



Updates and perspectives on the utilization of molecular makers of complex traits in rice

G. Li¹, S.W. Kwon^{1,2} and Y.J. Park^{1,2}

¹Department of Plant Resources, College of Industrial Science,
Kongju National University, Yesan, Republic of Korea

²Legume Bio-Resource Center of Green Manure,
Kongju National University, Yesan, Republic of Korea

Corresponding author: Y.J. Park

E-mail: yjpark@kongju.ac.kr

Genet. Mol. Res. 11 (4): 4157-4168 (2012)

Received May 19, 2012

Accepted August 10, 2012

Published September 10, 2012

DOI <http://dx.doi.org/10.4238/2012.September.10.4>

ABSTRACT. After complete sequencing of its genome and annotation of the majority of its ~32,000 genes, rice genome has become the model genome among the cereal genomes, and the focus has shifted from structural to functional genomics and application of genomic-derived information in rice breeding. During the past 2 decades, intensive worldwide efforts have led to significant improvements in rice. An abundance of molecular markers and information related to many genes/quantitative trait loci that control agronomically important traits such as yield, quality, and biotic and abiotic stress tolerance have been identified. Bridging the application gap between quantitative trait locus identification and marker-assisted selection breeding is an urgent, arduous, and long-term task. Marker development, allele mining, gene discovery, and molecular breeding have progressed to a great extent because of the rapid development of next-generation sequencing, large-scale high-density genotyping, and genome-wide selection strategies. The availability of high-density markers and the rapidly decreasing cost

of genotyping have facilitated marker-assisted selection of many traits that were previously not possible.

Key words: Quantitative trait locus; Marker-assisted selection; Rice; Next-generation sequencing; Whole-genome resequencing

INTRODUCTION

Rice (*Oryza sativa*) is not only the most important food crop, but also an ideal model organism for the investigations of other cereal crops. Most domesticated rice cultivars belong to 2 major subspecies, *O. sativa* ssp *japonica* and *O. s.* ssp *indica*, which differ in physiological and morphological traits (Londo et al., 2006). As the model cereal species, rice was the second plant species to be sequenced (after the dicotyledon *Arabidopsis thaliana*) because of its moderately complex genome and genetic relationship to other major cereal crops. The finished-quality sequence based on the *japonica* cultivar Nipponbare shows that the rice genome is 389 Mb (International Rice Genome Sequencing Project, IRGSP, 2005), and encodes ~32,000 genes (The Rice Annotation Project 2007, 2008). This is one-sixth the size of the maize genome and 40 times smaller than that of the wheat genome (IRGSP, 2005), making rice an excellent model for the study of cereal genomes. Comparison of the *indica* and *japonica* genomes has been selected as a model for understanding the origin, speciation, domestication, and genome evolution of rice (Han et al., 2007).

Sequencing and analysis of the rice genome has been conducted through international collaborative efforts. From the earliest Rice Genome Project (started in 1991 in Japan) to the later IRGSP (established in 1997), and the current whole-genome resequencing effort, rice genome sequencing has developed rapidly. Genome sequence drafts were completed for both the subspecies in 2002 by the IRGSP and the Beijing Genome Institute (BGI). Information from the draft sequence was ultimately integrated into the IRGSP effort, and an annotated “gold standard”, high-quality, more accurate japonica map was published in 2005 (IRGSP, 2005), which represents a landmark achievement in biological research. The road towards completion of a high-quality sequence of the entire genome not only facilitates functional gene studies, but also provides valuable experience and references for studying other crop species.

Many of the tools that are useful for rice genomics were developed from large-scale rice genome projects, and these tools greatly facilitated the analysis of the important agronomic traits of rice. However, obtaining pertinent data from this sea of biological information is problematic and should be addressed. As rice genomic data continue to accumulate rapidly, databases are becoming more valuable for storage of and access to large and rigorous datasets. More than 20 rice genomics databases and websites are available online. The US National Center for Biotechnology Information (NCBI; <http://www.ncbi.nlm.nih.gov/>) provides comprehensive coverage of genomics information, including genetic and physical maps, DNA markers, expressed sequence tags (EST), and morphological markers. Additional specialized services are provided by other databases, such as Gramene (<http://www.gramene.org/>), a widely used open-source data resource for comparative genome analysis in grasses that use the rice genome as an anchor. It provides information about EST sequencing, protein structure, functional analysis, genetic and physical mapping, interpretation of biochemical pathways, gene and quantitative trait locus (QTL) localization, and phenotypic characteristics, and mutations. The integrated rice genome explorer (<http://rgp.dna.affrc.go.jp/giot/>

INE.html) is another excellent database that provides quick and simple correlations between genetic markers and ESTs, and physical maps of the rice genome sequence. Moreover, a user-friendly “genome browser” (Gbrowse; <http://www.tigr.org/tdb/e2k1/osa1/>) developed by the Institute for Genomics Research consists of annotated genes, identified domains within predicted genes, and identified related sequences between rice and other cereals, which allow users to view specific features such as markers and putative genes within defined regions. Rice genome annotation can also be obtained from the Rice Annotation Project database (<http://rapdb.dna.affrc.go.jp/>), which has been incorporated into other databases such as the NCBI map viewer, the DNA Data Bank of Japan, and the European Molecular Biology Laboratory, thereby integrating a curated gene set and more than 70 annotation tracks. Most of these databases provide specialized and detailed information for use in a broad range of rice genomics or biology research, whereas some are designed for ordinary users and offer general information on rice. We can make good and effective use of the huge quantity of information stored in these databases to generate useful bioinformatics resources.

The completion of the rice genome, which is recognized as a landmark achievement in plant biological/molecular research, has opened a new chapter in functional studies and has considerable scientific and economic/social value. Combined with the development of novel technical platforms based on whole-genome sequence data, a systematic approach to characterization and application of these genes facilitated the analysis of important agronomic characteristics. Recent years have seen great progress in the elucidation of genetic mechanisms and evolutionary selection processes involved in complex traits such as yield, grain quality, biotic and abiotic stresses, reproductive barriers, epigenetics, and flowering time. The application of next-generation sequencing (NGS) strategies and whole-genome resequencing platforms will facilitate more complex analyses of genome-wide function, regulation, and cooperation. These resources will also facilitate the identification of all genes and proteins encoded by the DNA sequence, enable the full use of genetic germplasm resources, and identification of causative genes and mutant variations, and help accomplish our ultimate breeding objective: high yield and quality.

Herein, we provide an overview of the available rice genome information resources and their role in elucidating genomic information and molecular marker-assisted breeding, particularly when combined with whole-genome resequencing.

QTL IDENTIFICATION, CLONING, AND APPLICATION IN RICE BREEDING

Whole-genome sequencing of 2 rice subspecies - *indica* and *japonica* - have provided a vast quantity of information regarding the rice genome and facilitated detailed analyses of genome structure, function, and evolution. The rice genome characteristics described above provide important clues for the development of strategies to maximize utilization of the genome sequence (Sasaki, 2003). One of the first outcomes of genome sequence-based marker development was an increase in the map-based cloning of QTLs underlying agronomically important rice traits. This information has enabled the isolation of many important QTL genes that have the potential to greatly improve rice production. The identification of QTLs controlling agricultural traits is vital to increase yield and meet the needs of the growing human population, although progress has encountered challenges due to complexities in QTL inheritance. To date, many QTLs have been detected in rice, including those responsible for yield and grain quality (salt, drought and submergence tolerance), disease and insect resistance and nutrient utilization efficiency. Map-based cloning techniques have enabled scientists to successfully

fine map and clone approximately 17 QTLs for several traits. Additional in-depth functional analyses and characterization of these genes will facilitate rice molecular breeding. We integrated 494 QTLs, including those responsible for yield, grain quality, biotic, and abiotic stress tolerance (Figure 1), and investigated their applications in crop breeding. A large number of these QTLs have been mapped to putative genomic regions, while few have been fine mapped or cloned. Researchers have attributed this to several factors, including insufficient population size, unstable statistical thresholds for detecting putative loci, use of a minimal number of molecular markers in analyses, and the low heritability of target traits (Yano and Sasaki, 1997).

Genetic mapping

Many important agronomic traits, including grain yield, quality, and stress tolerance, show continuous phenotypic variation, which is usually governed by a number of QTL genes. For QTL identification, mapping populations such as F2 plants, recombinant inbred lines (RILs), backcross inbred lines, or double-haploid lines must be prepared (Hao and Lin, 2010), and a linkage map based on dozens of molecular markers must be generated. Meanwhile, to resolve the problem of minor QTL effects in QTL analyses using F2 and RIL, Ando et al. (2008) developed chromosome segment substitution lines.

Isolation of QTLs from mutants by map-based cloning (also called positional cloning) is a powerful and effective tool for revealing the mechanisms regulating target traits (Collins, 1995). Gene expression analysis and transformation system confirmation of candidate genes have facilitated many accomplishments in crop species. However, this technique is resource- and time-intensive due to the use of large separate populations and high-density molecular linkage maps, and the complementary demonstration of mutant functions. Thus, full use of bioinformatics resources combined with a candidate gene strategy will likely double the productivity, particularly for selected genes related to particular pathways.

In some cases, depending on the population used, an association mapping approach might be preferable, since the functional markers minimize the risk of recombination between marker and QTL alleles. Association mapping, based on linkage disequilibrium (LD), involves the direct use of the natural population to map genes associated with a trait of interest on the basis of the “common disease/common variant” hypothesis (Peltonen et al., 2000). Based on candidate gene or QTL analysis, high-density, whole-genome scans are required for efficient association mapping. The increased availability of high-throughput genotyping technologies, together with the development of statistical methodologies, enables the assembly of different sets of genotypes, such as germplasm accessions and cultivated varieties characterized by high or low LD, which can be used for QTL discovery or candidate gene validation (Rafalski, 2010).

Additionally, targeting induced local lesions in genomes (TILLING) or eco-TILLING is a powerful approach for identifying allelic variation in natural populations at target loci and making them available for association mapping (Gilchrist et al., 2006). This approach has proven to be useful for the identification and functional validation of candidate genes for complex traits governing disease resistance (Manosalva et al., 2009). As sequencing costs have declined, large-scale resequencing of germplasm accessions has resulted in the discovery of many alleles for specific genic loci in cereals; functional analysis of these loci is now underway. These validated alleles can be used to develop functional or breeder-ready markers.

Candidate drought-responsive genes, including dehydration-responsive element-binding protein 2a (*DREB2a*), ethylene response factor (*ERF3*), and sucrose synthase, are striking examples of markers that were identified using convergent information based on genome annotation, expression data, and co-localization with QTLs. The identification of associations between haplotype variation and a quantitative trait has been used widely in other crop species since the first report in *Arabidopsis* (Olsen et al., 2004). The integration of QTLs for numerous traits and candidate genes from genomic databases will enhance our understanding of association mapping and facilitate progress towards achieving breeding goals.

Major QTLs involved in agronomic traits

QTLs for grain yield

At present, most rice investigations have addressed grain yield and quality. Grain yield in rice is a complex trait that is multiplicatively determined by 3 typically quantitative component traits - the number of panicles, the number of grains per panicle and grain weight (Xing and Zhang, 2010). However, several other traits can also affect grain yield, including days to heading, plant height, tiller angles, sterility, and grain shattering. Hundreds of QTLs for yield have been identified in different populations during the last decade. The Gramene QTL database includes all QTLs related to rice grain shape that have been identified since 1994 (currently, N = 8646). However, only 7 QTLs (*Gn1a*, *GS3*, *GW2*, *qSW5*, *Ghd7*, *DEP1*, and *OsSPL14*) that control rice yield traits have been cloned and characterized to date. These QTLs, together with *MONOCULMI* (*MOC1*), a key gene controlling rice tiller number (Li et al., 2003) may allow for the elucidation of molecular genetic mechanisms underlying crop-yield traits and provide useful tools for improving crop yield. The molecular characterization of these genes has greatly advanced our mechanistic understanding of the regulation of these rice-yield traits. These findings have significant implications for crop genetic improvement.

Among these 3 component yield traits, the number of QTLs resolved for grain weight is the largest, followed by the number of spikelet per panicle and QTL number for tillers. Analyses based on molecular marker linkage maps showed that significant digenic interactions, including additive by additive, additive by dominance, and dominance by dominance, are frequent and widespread in the genome, contributing significantly to the genetic bases of the yield traits (Xing and Zhang, 2010). In addition, yield levels of rice varieties are also greatly influenced by environmental conditions and field management practices. Remarkable interactions occur between genotypes and environments, such that varieties are adapted to specific environmental conditions.

The number of panicles is dependent on the ability of the plant to produce tillers; thus, the number of tillers must be regulated to maximize the yield potential. Mutant analyses and map-based cloning of QTLs have identified a large number of genes required for the basic processes underlying the initiation and development of tillers and panicles. For example, the tiller-related gene *MOC1* was detected using the mutant *moc1* loss-of-function mutant (Li et al., 2003). Also, the *OsTBI/FINE CULM* (*FC1*) gene is known to negatively regulate rice tillering (Takeda et al., 2003). Grain weight is largely determined by grain size, which is specified by its 3 dimensions (length, width and thickness) and degree of filling. *GS3*, *GW2*, *qSW5*, and the grain-filling regulation gene *GIF1* were mapped to chromosome 3, 2, 5, and 4, respectively.

These QTL mapping and cloning studies have provided useful information regarding the molecular mechanism of genes regulating grain weight for both grain size and filling, although there is still the need for additional explanation obtained through a systematic study on the biological process of initiation, growth and development of grains (Xing and Zhang, 2010).

Heading date, which is controlled by many genes, is a key determinant of rice adaptation to different cultivation areas and cropping seasons. Thus far, many QTLs that contribute to heading date in rice have been detected and cloned using progeny from a *japonica-indica* cross. Here, we integrated 42 heading date QTLs distributed on 7 chromosomes, with the exception of chromosomes 4, 5, 9, 11, and 12. Nine QTLs (*Hd1*, *Hd2*, *Hd3a*, *Hd3b*, *Hd4*, *Hd5*, *Hd6*, *Hd8*, and *Hd9*) were mapped as single Mendelian factors (Lin et al., 2000). These QTLs, based on the previous research, jointly take part in the flowering regulation process. For example, *Hd1* contained a CCT domain with ~60% identity to *Ghd7*, corresponding to a gene homologous with *CONSTANS (CO)* in *Arabidopsis*, *Hd6* encodes casein kinase 2 alpha, and *Hd3a* was similar to an *Arabidopsis* FT-like protein. However, another major QTL, *Early heading date 1 (Ehd1)*, that encodes a B-type response regulator is suppressed under long-day conditions, which resulted in reduced *Hd3a* expression, for which *Ghd7* was responsible (Doi et al., 2004). The relationships of expression in these genes were complicated, which suggested that there was a complex mechanism of gene regulation involved.

QTLs for quality

Cooking and eating quality are the most important components of rice quality, although this characteristic is difficult to define, because it involves objective and subjective criteria. Gelatinization temperature (GT), gel consistency (GC), and amylose content (AC) are the 3 major rice traits related directly to cooking and eating quality (International Rice Research Institute, <http://irri.org/>). Molecular markers linked to key quality traits have been validated in a set of rice varieties with different starch, cooking and eating qualities. The AC, GT, and fragrance traits are known to be largely controlled by the *Wx*, *Starch Synthase IIa (SSIIa)*, and *fragrance (fgr)* genes, respectively. Amylose is considered to be the most important predictor of sensory quality in rice. The granule-bound starch synthase I is required for amylose synthesis, and several alleles are encoded by the *Wx* locus (Mikami et al., 2008). Evaluation in a range of mapping populations has demonstrated that QTLs for GC are consistently associated with the *Wx* locus, providing further support for the hypothesis that more allelic variation at this locus awaits discovery to further explain the sensory properties of grain quality (Wang et al., 2007). *Wx* functions as the sole major gene affecting both AC and GC, but as a minor gene affecting GT, whereas *SSII-3* is the major gene controlling GT but as a minor gene affecting AC and GC, together with other minor genes such *ISA* and *SBE3*, which affect the different characteristics of grain eating and cooking quality at the different starch biosynthesis stages (Tian et al., 2009).

QTLs of biotic/abiotic stress tolerance

Here, we integrated 155 QTLs related to tolerance to abiotic stresses, including cold tolerance, drought tolerance, lodging resistance, salt tolerance and submergence tolerance. Furthermore, 86 QTLs related to biotic stress tolerance were integrated, consisting of resis-

tance to bacterial leaf blight, sheath blight, blast field, and insects (Figure 1). QTLs controlling quantitative disease resistance are a valuable source of broad-spectrum and durable disease resistance. To date, approximately 30 R genes for bacterial blight resistance have been identified, 6 of which (*Xa1*, *Xa3/Xa26*, *xa5*, *xa13*, *xa21*, and *xa27*) have been isolated and characterized (Hu et al., 2008). QTLs for 3 aspects of cold tolerance, namely low-temperature germinability, cold tolerance at the seedling stage and cold tolerance at the booting stage, have been reported to date. The higher degree of crosstalk observed between drought and salt stress than between drought and cold stresses or between cold and salt stress was summarized (Hadiarto and Tran, 2011) based on previous research demonstrating that screening of 1700 rice genes under different stress conditions revealed that 36, 62, 57, and 43 genes were induced by cold, drought, high-salinity, and abscisic acid, respectively (Rabbani et al., 2003).

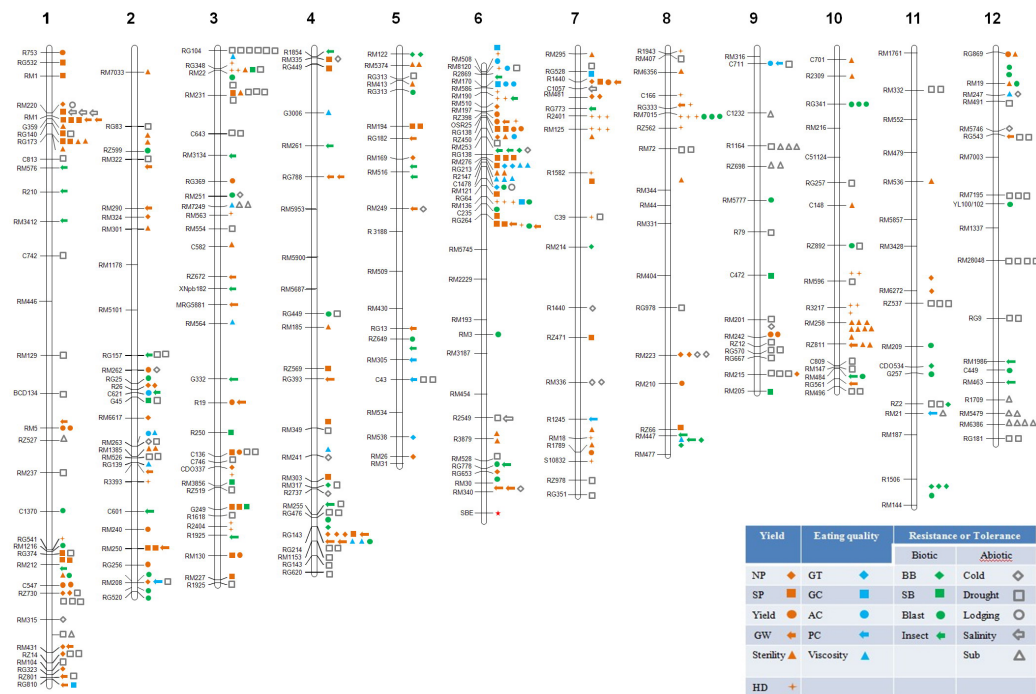


Figure 1. Integrated linkage map showing the positions of yield, eating quality, biotic/abiotic stress resistance QTLs based on the data from the Q-ARO database.

Salinity is a key abiotic stress affecting rice growth and yield. Several QTLs for salt tolerance have been reported to date. *SKC1*, located on chromosome 1, is a major QTL for shooting content and is involved in the regulation of Na⁺/K⁺ homeostasis (Ren et al., 2005). The positional cloning of QTL *Strong Culm2* (*SCM2*), identified by Ookawa et al. (2010), had desirable effects on rice yield, as lodging resistance is an important agronomic trait that should be included in breeding programs to obtain high yields. Submergence stress is another abiotic stress on rice production in southern and southeastern Asia. *Submergence 1* (*Sub1*), a cluster

of 3 genes (*Sub1A*, *Sub1B*, and *Sub1C*) that encodes putative ethylene response factors on chromosome 9, was detected in a cross between a submergence-tolerant *indica* cultivar and a submergence-intolerant *japonica* cultivar (Xu et al., 2006). Among these, *Sub1A* confers submergence tolerance and is strongly expressed during submergence, disrupting the underwater escape strategy, which involves augmentation of the ethylene response.

Crosstalk among drought, salt, and cold stresses can be observed at the protein kinase cascade and transcription activation levels. *OsMPK7*, *OsNAC10* and *OsZIP23* are drought- and high-salinity-inducible, but are unaffected by cold stress (Xiang et al., 2008; Hadiarto and Tran, 2011). Overexpression of cold-inducible *OsDREB1A* or *OsDREB1B* in rice improved not only tolerance to cold stress, but also tolerance to drought and salt stresses (Ito et al., 2006). These findings suggest the existence of crosstalk among drought, salt and cold stresses and provide useful fundamental information for breeding.

The role of QTLs in marker-assisted selection (MAS) breeding

QTL identification is merely the first step in the study of functional genomics. Fine mapping, cloning, and functional characterization of QTLs also facilitate the elucidation of mechanisms underlying specific traits and provide feedback to practical breeding programs. Once genes controlling traits of agronomic interest have been cloned or mapped precisely on their respective chromosomes, this information can be used in a breeding system. The representative methods of MAS include marker-assisted backcrossing, or introgression for major genes, or QTLs with a relatively large effect, marker-assisted recurrent selection for complex traits, and marker-evaluated selection to identify genomic regions during the selection of breeding populations using a bulk-population breeding system in target environments (Steele et al., 2004).

Four QTLs for rice heading date (*Hd6*, *Hd1*, *Hd4*, and *Hd5*) were introgressed from Kasalath into Koshihikari by MAS to enhance the cropping potential in one of the leading cultivars in Japan (Takeuchi et al., 2006). This approach demonstrates the potential power of MAS in rice breeding. A major QTL, *Sub1*, was detected at chromosome 9 and the underlying gene was cloned (Xu et al., 2006); the *Sub1A* allele was then introgressed by MAS into an elite cultivar grown widely in Asia. The resultant lines showed a promising yield and enhanced agronomic traits, as well as tolerance to submergence. To develop a new line with lodging resistance and high yield, 2 genes controlling semi-dwarfing and grain number were successfully introduced into Koshihikari (Ashikari et al., 2005). Four QTLs controlling partial resistance to rice blast in upland rice have been successfully pyramided into lowland rice cultivars by MAS (Fukuoka and Saka, 2006).

ONGOING RICE WHOLE-GENOME RESEQUENCING

In individuals of a known reference genome, whole-genome resequencing is a promising strategy to identify the relationship between sequence variation and normal or mutant phenotypes. High-throughput genome resequencing - if accurate - has the advantage of allowing researchers to identify the specific nucleotides associated with a given phenotype, and allowing the effective detection and analysis of genetic variations important for molecular breeding. Given the rapid improvements in NGS technologies, we expect that they will eventually enable the identification of genetic variation between individuals using whole-genome or target-region resequencing. This platform allows not only processing, analysis and interpretation of the data, but also the develop-

ment of a rapid, cost-effective, informative, and reliable genotyping method at an affordable cost. An important application of NGS is the resequencing of targeted regions for the identification of mutant alleles, and we believe that mapping by sequencing will become a centerpiece in efforts to discover the genes responsible for QTLs. Generally speaking, the availability of a wide range of low- and high-multiplex single nucleotide polymorphism (SNP) assay methods (sequencing accuracy and depth of coverage relies on the experimental design) makes SNPs an ideal marker for QTL mapping, association analysis, MAS, and the construction of high-density genetic maps for fine mapping and cloning of agronomically important genes (McCouch et al., 2010).

SNP discovery by resequencing whole-genome or subgenome representations is often among the first uses of a reference genome sequence. For inbreeding species such as rice, lines to be resequenced are normally purified through 1 or 2 generations of inbreeding (via single seed descent). The DNA sample is resequenced using NGS technology, and SNPs are identified by comparing the sequenced genome with a reference genome. For example, using information on the features of the B73, Gore et al. (2009) targeted the gene fraction of the maize genome for resequencing in the founder inbred lines of the nested association mapping population. Two datasets comprising 3.3 million SNPs were used to produce a first haplotype map (“HapMap”) and to analyze the distribution of recombination and diversity along the maize chromosomes. Additionally, in 2009, 2 rice RILs were accurately genotyped using high-throughput techniques based on whole-genome resequencing, and ultrahigh-density linkage maps were constructed for QTL mapping (Xie et al., 2010). The combination of both technologies and high-throughput SNP detection assays will greatly facilitate the genetic dissection of complex traits by high-resolution QTL mapping and whole-genome association studies (Huang et al., 2009).

A suitable recent example is the construction of a comprehensive HapMap for rice that was used for the genome-wide associate study of 14 agronomic traits, such as heading date and tillering (Huang et al., 2010). The researchers made use of low-coverage (1-fold per rice line) sequence data across lines, for a combined coverage of ~508-fold, and detected 3.6 million SNPs. This manuscript provided a new approach to low-fold sequence coverage, which can be used to detect not only SNPs but also more complex polymorphisms, and partially overcome the need for deeper sequence coverage (Clark, 2010). Additionally, 40 cultivated accessions selected from the major groups of rice and 10 from their wild progenitors (*O. rufipogon* and *O. nivara*) were resequenced to >15X raw data coverage (Xu et al., 2012). After mapping the sequence read back to an IRGSP reference genome, the authors investigated the genome-wide variation pattern in a comparative analysis. The data revealed examples of structural variation in genomes and included 6.5 million high-quality SNPs after excluding sites with missing data in any accession. Using these population SNP data, the authors also identified thousands of new rice genes and tracked down those with a significantly lower diversity in cultivated, but not wild, rice. These variants represent a valuable resource for those interested in improving rice crops.

A research project involving “rice gene resequencing of 3000 germplasm samples” has been launched at BGI in Shenzhen, China. This project, the initiation of genetic rice breeding, effectively combines basic biological research with biological breeding and agricultural industrialization to facilitate the stable, long-term development of the agricultural industry and promote food security. “We can select the gene we want and take it to the files of relevant materials. And it can greatly improve the efficiency of breeding”, said Zhai, the president of the Chinese Academy of Agricultural Science (<http://en.genomics.cn/navigation/index.action>). The data generated by this rice gene resequencing research study are expected to be released at the end of 2012.

A worldwide effort to collect rice accessions for whole-genome resequencing and comprehensive phenotyping is under way, and association results from this broader sampling will be investigated in the future. Genome sequencing is an effective genotyping approach for the improvement of mapping resolution and the identification of new alleles through continuous population expansion, as it allows new SNPs to be added and imputation efficiency to be improved, even at lower sequence coverage. These data represent a valuable resource for rice improvement.

CONCLUSION AND PERSPECTIVES

The past two decades have been a time of great progress in rice genome research. The availability of complete genome sequences, together with various genomic resources available for rice and *Arabidopsis*, now serve as the basis of comparative studies within or among species, and provide a context when overlaid with associated genome annotation. High-precision genome sequences, saturated molecular marker linkage maps, hundreds of genes and QTLs, efficient transformation technology and abundant germplasm resources together with functional genomics, proteomics and metabolomics provide a useful platform for the understanding of genetic and major biological process, including the complex metabolic pathways that control growth and development. The gate has already opened to fast and efficient characterization of genetic variation in all contexts using whole-genome resequencing, as its cost declines and throughput increases. These new approaches allow the reassessment of some fundamental questions in plant biology, and we look forward to the discoveries that will be enabled by functional genomic studies in plants.

The long-term goal of the International Rice Functional Genomics Consortium is to determine the function of all rice genes. Worldwide collaborations are necessary to reach this goal. Zhang et al. (2008) called for an international coordinated rice functional genomic effort in the form of the project 2020. The mission of this project is to determine the function of every gene in the rice genome by the year 2020, to identify functional diversity of alleles for agriculturally useful genes from the primary gene pool of rice and to apply the findings of functional genomic research to genetic improvement of rice. We are without doubt looking forward to this great achievement.

ACKNOWLEDGMENTS

Research supported by a grant from the BioGreen 21 Program (#PJ009039), Rural Development Administration, Republic of Korea.

REFERENCES

- Ando T, Yamamoto T, Shimizu T, Ma XF, et al. (2008). Genetic dissection and pyramiding of quantitative traits for panicle architecture by using chromosomal segment substitution lines in rice. *Theor. Appl. Genet.* 116: 881-890.
- Ashikari M, Sakakibara H, Lin S, Yamamoto T, et al. (2005). Cytokinin oxidase regulates rice grain production. *Science* 309: 741-745.
- Clark RM (2010). Genome-wide association studies coming of age in rice. *Nat. Genet.* 42: 926-927.
- Collins FS (1995). Positional cloning moves from perdictional to traditional. *Nat. Genet.* 9: 347-350.
- Doi K, Izawa T, Fuse T, Yamanouchi U, et al. (2004). Ehd1, a B-type response regulator in rice, confers short-day promotion of flowering and controls FT-like gene expression independently of Hd1. *Genes Dev.* 18: 926-936.
- Fukuoka S and Saka N (2006). Marker assisted combination of QTLs for yield resistance blast in rice. *Breed. Res.* 8: 191.

- Gilchrist EJ, Haughn GW, Ying CC, Otto SP, et al. (2006). Use of Ecotilling as an efficient SNP discovery tool to survey genetic variation in wild populations of *Populus trichocarpa*. *Mol. Ecol.* 15: 1367-1378.
- Gore MA, Chia JM, Elshire RJ, Sun Q, et al. (2009). A first-generation haplotype map of maize. *Science* 326: 1115-1117.
- Hadiarto T and Tran LS (2011). Progress studies of drought-responsive genes in rice. *Plant Cell Rep.* 30: 297-310.
- Han B, Xue Y, Li J, Deng XW, et al. (2007). Rice functional genomics research in China. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362: 1009-1021.
- Hao W and Lin HX (2010). Toward understanding genetic mechanisms of complex traits in rice. *J. Genet. Genomics* 37: 653-666.
- Hu KM, Qiu DY, Shen XL, Li XH, et al. (2008). Isolation and manipulation of quantitative trait loci for disease resistance in rice using a candidate gene approach. *Mol. Plant* 1: 786-793.
- Huang X, Feng Q, Qian Q, Zhao Q, et al. (2009). High-throughput genotyping by whole-genome resequencing. *Genome Res.* 19: 1068-1076.
- Huang X, Wei X, Sang T, Zhao Q, et al. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* 42: 961-967.
- International Rice Genome Sequencing Project (IRGSP) (2005). The map-based sequence of the rice genome. *Nature* 436: 793-800.
- Ito Y, Katsura K, Maruyama K, Taji T, et al. (2006). Functional analysis of rice DREB1/CBF-type transcription factors involved in cold-responsive gene expression in transgenic rice. *Plant Cell Physiol.* 47: 141-153.
- Li X, Qian Q, Fu Z, Wang Y, et al. (2003). Control of tillering in rice. *Nature* 422: 618-621.
- Lin HX, Yamamoto T, Sasaki T and Yano M (2000). Characterization and detection of epistatic interactions of 3 QTLs, Hd1, Hd2, and Hd3, controlling heading date in rice using nearly isogenic lines. *Theor. Appl. Genet.* 101: 1021-1028.
- Londo JP, Chiang YC, Hung KH, Chiang TY, 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.
- Manosalva PM, Davidson RM, Liu B, Zhu X, et al. (2009). A germin-like protein gene family functions as a complex quantitative trait locus conferring broad-spectrum disease resistance in rice. *Plant Physiol.* 149: 286-296.
- McCouch SR, Zhao K, Wright M, Tung CW, et al. (2010). Development of genome-wide SNP assays for rice. *Breed. Sci.* 60: 524-535.
- Mikami I, Uwatoko N, Ikeda Y, Yamaguchi J, et al. (2008). Allelic diversification at the wx locus in landraces of Asian rice. *Theor. Appl. Genet.* 116: 979-989.
- Olsen KM, Halldorsdottir SS, Stinchcombe JR, Weing C, et al. (2004). Linkage disequilibrium mapping of *Arabidopsis* CRY2 flowering time alleles. *Genetics* 167: 1361-1369.
- Ookawa T, Hobo T, Yano M, Murata K, et al. (2010). New approach for rice improvement using a pleiotropic QTL gene for lodging resistance and yield. *Nat. Commun.* 1: 132.
- Peltonen L, Palotie A and Lange K (2000). Use of population isolates for mapping complex traits. *Nat. Rev. Genet.* 1: 182-190.
- Rabbani MA, Maruyama K, Abe H, Khan MA, et al. (2003). Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. *Plant Physiol.* 133: 1755-1767.
- Rafalski JA (2010). Association genetics in crop improvement. *Curr. Opin. Plant Biol.* 13: 174-180.
- Ren ZH, Gao JP, Li LG, Cai XL, et al. (2005). A rice quantitative trait locus for salt tolerance encodes a sodium transporter. *Nat. Genet.* 37: 1141-1146.
- Sasaki T (2003). Rice genome analysis: understanding the genetic secrets of the rice plant. *Breed. Sci.* 53: 281-289.
- Steele KA, Edwards G, Zhu J and Witcombe JR (2004). Marker-evaluated selection in rice: shifts in allele frequency among bulks selected in contrasting agricultural environments identify genomic regions of importance to rice adaptation and breeding. *Theor. Appl. Genet.* 109: 1247-1260.
- Takeda T, Suwa Y, Suzuki M, Kitano H, et al. (2003). The OsTB1 gene negatively regulates lateral branching in rice. *Plant J.* 33: 513-520.
- Takeuchi Y, Ebitani T, Yamamoto T, Sato H, et al. (2006). Development of isogenic lines of rice cultivar Koshihikari with early and late heading by marker-assisted selection. *Breed. Sci.* 56: 405-413.
- Tian Z, Qian Q, Liu Q, Yan M, et al. (2009). Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities. *Proc. Natl. Acad. Sci. U. S. A.* 106: 21760-21765.
- Wang LQ, Liu WJ, Xu Y, He YQ, et al. (2007). Genetic basis of 17 traits and viscosity parameters characterizing the eating and cooking quality of rice grain. *Theor. Appl. Genet.* 115: 463-476.
- Xiang Y, Tang N, Du H, Ye H, et al. (2008). Characterization of OsbZIP23 as a key player of the basic leucine zipper transcription factor family for conferring abscisic acid sensitivity and salinity and drought tolerance in rice. *Plant Physiol.* 148: 1938-1952.
- Xie W, Feng Q, Yu H, Huang X, et al. (2010). Parent-independent genotyping for constructing an ultrahigh-density

- linkage map based on population sequencing. *Proc. Natl. Acad. Sci. U. S. A.* 107: 10578-10583.
- Xing Y and Zhang Q (2010). Genetic and molecular bases of rice yield. *Annu. Rev. Plant Biol.* 61: 421-442.
- Xu K, Xu X, Fukao T, Canlas P, et al. (2006). Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature* 442: 705-708.
- Xu X, Liu X, Ge S, Jensen JD, et al. (2012). Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat. Biotechnol.* 30: 105-111.
- Yano M and Sasaki T (1997). Genetic and molecular dissection of quantitative traits in rice. *Plant Mol. Biol.* 35: 145-153.
- Zhang Q, Li J, Xue Y, Han B, et al. (2008). Rice 2020: a call for an international coordinated effort in rice functional genomics. *Mol. Plant* 1: 715-719.