Rice Genome Research: Current Status and Future Perspectives

Bin Han* and Qifa Zhang*

Abstract

Rice (*Oryza sativa* L.) is the leading genomics system among the crop plants. The sequence of the rice genome, the first cereal plant genome, was published in 2005. This review summarizes progress made in rice genome annotations, comparative genomics, and functional genomics researches. It also maps out the status of rice genomics globally and provides a vision of future research directions and resource building.

RICE (*ORYZA SATIVA* L.) IS ONE of the most important crops feeding about half of the world's population. With a compact genome, the cultivated rice species *Oryza sativa* represents a model for cereals as well as other monocot plants (Shimamoto and Kyozuka, 2002; Paterson et al., 2005; Xu et al., 2005). Because of the availability of the whole genome sequences of both *indica* and *japonica* subspecies and abundant genetic and genomic resources, including mutants and wild rice species, rice has become a model for comparative genome analysis. The completion of the genome sequence of rice in 2005 opens a new and exciting chapter in our quest to functionally characterize all of the annotated genes in rice (IRGSP, 2005). A systematic approach to characterizing these genes will allow us to dissect and understand the regulatory networks and evolutionary selection controlling such complex traits as yield, grain quality, biotic and abiotic stresses, reproductive barriers, epigenetics, and flowering time. The next essential steps toward deciphering the sequenced genome are to develop complete and accurate maps of actively transcribed regions during rice development, and to generate more and more genome-wide rice mutant resources. These will facilitate the identification of all the genes and proteins encoded in the DNA sequence. Such information will allow

Bin Han, National Center for Gene Research and Institute of Plant Physiology and Ecology, Shanghai Institutes of Biological Sciences, Chinese Academy of Sciences, 500 Caobao Rd., Shanghai 200233, China; and Beijing Institute of Genomics, Chinese Academy of Sciences, No 7, Beitucheng West Rd., Beijing 100029, China; Qifa Zhang, National Key Lab. of Crop Genetic Improvement and National Center of Plant Gene Research (Wuhan), Huazhong Agricultural Univ., Wuhan 430070, China. Received 7 Sept. 2008.
*Corresponding authors (bhan@ncgr.ac.cn; qifazh@mail.hzau.edu.cn).

Abbreviations: ESTs, expressed sequence tags; IRGSP, International Rice Genome Sequencing Project; miRNAs, MicroRNAs; MPSS, massively parallel signature sequencing; QTLs, quantitative trait loci; TE, transposable elements.

Published in The Plant Genome 1:71–76. Published 21 Nov. 2008.
doi: 10.3835/plantgenome2008.09.0008

677 S. Segoe Rd., Madison, WI 53711 USA
An open-access publication

further analysis of their function and regulation, and how they cooperate in complex biological processes in a systems manner.

Progress in Sequencing and Annotation of the Genome

The International Rice Genome Sequencing Project (IRGSP) has adopted the clone-by-clone approach for sequencing rice *Oryza sativa* ssp. *japonica* Nipponbare genome sequence, because it allows efficient gap-filling, avoids problems arising from distant repetitive sequences and results in the early completion of larger contiguous segments of a genome (Sasaki and Burr, 2000). A map-based, finished quality sequence of rice *japonica* Nipponbare, which covers 95% of the 389 Mb genome, including virtually all of the euchromatin and two complete centromeres, was completed in 2005 (IRGSP, 2005; Feng et al. 2002; Sasaki et al. 2002; The Rice Chromosome 10 Sequencing Consortium, 2003). At the same time, a whole-genome shotgun sequencing approach was performed for sequencing an *indica* variety 93–11 genome (Yu et al., 2002; Yu et al., 2005).

To further understand rice adaptation and facilitate genetic improvement of rice a comparative genomics approach will be necessary to make a more integrated and detailed map that collects all kinds of genetic variations, which will need to include copy-number variation, gene loss caused by frame-shift or point mutation, and other specific evolutionary events.

With the completion of the rice genome sequencing, the Rice Annotation Project Database (RAP-DB) was created to provide the genome sequence assembly of the IRGSP release 3, a manually curated annotation of the sequence, and other genomics information that could be useful for comprehensive understanding of rice biology (Ohyanagi et al., 2006; The Rice Annotation Project, 2007; Rice Annotation Project, 2008). Also, the Rice Genome Annotation Project of The Institute for Genomic Research (TIGR) continues to improve the quality of the annotation and to update the rice genome sequence with new data (Ouyang et al., 2007). In the current release 4.0 of the annotation, 42,653 non-transposable element-related genes encoding 49,472 gene models were identified as a result of the detection of alternative splicing. Moreover, transposon is an important content of the rice genome, which is populated by representatives from all known transposon superfamilies (Juretic et al., 2004; IRGSP, 2005). Among them, Pack-MULEs can mobilize thousands of gene fragments, which may have an impact on rice genome evolution (Jiang et al., 2004; Juretic et al., 2005). The completion and annotation of the rice genome have afforded an unprecedented opportunity for systematic studies of plant gene function. The rice genome sequence provides a complete catalog of genes that are important for improving not only rice but also other cereals, as functionally important sequences are conserved and may be identified by their similarity.

Comparative Genomics in Rice

The Asian cultivated rice *Oryza sativa* contains two major subspecies *indica* and *japonica*, which show distinct divergence from sequence variations to phenotypic changes (Oka and Chang, 1962; Cheng and Lu, 1984). *Indica–japonica* genome comparison has been chosen as a model system for understanding the origin, speciation, domestication, and genome evolution of rice. *Indica* and *japonica* cultivars can be classified based on their agronomic traits (Oka and Morishima, 1997) and the use of molecular markers (Glaszmann, 1987; Cheng et al., 2003). The sequencing of *indica* and *japonica* rice genomes provides a powerful resource for comparative analyses, which could help to detect polymorphisms between them and also give clues to recent genome variations. With the available genome sequences and comparative analysis, DNA polymorphisms across the entire rice genome were discerned to develop molecular markers (Han and Xue, 2003; Feltus et al., 2004; Shen et al., 2004; Garris et al., 2005), which greatly facilitates gene cloning and also molecular-assisted breeding in rice.

Comparisons of chloroplast, mitochondrial, and nuclear genomes revealed that the two subspecies diverged ~0.44 million years ago (Tang et al., 2004; Ma and Bennetzen, 2004; Vitte et al., 2004; Tian et al., 2006). In-depth sequence comparison of genomic DNA sequences has also shown a large number of variations in intergenic and genic regions between *indica* and *japonica* (Feng et al., 2002; Han and Xue, 2003). Among them, transposable elements (TEs) serve as an important evolutionary driving force for intra-specific variation. The genome sizes of both *indica* and *japonica* have increased substantially, mainly because of the insertions of TEs (Ma and Bennetzen, 2004; Huang et al., 2008). The activities of the mobile elements are responsible for a series of genetic differences between two subspecies via various ways, including interrupting host genes, creating different expression forms, drastically changing intron length, and affecting expression levels of adjacent genes (Huang et al., 2008).

From a comparative standpoint of the genus *Oryza*, a comprehensive set of 12 bacterial artificial chromosome (BAC) libraries that represent the 10 genome types of *Oryza* has been constructed, as a first step to performing comparative genomic analysis within the

genus (Ammiraju et al., 2006). Furthermore, rice is also an important model species for the Poaceae. Comparative genome analysis of closely related species in cereals would be a powerful tool to identify conserved functional units and regulatory elements. Through comparisons with other plant genome sequences and transcript sequences, structural and functional features of the rice genome itself have been confirmed and improved (Zhu and Buell, 2007).

To further understand rice adaptation and facilitate genetic improvement of rice a comparative genomics approach will be necessary to make a more integrated and detailed map that collects all kinds of genetic variations, which will need to include copy-number variation, gene loss caused by frame-shift or point mutation, and other specific evolutionary events.

Functional Genomics Studies

Rice functional genomics is a scientific approach that seeks to identify and define the function and interaction of genes to produce phenotypic traits. Rapid progress in rice genome sequencing has facilitated research in rice functional genomics. The rice functional genomics researches include development of technical platforms, and molecular cloning and functional analysis of agronomic genes. The platforms are aimed at enabling high-throughput analyses and effective determination of gene functions, which consist of three major components (i) generation and characterization of a large mutant library, (ii) expression profiling of the predicted exons and expressed sequence tags (ESTs) of the entire genome, (iii) and isolation of full length cDNAs.

Development of Technical Platforms

After the release of the rice genome sequence, one of the significant challenges has been the large-scale identification of gene functions. Various methods and technical platforms have been employed to enable high-throughput analyses and effective determination of gene functions. Two major platforms of functional genomics studies are generation of a large mutant library and isolation of full-length cDNAs.

Insertional mutagenesis, an effective strategy to study gene function, has been widely applied to construct mutant libraries (Jeon et al., 2000; Wu et al., 2003; Hirochika et al., 2004). For mutant generation, T-DNA is the most frequently used foreign DNA, as it both disrupts the gene function to facilitate gene identification and provides tags making gene isolation easier. About 29,482, 47,932, and 13,804 T-DNA tag lines in *japonica* rice were individually generated (Sallaud et al., 2004; Jeong et al., 2006; Zhang et al., 2007). Mapping T-DNA flanking sequence tags on chromosomes revealed that T-DNA integration frequency was generally proportional to chromosome size; however, T-DNA insertions were nonrandomly distributed on each chromosome (Zhang et al., 2007). The *Tos17* disruption system is another efficient method and widely used for mutational analysis (Agrawal et al., 2001; Sakamoto et al., 2004). More than 50,000 disruption lines of *japonica* were produced using the endogenous retrotransposon Tos17 (Miyao et al., 2007). Phenotypes of these lines in the M2 generation were observed and characterized. This combination of phenotypic and flanking sequence data will stimulate the functional analysis of rice genes. In the rice T-DNA mutant library, large-scale characterization of Tos17 insertion sites has also been conducted (Piffanelli et al., 2007).

Full-length cDNA clones are valuable resources for the functional analysis of genes not only at transcriptional level but also at the translational level (Nishiyama et al., 2003). The promoter sequences can be obtained by comparing the 5′-end sequences of cDNAs with the rice genome sequences. Large-scale new gene discovery was accelerated by utilizing these procedures. Over 32,000 *japonica* full-length cDNAs have been published (The Rice Full-Length cDNA Consortium, 2003). Collections of over 20,000 full-length cDNAs and over 40,000 5′ ESTs isolated from various cDNA libraries of two *indica* varieties Guangluai 4 and Minghui 63 have also been done (Xie et al., 2005; Liu et al., 2007). In addition, the *indica* cDNA clones are useful for comparative analysis between *indica* and *japonica* subspecies, improvement for genome sequence annotation, and for identification of lineage of specific genes. Moreover, about 1888 wild rice *Oryza rufipogon* W1943 full-length cDNAs have been collected (Lu et al., 2008b). All these full-length cDNA clones provided important resources for further functional studies and could be broadly utilized in rice biological studies. Moreover, through comparative analysis among rice varieties, some putative *indica*-specific genes and wild rice specific genes were identified (Liu et al., 2007; Lu et al., 2008b).

Cloning and Functional Analysis of Agronomics Related Genes and QTLs

The rice genome sequence provides a complete catalog of genes that are very important for identification of the rice genes through map-based cloning strategy. Great progress has been made by rice researchers in recent years in the cloning and functional analysis of agronomically important traits, including plant architecture (Li et al., 2003), stress tolerance (Ren et al., 2005; Xu et al., 2006; Fukao et al., 2006), disease resistance (bacteria blight) (Chu et al., 2006), grain yielding (Ashikari et al., 2005; Song et al., 2007; Xue et al., 2008; Shomura et al., 2008), shattering habit (Li et al., 2006; Konishi et al., 2006), fertility (Wang et al., 2006; Chen et al., 2008), and nutrition efficiency.

Here, we describe the cloning of a number of rice genes and quantitative trait loci (QTLs). Tillering in rice is an important agronomic trait for grain production, and also a model system for the study of branching in monocotyledonous plants. A gene *MONOCULM 1* (*MOC1*) that controls rice tillering has been cloned and characterized (Li et al., 2003). Soil salinity is a major abiotic stress in crop productivity worldwide. Salt tolerance is a complex trait controlled by QTLs and is the final manifestation of

several components, such as Na^+ uptake, K^+ uptake, ions balance, and ions compartmentation. The *SKC1* gene that encodes an ion transporter was cloned from a high salt tolerance *indica* variety 'Nona Bokra' by using advanced backcross progeny and map-based cloning (Ren et al., 2005). The QTL Ghd7, which has played crucial roles in increasing productivity and adaptability of rice globally, was isolated from an elite rice hybrid and was found to encode a CCT domain protein (Xue et al., 2008). *Ghd7* has major effects on an array of traits in rice, including number of grains per panicle, plant height, and heading date. Enhanced expression of *Ghd7* under long-day conditions delays heading and increases plant height and panicle size. Natural mutants with reduced function enable rice to be cultivated in temperate and cooler regions.

MicroRNA and Gene Expression

Since the completion of the rice genome sequence, various technologies have been used to generate genome-wide expression profiles and to perform rice transcriptome analysis, including ESTs, microarrays platforms, and massively parallel signature sequencing (MPSS). As yet, more than 1,220,877 rice ESTs constructed from various kinds of tissues under various conditions have been sequenced and released in the NCBI-dbEST database. These are currently the most important resources for transcriptome exploration in rice. The experimental evidence greatly improved structural annotation of gene models, and also facilitated the development of whole-genome exon arrays. For example, Affymetrix (Santa Clara, CA, USA) rice whole genome array is a widespread exon array, which covered ~51,279 transcripts in rice. Moreover, there are also many transcription units outside of the annotated rice genes, of which the expression levels cannot be detected by these exon arrays. For this reason, several rice genome tiling arrays were developed to identify novel transcripts, especially in the non-coding region (Jiao et al., 2005; Li et al., 2005; Li et al., 2006). Furthermore, taking advantage of the development of sequencing technology, expression data are becoming available at considerably accelerated speeds. As a powerful alternative to microarray technologies, MPSS uses a unique method to quantify gene-expression levels and generate millions of short sequence tags per library. The recent discovery and analysis of rice MPSS data (i.e., 46,971,553 mRNA transcripts from 22 libraries) has characterized many previously unknown transcripts and demonstrated the biological importance of these noncoding molecules (Nobuta et al., 2007). With the advent of these platforms, the high-throughput analysis of global gene expression has become a useful and powerful tool in the study of functional genomics.

MicroRNAs (miRNAs) are 21 to 24 nucleotide sequences, which are involved in various biological processes often by interfering with mRNA translation. Various approaches using MPSS, the 454 technology, or the Ilumina technology have been used in rice for miRNA discovery. By carrying out deep sequencing of the small RNA populations of rice tissues at different developmental stages, currently about 240 miRNAs have been discovered and annotated in rice (Nobuta et al., 2007; Lu et al., 2008a). The published studies reveal that most of the miRNAs in rice are conserved in plants (Johnson et al., 2007; Zhu et al., 2008). These newly identified miRNAs in rice exhibit tissue-specific expression patterns and suggest a recent evolutionary origin (Morin et al., 2008; Sunkar et al., 2008).

Future Research

Although tremendous progress has been made in rice genomics, there is still a huge gap of knowledge between the genotype and phenotype; a gap which must be bridged in order to breed elite varieties suitable for sustainable agriculture. A highly coordinated effort that brings together scientists and resources worldwide is a desirable step and perhaps the only practical and efficient one. Zhang et al. (2008) proposed an International Rice Functional Genomics Project (IRFGP), with the ultimate goals of determining 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, and to apply the findings of functional genomics research to rice crop genetic improvement and beyond. The following objectives for this international effort, with elaboration of specific aims to be achieved, were proposed: (i) development of enabling tools and genetic resources for an international community of scientists to conduct functional genomics research in rice; (ii) assignment of biological functions to every annotated gene; (iii) global analyses of the proteome and protein-protein interactions; (iv) natural variation of *O. sativa* and its relatives; (v) bioinformatics, data management, and exchange of information; (vi) establishment of the toolkit for high throughput knowledge-based rice breeding. These provide vision for future research directions and resource-building in rice genomics studies.

In the application aspect, Zhang (2007) outlined strategies for the development of what he referred to as Green Super Rice based on advances in genome research. On the premise of continued yield increase and quality improvement, Green Super Rice should possess resistances to insects and diseases, high nutrient efficiency, and drought resistance; all of which promises to greatly reduce the consumption of pesticides, chemical fertilizers, and water. This would ensure that rice production is in harmony with the environment, thus safeguarding a proper level of sustainability. As rice is a major world crop, rice functional genomics research will have an immense global impact on sustainable agriculture.

Acknowledgments

We thank Tingting Lu and Xuehui Huang (National Center for Gene Research) for their assistance with preparation of this manuscript. This research was supported by the grants from the Ministry of Science and Technology of China (The China Rice Functional Genomics Programs, grant numbers 2006AA10A101 and 2006AA10A102).

References

Agrawal, G.K., M. Yamazaki, M. Kobayashi, R. Hirochika, A. Miyao, and H. Hirochika. 2001. Screening of the rice viviparous mutants generated by endogenous retrotransposon *Tos17* insertion. Tagging of a zeaxanthin epoxidase gene and a novel *OsTATC* gene. Plant Physiol. 125:1248–1257.

Ammiraju, J.S., M. Luo, J.L. Goicoechea, W. Wang, D. Kudrna, C. Mueller, J. Talag, H. Kim, N.B. Sisneros, B. Blackmon, E. Fang, J.B. Tomkins, D. Brar, D. MacKill, S. McCouch, N. Kurata, G. Lambert, D.W. Galbraith, K. Arumuganathan, K. Rao, J.G. Walling, N. Gill, Y. Yu, P. SanMiguel, C. Soderlund, S. Jackson, and R.A. Wing. 2006. The *Oryza* bacterial artificial chromosome library resource: Construction and analysis of 12 deep-coverage large-insert BAC libraries that represent the 10 genome types of the genus *Oryza*. Genome Res. 16:140–147.

Ashikari, M., H. Sakakibara, S. Lin, T. Yamamoto, T. Takashi, A. Nishimura, E.R. Angeles, Q. Qian, H. Kitano, and M. Matsuoka. 2005. Cytokinin oxidase regulates rice grain production. Science 309:741–745.

Chen, J., J. Ding, Y. Ouyang, H. Du, J. Yang, K. Cheng, J. Zhao, S. Qiu, X. Zhang, J. Yao, K. Liu, L. Wang, C. Xu, X. Li, Y. Xue, M. Xia, Q. Ji, J. Lu, M. Xu, and Q. Zhang. 2008. A triallelic system of S5 is a major regulator of the reproductive barrier and compatibility of *indica-japonica* hybrids in rice. Proc. Natl. Acad. Sci. USA 105:11436–11441.

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

Cheng, K.S., and Y.X. Lu. 1984. Reflection of *Oryza sativa* L. classification. Acta Agronomica Sinica 10:271–280.

Chu, Z., M. Yuan, J. Yao, X. Ge, B. Yuan, C. Xu, X. Li, B. Fu, Z. Li, J.L. Bennetzen, Q. Zhang, and S. Wang. 2006. Promoter mutations of an essential gene for pollen development result in disease resistance in rice. Gene Dev. 20:1250–1255.

Feltus, F.A., J. Wan, S.R. Schulze, J.C. Estill, N. Jiang, and A.H. Paterson. 2004. An SNP resource for rice genetics and breeding based on subspecies *Indica* and *Japonica* genome alignments. Genome Res. 14:1812–1819.

Feng, Q., Y. Zhang, P. Hao, S. Wang, G. Fu, Y. Huang, Y. Li, J. Zhu, Y. Liu, X. Hu, P. Jia, Y. Zhang, Q. Zhao, K. Ying, S. Yu, Y. Tang, Q. Weng, L. Zhang, Y. Lu, J. Mu, Y. Lu, L.S. Zhang, Z. Yu, D. Fan, X. Liu, T. Lu, C. Li, Y. Wu, et al. 2002. Sequence and analysis of rice chromosome 4. Nature 420:316–320.

Fukao, T., K. Xu, P. Ronald, and J. Bailey-Serres. 2006. A variable cluster of ethylene responsive-like factors regulates metabolic and developmental acclimation responses to submergence in rice. Plant Cell 18:2021–2034.

Garris, A.J., T.H. Tai, J. Coburn, S. Kresovich, and S. McCouch. 2005. Genetic structure and diversity in *Oryza sativa* L. Genetics 169:1631–1638.

Glaszmann, J.C. 1987. Isozymes and classification of Asian rice varieties. Theor. Appl. Genet. 74:21–30.

Han, B., and Y.B. Xue. 2003. Genome-wide intraspecific DNA-sequence variations in rice. Curr. Opin. Plant Biol. 6:134–138.

Hirochika, H., E. Guiderdoni, G. An, Y.I. Hsing, M.Y. Eun, C.D. Han, N. Upadhyaya, S. Ramachandran, Q. Zhang, A. Pereira, V. Sundaresan, and H. Leung. 2004. Rice mutant resources for gene discovery. Plant Mol. Biol. 54:325–334.

Huang, X., G. Lu, Q. Zhao, X. Liu, and B. Han. 2008. Genome-wide analysis of transposon insertion polymorphisms reveals intra-specific variation in cultivated rice. Plant Physiol. 148:25–40.

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

Jeon, J.S., S. Lee, K.H. Jung, S.H. Jun, D.H. Jeong, J. Lee, C. Kim, S. Jang, K. Yang, J. Nam, K. An, M.J. Han, R.J. Sung, H.S. Choi, J.H. Yu, J.H. Choi, S.Y. Cho, S.S. Cha, S.I. Kim, and G. An. 2000. T-DNA insertional mutagenesis for functional genomics in rice. Plant J. 22:561–570.

Jeong, D.H., S. An, S. Park, H.G. Kang, G.G. Park, S.R. Kim, J. Sim, Y.O. Kim, M.K. Kim, S.R. Kim, J. Kim, M. Shin, M. Jung, and G. An. 2006. Generation of a flanking sequence-tag database for activation-tagging lines in *japonica* rice. Plant J. 45:123–132.

Jiang, N., Z.R. Bao, X.Y. Zhang, S.R. Eddy, and S.R. Wessler. 2004. Pack-MULE transposable elements mediate gene evolution in plants. Nature 431:569–573.

Jiao, Y., P. Jia, X. Wang, N. Su, S. Yu, D. Zhang, L. Ma, Q. Feng, Z. Jin, L. Li, Y. Xue, Z. Cheng, H. Zhao, B. Han, and X.W. Deng. 2005. A tiling microarray expression analysis of rice chromosome 4 suggests a chromosome-level regulation of transcription. Plant Cell 17:1641–1657.

Johnson, C., L. Bowman, A.T. Adai, V. Vance, and V. Sundaresan. 2007. CSRDB: A small RNA integrated database and browser resource for cereals. Nucleic Acids Res. 35:D829–D833.

Juretic, N., T.E. Bureau, and R.M. Bruskiewich. 2004. Transposable element annotation of the rice genome. Bioinformatics 20:155–160.

Juretic, N., D.R. Hoen, M.L. Huynh, P.M. Harrison, and T.E. Bureau. 2005. The evolutionary fate of MULE-mediated duplications of host gene fragments in rice. Genome Res. 15:1292–1297.

Konishi, S., T. Izawa, S.Y. Lin, K. Ebana, Y. Fukuta, T. Sasaki, and M. Yano. 2006. An SNP caused loss of seed shattering during rice domestication. Science 312:1392–1396.

Li, C., A. Zhou, and T. Sang. 2006. Rice domestication by reducing shattering. Science 311:1936–1939.

Li, L., X. Wang, M. Xia, V. Stolc, N. Su, Z. Peng, T. Waraporn, S. Li, J. Wang, X. Wang, and X.W. Deng. 2005. Tiling microarray analysis of rice chromosome 10 to identify the transcriptome and relate its expression to chromosomal architecture. Genome Biol. 6:R52.

Li, L., X. Wang, V. Stolc, X. Li, D. Zhang, N. Su, W. Tongprasit, S. Li, Z. Cheng, J. Wang, and X.W. Deng. 2006. Genome-wide transcription analyses in rice using tiling microarrays. Nat. Genet. 38:124–129.

Li, X.Y., Q. Qian, Z.M. Fu, Y.H. Wang, G.S. Xiong, D.L. Zeng, X.Q. Wang, X.F. Liu, S. Teng, F. Hiroshi, M. Yuan, D. Luo, B. Han, and J.Y. Li. 2003. Control of tillering in rice. Nature 422:618–621.

Liu, X., T. Lu, S. Yu, Y. Li, Y. Huang, T. Huang, L. Zhang, J. Zhu, Q. Zhao, D. Fan, J. Mu, Y. Shangguan, Q. Feng, J. Guan, K. Ying, Y. Zhang, Z. Lin, Z. Sun, Q. Qian, Y. Lu, and B. Han. 2007. A collection of 10,096 *indica* rice full-length cDNAs reveals highly expressed sequence divergence between *Oryza sativa indica* and *japonica* subspecies. Plant Mol. Biol. 65:403–415.

Lu, C., D.H. Jeong, K. Kulkarni, M. Pillay, K. Nobuta, R. German, S.R. Thatcher, C. Maher, L. Zhang, D. Ware, B. Liu, X. Cao, B.C. Meyers, and P.J. Green. 2008a. Genome-wide analysis for discovery of rice microRNAs reveals natural antisense microRNAs (nat-miRNAs). Proc. Natl. Acad. Sci. USA 105:4951–4956.

Lu, T., S. Yu, D. Fan, J. Mu, Y. Shangguan, Z. Wang, Y. Minobe, Z.X. Lin, and B. Han. 2008b. Collection and comparative analysis of 1888 full-length cDNAs from wild rice *Oryza rufipogon* Griff. W1943. DNA Res.

Ma, J.X., and J.L. Bennetzen. 2004. Rapid recent growth and divergence of rice nuclear genomes. Proc. Natl. Acad. Sci. USA 101:12404–12410.

Miyao, A., Y. Iwasaki, H. Kitano, J. Itoh, M. Maekawa, K. Murata, O. Yatou, Y. Nagato, and H. Hirochika. 2007. A large-scale collection of phenotypic data describing an insertional mutant population to facilitate functional analysis of rice genes. Plant Mol. Biol. 54:325–334.

Morin, R.D., G. Aksay, E. Dolgosheina, H.A. Ebhardt, V. Magrini, E.R. Mardis, S.C. Sahinalp, and P.J. Unrau. 2008. Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. Genome Res. 18:571–584.

Nishiyama, T., T. Fujita, I. Shin, M. Seki, H. Nishide, I. Uchiyama, A. Kamiya, P. Carninci, Y. Hayashizaki, K. Shinozaki, Y. Kohara, and M. Hasebe. 2003. Comparative genomics of *Physcomitrella patens* gametophytic transcriptome and *Arabidopsis thaliana*: Implication for land plant evolution. Proc. Natl. Acad. Sci. USA 100:8007–8012.

Nobuta, K., R.C. Venu, C. Lu, A. Beló, K. Vemaraju, K. Kulkarni, W. Wang, M. Pillay, P.J. Green, G.L. Wang, and B.C. Meyers. 2007. An expression atlas of rice mRNAs and small RNAs. Nat. Biotechnol. 25:473–477.

Ohyanagi, H., T. Tanaka, H. Sakai, Y. Shigemoto, K. Yamaguchi, T. Habara, Y. Fujii, B.A. Antonio, Y. Nagamura, T. Imanishi, K. Ikeo, T. Itoh, T. Gojobori, and T. Sasaki. 2006. The rice annotation project

database (RAP-DB): Hub for *Oryza sativa* ssp. *japonica* genome information. Nucleic Acids Res. 34:D741–D744.
Oka, H., and H. Morishima. 1997. Wild and cultivated rice. p. 88–111. *In* T. Matsuo et al. (ed.) Science of the rice plant, Vol. 3.
Oka, H.I., and W.T. Chang. 1962. Rice varieties intermediate between wild and cultivated forms and the origin of the *japonica* type. Bot. Bull. Acad. Sinica (Taiwan) 3:109–131.
Ouyang, S., W. Zhu, J. Hamilton, H. Lin, M. Campbell, K. Childs, F. Thibaud-Nissen, R.L. Malek, Y. Lee, L. Zheng, J. Orvis, B. Haas, J. Wortman, and C.R. Buell. 2007. The TIGR rice Genome annotation resource: Improvements and new features. Nucleic Acids Res. 35:D883–D887.
Paterson, A.H., M. Freeling, and T. Sasaki. 2005. Grains of knowledge: Genomics of model cereals. Genome Res. 15:1643–1650.
Piffanelli, P., G. Droc, D. Mieulet, N. Lanau, M. Bès, E. Bourgeois, C. Rouvière, F. Gavory, C. Cruaud, A. Ghesquière, and E. Guiderdoni. 2007. Large-scale characterization of *Tos17* insertion sites in a rice T-DNA mutant library. Plant Mol. Biol. 65:587–601.
Ren, Z.H., J.P. Gao, L.G. Li, X.L. Cai, W. Huang, D.Y. Chao, M.Z. Zhu, Z.Y. Wang, S. Luan, and H.X. Lin. 2005. A rice quantitative trait locus for salt tolerance encodes a sodium transporter. Nat. Genet. 37:1141–1146.
Rice Annotation Project. 2008. The Rice Annotation Project Database (RAP-DB): 2008 update. Nucleic Acids Res. 36:D1028–D1033.
Sasaki, T., and B. Burr. 2000. International rice genome sequencing project: The effort to completely sequence the rice genome. Curr. Opin. Plant Biol. 3:138–141.
Sasaki, T., T. Matsumoto, K. Yamamoto, K. Sakata, T. Baba, Y. Katayose, J. Wu, Y. Niimura, Z. Cheng, Y. Nagamura, B.A. Antonio, H. Kanamori, S. Hosokawa, M. Masukawa, K. Arikawa, Y. Chiden, M. Hayashi, M. Okamoto, T. Ando, H. Aoki, K. Arita, M. Hamada, C. Harada, S. Hijishita, M. Honda, Y. Ichikawa, A. Idonuma, M. Iijima, et al. 2002. The genome sequence and structure of rice chromosome 1. Nature 420:312–316.
Sakamoto, T., K. Miura, H. Itoh, T. Tatsumi, M. Ueguchi-Tanaka, K. Ishiyama, M. Kobayashi, G.K. Agrawal, S. Takeda, K. Abe, A. Miyao, H. Hirochika, H. Kitano, M. Ashikari, and M. Matsuoka. 2004. An overview of gibberellin metabolism enzyme genes and their related mutants in rice. Plant Physiol. 134:1642–1653.
Sallaud, C., C. Gay, P. Larmande, M. Bes, P. Piffanelli, B. Piegu, G. Droc, F. Regad, E. Bourgeois, D. Meynard, C. Perin, X. Sabau, A. Ghesquiere, J.C. Glaszmann, M. Delseny, and E. Guiderdoni. 2004. High throughput T-DNA insertion mutagenesis in rice: A first step towards *in silico* reverse genetics. Plant J. 39:450–464.
Shen, Y.J., H. Jiang, J.P. Jin, Z.B. Zhang, B. Xi, Y.Y. He, G. Wang, C. Wang, L. Qian, X. Li, Q.B. Yu, H.J. Liu, D.H. Chen, J.H. Gao, H. Huang, T.L. Shi, and Z.N. Yang. 2004. Development of genome-wide DNA polymorphism database for map-based cloning of rice genes. Plant Physiol. 135:1198–1205.
Shimamoto, K., and J. Kyozuka. 2002. Rice as a model for comparative genomics of plants. Annu. Rev. Plant Biol. 53:399–419.
Shomura, A., T. Izawa, K. Ebana, T. Ebitani, and H. Kanegae. 2008. Deletion in a gene associated with grain size increased yields during rice domestication. Nat. Genet. 40:1023–1028.
Song, X.J., W. Huang, M. Shi, M.Z. Zhu, and H.X. Lin. 2007. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat. Genet. 39:623–630.
Sunkar, R., X. Zhou, Y. Zheng, W. Zhang, and J.K. Zhu. 2008. Identification of novel and candidate miRNAs in rice by high throughput sequencing. BMC Plant Biol. 8:25.
Tang, J., H. Xia, M. Cao, X. Zhan, W. Zeng, S. Hu, W. Tong, J. Wang, J. Wang, J. Yu, H. Yang, and L. Zhu. 2004. A comparison of rice chloroplast genomes. Plant Physiol. 135:412–420.
The Rice Chromosome 10 Sequencing Consortium. 2003. In-depth view of structure, activity, and evolution of rice chromosome 10. Science 300:1566–1569.
The Rice Full-Length cDNA Consortium. 2003. Collection, mapping, and annotation of over 28,000 cDNA clones from *japonica* rice. Science 301:376–379.
Tian, X., J. Zheng, S. Hu, and J. Yu. 2006. The rice mitochondrial genomes and their variations. Plant Physiol. 140:401–410.
Vitte, C., T. Ishii, F. Lamy, D. Brar, and O. Panaud. 2004. Genomic paleontology provides evidence for two distinct origins of Asian rice (*Oryza sativa* L.). Mol. Gen. Genomics 272:504–511.
Wang, Z., Y. Zou, X. Li, Q. Zhang, L. Chen, H. Wu, D. Su, Y. Chen, J. Guo, D. Luo, Y. Long, Y. Zhong, and Y.G. Liu. 2006. Cytoplasmic male sterility of rice with Boro II cytoplasm is caused by a cytotoxic peptide and is restored by two related PPR motif genes via distinct modes of mRNA silencing. Plant Cell 18:676–687.
Wu, C., X.J. Li, W.Y. Yuan, G.X. Chen, A. Kilian, J. Li, C. Xu, X. Li, D.X. Zhou, S. Wang, and Q. Zhang. 2003. Development of enhancer trap lines for functional analysis of the rice genome. Plant J. 35:418–427.
Xie, K., J. Zhang, Y. Xiang, Q. Feng, B. Han, Z. Chu, S. Wang, Q. Zhang, and L. Xiong. 2005. Isolation and annotation of 10,828 putative full length cDNAs from *indica* rice. Science in China 35:6–12 (ser C).
Xu, K., X. Xu, T. Fukao, P. Canlas, R. Maghirang-Rodriguez, S. Heuer, A.M. Ismail, J. Bailey-Serres, P.C. Ronald, and D.J. Mackill. 2006. *Sub1A* is an ethylene-response-factor-like gene that confers submergence tolerance to rice. Nature 442:705–708.
Xu, Y., S.R. McCouch, and Q. Zhang. 2005. How can we use genomics to improve cereals with rice as a reference genome? Plant Mol. Biol. 59:7–26.
Xue, W., Y. Xing, X. Weng, Y. Zhao, W. Tang, L. Wang, H. Zhou, S. Yu, C. Xu, X. Li, and Q. Zhang. 2008. Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. Nat. Genet. 40:761–767.
Yu, J., S. Hu, J. Wang, G.K. Wong, S. Li, B. Liu, Y. Deng, L. Dai, Y. Zhou, X. Zhang, M. Cao, J. Liu, J. Sun, J. Tang, Y. Chen, X. Huang, W. Lin, C. Ye, W. Tong, L. Cong, J. Geng, Y. Han, L. Li, W. Li, G. Hu, X. Huang, W. Li, J. Li, Z. Liu, L. Li, et al. 2002. A draft sequence of the rice genome (*Oryza sativa* ssp. *indica*). Science 296:79–91.
Yu, J., J. Wang, W. Lin, S. Li, H. Li, J. Zhou, P. Ni, W. Dong, S. Hu, C. Zeng, J. Zhang, Y. Zhang, R. Li, Z. Xu, S. Li, X. Li, H. Zheng, L. Cong, L. Lin, J. Yin, J. Geng, G. Li, J. Shi, J. Liu, H. Lv, J. Li, J. Wang, Y. Deng, L. Ran, X. Shi, X. Wang, Q. Wu, C. Li, X. Ren, J. Wang, X. Wang, D. Li, D. Liu, X. Zhang, Z. Ji, W. Zhao, Y. Sun, Z. Zhang, J. Bao, Y. Han, L. Dong, J. Ji, P. Chen, S. Wu, J. Liu, Y. Xiao, D. Bu, J. Tan, L. Yang, C. Ye, J. Zhang, J. Xu, et al. 2005. The genomes of *Oryza sativa*: A history of duplications. PLoS Biol. 3:e38.
Zhang, J., D. Guo, Y. Chang, C. You, X. Li, X. Dai, Q. Weng, J. Zhang, G. Chen, X. Li, H. Liu, B. Han, Q. Zhang, and C. Wu. 2007. Non-random distribution of T-DNA insertions at various levels of the genome hierarchy as revealed by analyzing 13 804 T-DNA flanking sequences from an enhancer-trap mutant library. Plant J. 49:947–959.
Zhang, Q. 2007. Strategies for developing Green Super Rice. Proc. Natl. Acad. Sci. USA 104:16402–16409.
Zhang, Q., J. Li, Y. Xue, B. Han, and X.W. Deng. 2008. Rice 2020: A call for an international coordinated effort in rice functional genomics. Mol. Plant. 1:715–719.
Zhu, Q., A. Spriggs, L. Matthew, L. Fan, G. Kennedy, F. Gubler, and C. Helliwell. 2008. A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains. Genome Res. 18:1456–1465.
Zhu, W., and C.R. Buell. 2007. Improvement of whole-genome annotation of cereals through comparative analyses. Genome Res. 17:299–310.