K., Tung, C.W., Eizenga, G.C., Wright, M.H., Ali, M.L., Price, A.H., Norton, G.J., Islam, M.R., Reynolds, A., Mezey, J., *et al.* (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. Nat. Commun. *2*, 467. http://dx.doi.org/10.1038/ncomms1467

Zhao, K., Wright, M., Kimball, J., Eizenga, G., McClung, A., Kovach, M., Tyagi, W., Ali, M.L., Tung, C.W., Reynolds, A., *et al.* (2010). Genomic diversity and introgression in *O. sativa* reveal the impact of domestication and breeding on the rice genome. PLOS ONE *5*, e10780. http://dx.doi.org/10.1371/journal.pone.0010780

Zhou, B., Qu, S., Liu, G., Dolan, M., Sakai, H., Lu, G., Bellizzi, M., and Wang, G.L. (2006). The eight amino-acid differences within three leucine-rich repeats between *Pi2* and *Piz-t* resistance proteins determine the resistance specificity to *Magnaporthe grisea*. Mol. Plant Microbe In. *19*, 1216–1228. http://dx.doi.org/10.1094/MPMI-19-1216