NBU Journal of Plant Sciences

Vol. 10, No. 1, March 2016 p. 17-30
ISSN No. 0974-6927

Review Article

Rice research in the high-throughput sequencing era: Genomic breeding *Rice breeding for better health*

Subhas Chandra Roy*

Plant Genetics & Molecular Breeding Laboratory, Department of Botany, University of North Bengal, Siliguri-734013, West Bengal

Abstract

Rice [*Oryza sativa* L.] is the most important cereal crop belongs to the family Poaceae (Grass) which provide staple food for half of the World's population (>3.3 billion). This staple food grain (rice) supplies the main energy resource providing 40-75% of the daily calorie intake to the world's poor people. It is equivalent to the proposition that 'Rice is life' in Asian continent because 90% people dependent on for their sustainable livelihood. Simultaneously Asia is considered as 'Rice Basket' because it produces 90% of the world's production (662 million tons, paddy rice, Mt). Total world production was 729 Mt from 154.3 million hectares with productivity of 4.1 tons/hectare (t/ha) in 2012 of which 662 million tons produced by Asian countries. Rice production has been doubled in the recent decades (1960s-1990s) during the time of Green Revolution (1960s) primarily as the result of genetic improvement. It was factual that the varieties released in the last 30 years in the farmers field, had a narrow genetic base in-spite of high genetic diversity prevailed in the rice germplasm, and yield enhancing capacity has reached to plateau. We need more production of rice to feed 9 Billion people in 2050. Breeder could manage the yield increase over released varieties through genetic gain by combining the yield related genes/QTLs from various genetic resources of rice germplasms either from cultivated local landraces or from wild varieties. Germplasm diversity is the mainstay for crop improvement and genetic dissection of complex traits. Rice germplasm shows tremendous genetic diversity in both within the species and among the varietal groups. This genetic diversity may be associated with the diverse alleles of important traits and can be exploited to introgress these traits using knowledge of molecular breeding techniques such as marker assisted breeding (MAB) or marker assisted selection (MAS). The Next Generation Sequencing based technology is used for whole genome analysis to unveil the genetic and genomic information pertaining to important traits for advancing the molecular breeding procedures to increase the production. That ultimately leads to the development of genomic breeding and genomic selection to accelerate the breeding process.

Key words: Rice, Molecular Breeding, Genomic selection, MABC breeding, MAS, GAB.

Introduction

Genomic Breeding [Genomic Assisted Breeding (GAB) and Genomic Selection (GS)]

Rice genome sequencing information has revolutionized the research dimension in rice genetics and breeding and as a result many varieties were improved that will continue to feed the growing world population (Scott, 2016). Now, rice has been considered as a model crop for genetics and breeding. Genome sequencing information of rice allows the breeders to better understand the genetic variation and can exploit this genetic variation to improve the HYV (McCouch *et al.*, 2012). Molecular understanding of the genetic basis of traits related to N and P-use has been acquired from the genome sequencing analysis and that has been utilized in engineering 'Green Super Rice' (GSR) to enhance the yield and quality of rice with minimum inputs (Zhang, 2007). Marker assisted selection (MAS) is an indirect process of selection where screening and individual plant identification is carried out on the basis of markers (DNA markers) instead of the expressed phenotypic trait. Therefore, thriving benefits of MAS depends on the tight association between the markers and the gene (or *QTLs*) responsible for that traits. Recombination associated phenomenon between linked markers and the gene(s) limits the applicability of MAS breeding. The application of intragenic markers, termed as functional markers can help to surmount this

*Corresponding author
E-mail: subhascr@rediffmail.com

problem (Andersen and Lübberstedt, 2003). Thus, new type of genomic tools such as NGS can accelerate the identification of markers tightly linked to target genomic regions. Limitations of MAS procedure can be overcome using a discrete system known as genomic selection (GS or genome-wide selection), based on the high-throughput Next Generation Sequencing (NGS) technique. Huge amount of genomic information can be assembled using the Next Generation Sequencing (NGS) technologies, which allows mass sequencing of genomes and transcriptomes. The bioinformatics analysis of this genomic data escorts to the development of large collections of molecular markers leading to the discovery of new gene/*QTLs*, position and dissection of complex traits. The NGS technology has revolutionized the genotyping ability (gene and *QTLs* discovery) of large population that help breeders to accelerate the breeding program. The SNPs markers are available in this NGS era and thus becoming the preferred choice of marker system in modern genomics research (Ganal *et al.* 2009). SNPs are more abundant, stable, amenable to automation, and most cost-effective in this NGS system. Genes and *QTLs* are identified through genome-wide association studies (GWAS) based on association mapping which also termed as linkage disequilibrium (LD) mapping (phenotype-genotype associations). High-density genome markers are being successfully used in background and foreground selection (also termed as positive selection) in rice breeding. Due to the advantage of the NGS technique, it is now possible to re-sequence the whole genome of any rice varieties to study the genetic diversity and variations. Re-sequencing of genome is constructive in order to the genome-wide discovery of markers for the building of high density genetic maps based on SNPs or SSR markers. These types of SSR/SNP markers are exploited in marker assisted selection (MAS), which include marker assisted backcross selection and helps the breeder to design the genotype of a cross and commonly known as 'Breeding-by-Design', because genotype of the individual is predetermined and predesigned. The same markers may be utilized in 'genomic selection' (GS) in the genomic breeding technique (Pérez-de-Castro *et al.*, 2012) to improve the rice varieties. Genomic selection (GS) is a new type of breeding method (Genomic breeding) in which genome-wide markers is used to predict the breeding value of individuals in a breeding population.

In this way, many genes and *QTLs* of agronomically valuable traits are identified and information about LD (linkage disequilibrium) and millions of SNPs data are detected from the genome sequencing analyses, which as a whole accelerating the genomic breeding to improve the varieties. The genomic breeding approach can be used as an advanced tool to the alternative of the conventional marker-based genotyping process for discovery of gene (s) and *QTL*. Marker assisted selection (MAS) in molecular breeding system has accelerated the development of crop varieties with improved yield, quality and biotic and abiotic stress tolerance traits. But these traits are complex and governed by many genes, each with small effect. The traditional marker-assisted selection (MAS) has not been so effective to analyse such traits. Marker-assisted selection (MAS) has failed to significantly improve polygenic traits (Bernardo, 2008; Xu and Crouch, 2008). To solve these difficulties, a novel statistical method has been employed to enable the simultaneous estimation of all marker effects and we can get genomic selection (GS) (Meuwissen *et al.*, 2001). The introduction of genomic selection (GS), however, has shifted that paradigm in the era of low cost genome sequencing regime (Jannink *et al.*, 2010). Genomic selection uses a 'training population' of individuals that have been both genotyped and phenotyped to develop a model that takes genotypic data from a 'candidate population' of untested individuals and produces genomic estimated breeding values (GEBVs) (Heffner *et al.*, 2009). In simulation studies, GEBVs based solely on individuals' genotype have been remarkably accurate. Some statistical methods have been used (Best Linear Unbiased Predictors (BLUP)' Bayesian regression, machine learning methods) to develop prediction models for genomic

selection (GS). Both the markers, SSR-like multiallelic markers and SNP-like biallelic markers has given the same prediction value, thus both are reliable in the GS.

Mapping population

In molecular breeding system, mapping populations are mandatory which facilitates *QTL* mapping and gene function analysis based on association studies. Different categories of populations are developed and used as mapping population such as bi-parental and multi-parent mapping populations, mutant populations, and immortalized recombinant inbred lines (RILs), BILs, F2 populations. Primary knowledge is gained from the analyses of mapping population using DNA markers subsequently are being utilized for the identification and map location of agronomically important genes/*QTLs* that provides the basis for marker-assisted selection (MAS) in plant breeding or in genomic selection (GS).

QTL mapping through conventional markers such as SSR, produced low resolution linkage map and time consuming. Most of the agronomically important traits are associated with multiple genes and named as *QTL* (Quantitative Trait Loci). Thus cloning and identification of a gene/*QTL* related to the traits is somewhat problematic and laborious using marker based technique. It is easy if any one uses the sequencing-based genotyping to discover gene or *QTLs* (Huang *et al.,* 2012). Genomic selection is carried out based on simultaneous estimation of effects on phenotype considering all *loci*, haplotype pattern, and markers available to calculate genomic value which ultimately used to select the desired phenotypes. These estimated value is termed as the genome estimated breeding values (GEBVs), are the output from a model of the relationship between the genome-wide markers and phenotypes of the individuals undergoing selection. Varshney *et al.* (2014) studied the need of genomics breeding in crop improvement using the NGS technology which act as highly multiplexed genotyping system and used in many different ways for genotyping the individuals such as whole genome sequencing (WGS), whole genome re-sequencing (WGRS), and genotyping by sequencing (GBS) system. Genome-wide association studies (GWAS) method based on NGS system has offered an effective technique to analyze the genetic architecture of complex traits and allowed identification of candidate genes for further improvement of agronomically important traits (Huang *et al.,* 2010; 2012).

Genomic Era in Rice Breeding

The International Rice Genome Sequencing Project has released the completed genome sequencing report of cultivar Nipponbare of *O. sativa* ssp. japonica (IRGSP 2005) in the year 2005. Before that draft genomic sequence of this Nipponbare cultivar was published in 2002 (Yu *et al.,* 2002). In the same year, the draft genomic sequence of one Chines cultivar 93-11 of subspecies *indica* was also released (Goff *et al.,* 2002). Three rice genome sequences are now available as reference genome such as *Aus* rice cultivar Kasalath, *indica rice* variety 93-11 (Kanamori *et al.,* 2013; Gao *et al.,* 2013) and Nipponbare of *japonica* rice cultivar (IRGSP, 2005). Genomic information acquired from these reports is helpful in rice breeding by analysing the functional genomics markers. Reference genomes are served as a source of gene catalogue to identify high density polymorphic markers, which may be related to the genes of particular traits (Feltus *et al.,* 2004). The Nest Generation Sequencing (NGS) technologies such as Illumina/Solexa (Gao *et al.,* 2012) have revolutionized the genotyping and functional genomics approaches for discovering new genes and alleles to accelerate the breeding process. The availability of high quality whole genome sequence provided a thorough understanding of the genome structure and evolution patterns. Novel means of understanding the genome functions could be designed by utilizing the DNA sequence information. These developments led to the birth of a new discipline in biology aptly termed as "Genomics".

Short reads of the re-sequencing (NGS) data from rice varieties can accurately align using any one as the reference genome from these three. The comparative analysis can also be done using these three reference

genome at a time to analyze genetic variation (Xian and Huang, 2013) including SNP, InDel (insertions-deletions), SV (structural variation), and CNV (copy number variations). Based on the reference genome, simple sequence repeat (SSR) markers are now easily available for any region of the rice genome because we now know the each and every base pair of rice genome (such as Nipponbare, indica 93-11, Kasalath). Simple sequence repeat (SSR) markers are easily available for any region of the genome, and candidate gene markers are being developed rapidly. Now, rice is considered as a model plant in the cereal crops. Genome sequencing information of the these reference genome (*indica*, aus, *japonica* rice subspecies) have provided breeders with the necessary tools for marker assisted breeding (MAS).

Application of Marker Assisted Selection (MAS)

In MAS breeding, markers associated with the genes are the main indication for gene mapping and identification. Many genes and *QTLs* with major effects are being utilized to develop improved varieties of rice using marker assisted backcrossing breeding (MABC). Submergence tolerant mega varieties has been development and released within 2-3 years was a significant paradigm shift in rice breeding with MAS. The marker assisted selection (MAS) and marker-assisted back-crossing (MABC) systems have been used in rice breeding to incorporate the genes/*QTLs* from wild or unadapted genetic resources (Septiningsih *et al.*, 2009; Imai *et al.*, 2013; Uga *et al.*, 2013).

The NGS based re-sequencing knowledge is used to identify genome wide variation within a species and genetic diversity within the population and to determine availability of linkage disequilibrium (Huang *et al.*, 2010; Jeong *et al.*, 2013; Guo *et al.*, 2014) among the varieties. Re-sequencing based generated data are applicable in the studies of high-throughput genotyping, and have been used for large-scale gene discovery in rice from RIL lines (Huang *et al.*, 2009), haplotype construction (Xie *et al.*, 2010); *QTL* mapping for culm length (Xu *et al.*, 2010), *QTL* detection for grains (Yu *et al.*, 2011); identifying 49 *QTLs* for 14 agronomic traits and gene discovery in rice (Wang *et al.*, 2011); genome-wide association studies (GWAS) for agronomic traits (Huang *et al.* 2010, 2011); identifying agronomic *QTL* in rice (Xu *et al.*, 2012); and rapid *QTL* mapping in rice (Takagi *et al.*, 2013). Millions of single nucleotide polymorphism (SNP) and insertion-deletion (InDel) markers have been identified in rice based on genome sequencing results.

Thus rice genome has been saturated with such sequence based markers SSR and SNP which is accelerating fine mapping of genes/*QTLs* and developing gene-based allele-specific markers. Markers based genomic information is necessary for the improvement of rice breeding programs. This genetic information are used to study the genetic variation, *QTL* identification (Quantitative Trait Loci) by GWAS, origin of cultivated rice from wild species and help in molecular breeding strategies. Both the processes are employed in background selection integrated with foreground selection to identify the superior lines with maximum recovery of Basmati rice genome with bacterial blight resistance genes (*xa13* and *Xa21* genes) (Gopalakrishnan *et al.*, 2008). Allelic variant can also be detected by using genomic study based on TILLING and EcoTILLING technology, through these approaches one can screen mutant lines and germplasm collections for their genetic variation to detect agronomically important trait specific genes.

Genomic Assisted Breeding (GAB) and Genomic Selection (GS)

Sequencing-based GWAS is employed for the purpose of genetic mapping and uncovering the genetic variability existed among the landraces of rice from the mapping populations (Abe *et al.*, 2012; Huang *et al.*, 2013). Some studies have been performed to identify the allelic variations existing in the rice germplasm through the technique of next generation sequencing (NGS) technologies (Huang *et al.*, 2010; Xu *et al.*, 2011; Huang *et al.*, 2012; 3K RGP, 2014). The NGS based technology is used in re-sequencing the whole genome of rice to unveil the genetic

and genomic information pertaining to important traits for advancing the molecular breeding procedures to increase the production.

The genomics assisted breeding (GAB) is based on genomic selection. In the genomic selection (GS), breeder uses all available markers information within a mapping population to predict genomic estimated breeding values (GEBVs) as a whole. The GAB system has an advantageous value over conventional breeding where genotyping data obtained from a seed or seedling stage can be used to estimate the trait related performance of mature plants without waiting for full grown plants in the fields, which reduces the time and cost for breeding program (Varshney *et al.,* 2005). The MAGIC (multi-parent advanced generation inter-cross) mapping populations are used to shuffle the genetic background among a set of diverse parental lines and increase recombination, which consequently increase resolution of *QTL* mapping (Bandillo *et al.,* 2013). Blast resistance gene *Pil* in rice was identified using WGRS technique, which is a NBS-LRR (nucleotide-binding site-leucine rich repeat) type protein gene (Takagi *et al.,* 2013). Grain yield and yield under drought conditions has been validated in rice using MAS breeding (Imai *et al.,* 2013; Mishra *et al.,* 2013; Venuprasad *et al.,* 2012). Genomics-assisted breeding has radically changed the approach so that breeders can use unadapted genetic resources in breeding program to improve rice varieties. Blast disease is a devastating disease of rice caused by fungal pathogen *Magnaporthae grease*. The resistant *pi21* allele has been identified through NGS approach from japonica rice lines, and can be used to improve blast resistance of rice worldwide without any lingkage drag (Fukuoka *et al.,* 2009).

Genomics-based genotyping system not only reduces the number of breeding cycles but also precisely integrate target genes for particular traits into an ideal genetic background. Twenty-eight genes of important traits were detected in rice using whole genome based SNP array, RICE6K, which is a (Yu *et al.,* 2014). Results also detected 12 SNPs per 1 Mb and provided more intensive information about polymorphisms between *indica* and *japonica* subspecies as well as varieties within *indica* and *japonica* groups. It showed that SNP chip RICE6K is suitable for rice germplasm fingerprinting, functional allele detection, genetic background selection among breeding lines and considered that this genotyping technique can be used reliably in rice genomic breeding. The RICE6K was developed using four million SNPs identified from resequencing results of 500 rice germplasm. Whole genome sequencing results (Huang *et al.,* 2012) demonstrate that *O. sativa* has been domesticated from a single origin of *O. rufipogon.* Many other research reports supporting the view that cultivated rice *Oryza sativa* has been developed from its close wild relatives *O. rufipogon* and *O. nivara* based on genome information (McNally *et al.,* 2009; Zhao *et al.,* 2010, 2011; Molina *et al.,* 2011).

The NGS based GWAS mapping has created a new avenues to accelerate the mining of diverse germplasm to identify important functional alleles (Zhang *et al.,* 2008; Ikeda *et al.,* 2013). Identified functional alleles may help to design new idiotype super rice by combining heterosis vigour between *indica-japonica* (Guo and Ye, 2014) subspecies and it can be used to develop 'Green Super Rice' (GSR) using genomic selection (Zhang, 2007). Rice breeder can utilize the genomic knowledge including DNA sequences and gene functions to create new genotype and control the selection procedure to modify the whole genomic information to improve the varieties through the advent of genomic technologies (*i.e.,* Genomic Breeding). In genomic breeding two types of high-throughput genotyping markers are used, DNA sequencing and DNA array (Davey *et al.,* 2011; Gupta *et al.,* 2008). In genomic breeding, target gene can be selected based on molecular markers and genetic background selection can be achieved using genome-wide DNA polymorphism analyses (Yu *et al.,* 2013). Whole-genome based SNP array (RICE6K) has been used in genomic breeding for fingerprinting the rice germplasm, selection of genetic background

of the progeny lines and target gene introgression (Yu *et al.*, 2013). Using this RICE6K array some functional alleles of seven genes have been detected such as Sd1 (plant height), Gn1a for grain number, grain size gene GW2, plant architecture gene TAC1 and hybrid fertility gene (S5 and Sa) (Yu *et al.*, 2013); rice germplasm characterization based on DNA array (McNally *et al.*, 2009; Wang *et al.*, 2010), GWAS based SNP-genotyping is carried out in rice using 44K SNP array (McCouch *et al.*, 2010; Zhao *et al.*, 2011); Illumina GoldenGate SNP Chip is used in detects SNPs and genetic analysis in rice breeding (Thomson *et al.*, 2012). Whole genome sequencing (WGS) works have been done in many accessions of rice for genetic polymorphism analysis (Arai-Kichise *et al.*, 2014; Lyu *et al.*, 2013; Xu *et al.*, 2012; Yang *et al.*, 2012' Duitama *et al.*, 2015 and 3K RGP). Genetic diversity within *O. sativa* (McNally *et al.*, 2009), integration of important traits in the HYVs (Zhao *et al.*, 2010), and identification of many genes related to complex traits (*QTLs*) has been conducted by the rice scientists using GWAS/SNP chip (Zhao *et al.*, 2011) and aluminium tolerance traits in rice has also been studied in details (Famoso *et al.*, 2011). Yamamoto *et al.* (2012) has developed online data base (OGRO) of rice by incorporating all the functional alleles and *QTLs* to facilities the rice breeding program world-wide. Based on the SNP diversity, it was observed overall (pairwise SNP differences per kb) variation 3.93, and on average 2.58 within *indica* subspecies, 1.96 within *japonica* and 5.9 between *indica* and *japonica* (Xu *et al.*, 2012; Duitama *et al.*, 2015). Supporting the earlier report that subspecies *indica* has more genetic diversity than japonica.

It is obvious to increase approximately 25% rice grain of the present production to meet up the demand of population growth in 2030 from less amount of arable land, less water and under adverse effects of climatic change (Flood/drought/salinity). Estimated annual yield increase is needed about 1.2–1.5%, which means yield increase of 0.6 t/ ha world-wide on an average (Seck *et al.*, 2012). Population growth will reach nine billion (9 Billion) by 2050, and it is imperative to increase the food supply to meet up the demands (Godfray *et al.*, 2010).

Genomic Assisted Breeding for Rice Crop Improvement

Molecular genetics can play a vital role for the improvement of rice yield and productivity in the post-genomics era to sustain world food security (Miura *et al.* 2011; Huang *et al.* 2013). Grain yield enhancing processes are the main target to increase yield in the rice breeding and genomic research (Huang *et al.*, 2013; Zuo and Li, 2014; Li *et al.*, 2011). Advanced genomic tools are being used for study the mapping population and functional genomics (SNP based marker-trait association) analyses to dissect the unknown complex traits (*QTLs*) to uncover the *QTLs*/genes (*GS3, GW2, qSW5/GW5, GS5, TGW6* and *GIF1*) for grain size/weight in rice (Fan *et al.* 2009; Zhang *et al.* 2012; Weng *et al.*, 2008; Zhao *et al.*, 2011; Huang *et al.*, 2010). The GWAS have successfully identified many *QTLs*/gene for agronomic traits, including grain yield (Li *et al.*, 1997; Venuprasad *et al.*, 2012), flowering time (Chen *et al.*, 2014), plant height (Ashikari *et al.*, 2002), aluminum tolerance (Famosa *et al.*, 2011), grain yield under drought stress, and submergence tolerance (Xu and Mackill, 1996) in rice. Genotyping-by-sequencing (GBS) was used to discover and call SNPs on 369 advanced inbred breeding lines of rice for grain yield (kg/ha), flowering time (days to 50% flowering), and plant height (cm) (Spindel *et al.*, 2015).

Plant height specific *QTLs* were identified based on genomic breeding (sequencing based) in rice (Huang *et al.*, 2009), which was considered as 'Green revolutionary' gene. Whole genome sequencing of rice cultivar is performed to identify gene specific markers (may be SNPs) which may help in marker assisted selection (MAS) in breeding program. Amylose content determining gene specific SNPs were detected in rice varieties using whole genome re-sequencing method and discovered three SNPs (Waxy-1, Waxy-2 and Waxy-3) (Duitama *et al.*, 2015) for quality improvement and reported as amylose

content marker in rice by other group (Larkin and Park, 2003). Many gene/ *QTLs* were identified using cloning and functional analysis of the genes based on comparative genome characterization (Weng *et al.*, 2008; Jiang *et al.*, 2012). The *QTL* related to grain yield DEP1 was characterized and responsible for grain number per panicle and for erect panicle (Huang *et al.*, 2009). GS3 is major *QTL* related to grain weight and size was identified by Fan et al (2006). Heading date *QTL* Ghd7 and days to heading (DTH8) were identified in rice (Xue *et al.*, 2008; Wei *et al.*, 2010; Huang *et al.*, 2013). Wild species *O. rufipogon* has also been resequenced for the generation of functional SNPs, which can be used to improve the crop varieties and relationship between rice diversity and domestication of rice crop (He *et al.*, 2011; Huang *et al.*, 2012a). The 'Oryza Map Alignment Project' (OMAP) was developed for alignment of sequencing data of wild rice based on reference genome to identify genes and *QTLs* (Wing *et al.*, 2005).

RIL lines are re-sequenced using the NGS technique and ultra-high-density linkage map has been constructed for the identification of many *QTLs* associated with yield and gene (DTH8 and LAX1). The fine mapping of *QTLs* and genes can accelerate the *QTL* cloning and molecular breeding to develop improved rice varieties (Gao *et al.*, 2013). Resequencing based gene mapping has been applied to identify some important gene/*QTLs* (qPH1, qPH5, qGL-3, qGN1) from the rice lines (Wang *et al.*, 2011; Gao *et al.*, 2013; Yu *et al.*, 2011; Duan *et al.*, 2013). The International Rice Functional Genomics Project (IRFGP) has been initiated to determine the function of each of the alleles in the rice genome (Ikeda *et al.*, 2013). Bacterial blight (BB) is one of the most devastating diseases of rice which causes huge loss in rice production worldwide and caused by *Xanthomonus oryzae* pv. *oryzae* (*Xoo*). Different approaches have been applied to make BB resistant rice varieties through molecular breeding. Some BB resistance genes (*Xa3*/*Xa26*, *Xa4*, *Xa4b*, *Xa6*, and *Xa9*) have been transferred to improved varieties through MAS breeding. These genes are located at the end of the long arm of chromosome 11. The receptor-like kinase gene was found in the cloned region of *Xa3*/*Xa26* gene (Sun *et al.*, 2004) based on GWAS.

The genetic improvement of Basmati rice for yield, quality and resistance to bacterial leaf blight (*Xa21*, *xa4*, *xa13*, *xa5*, *Xa33t*, *xa34t* and *Xa38*) and blast (*Pi1*, *Pi2*, *Pi5*, *Pi9*, *Pi54*, *Pib*, *Piz*, *Piz5*, *Pi-ta* and *Pi54*/*Pi-K^{h}*), brown plant hopper [*Bph-3/17/18/20/21* and *Bph18(t)*], sheath blight (*qSHB*) and gall midge (*Gm4* and *Gm8*) diseases has been performed by pyramiding the multiple genes/*QTLs* through marker-assisted back-crossing (MABC)/marker-assisted foreground and background selection (Cheema *et al.*, 2008; Madhavi *et al.*, 2011; Natarajkumar *et al.*, 2012; Sujatha *et al.*, 2013; Pandey *et al.*, 2013; Pradhan *et al.*, 2015). Introgression of known cloned genes and *QTLs* for drought tolerance (*DTY1.1*, *DTY2.1*, *DTY2.2*, *DTY3.1*, *DTY3.2*, *DTY9.1* and *DTY12.1*), flood tolerance (sub-mergence) (*Sub1*) and salinity stress (*Saltol*) tolerance are initiated to develop high-yielding mega rice varieties (ADT46, Bahadur, MTU1075, Pooja, Rajendra, Mahsuri, Ranjit, ADT39, Pusa44, ADT45, Gayatri and Savitri) of India through MAS (http://india.irri.org/mega-projects-in-india, Singh *et al.* 2015). This eventually may lead to development of certain diverse genetically-tailored high-yielding and climate resilient early maturing Indian rice varieties for sustaining food security.

Furthermore, millions of single nucleotide polymorphism (SNP) and insertion-deletion (InDel) markers have already been identified in rice. Saturation of the genome with such sequence based SSR and SNP markers is accelerating fine mapping and map-based cloning of genes, and thus, development of gene-based allele-specific markers. Rice improvement programs are expected to benefit greatly from the use of these markers in near future. The availability of gold standard reference genome sequence of *japonica* rice cv. Nipponbare (International Rice Genome Sequencing Project 2005) has propelled the genome resequencing and transcriptome sequencing of diverse rice

genotypes in recent years by use of NGS (next-generation sequencing) approaches in India. This in turn led to the development of enormous resources in the form of genomic (genic) simple sequence repeat (SSR) and single nucleotide polymorphism (SNP) markers at a genome-wide scale in rice. For instance, non-redundant 2495052 SNP and 324034 InDel markers have been discovered by comparing the NGS-based whole-genome resequencing data of six elite *indica* inbred lines (three of each cytoplasmic male sterile and restorer lines) to accelerate genomics-assisted breeding for hybrid performance in rice (Subbaiyan *et al.,* 2012). Subsequently, the whole genome resequencing of three drought/salinity tolerant (Nagina 22 and Pokkali) and sensitive (IR64) rice accessions identified non-redundant 1784583 SNPs and 154275 InDels between reference Nipponbare and three resequenced rice accessions. Based on this outcome, genome-wide 401683 SNPs between IR64 and Pokkali and 662509 SNPs between IR64 and Nagina 22 that are well-distributed across coding and non-coding regions of these sequenced genomes were discovered with the eventual aim to deploy them in marker-assisted breeding for abiotic stress tolerance in rice (Jain *et al.,* 2014). More recently, the comparison of whole genome resequencing data of a widely cultivated low glycemic index-containing *indica* rice variety, Swarna, with reference genome Nipponbare, identified 1,149,698 SNPs (65,984 non-synonymous SNPs) and 104,163 InDels for deciphering the genetic basis of complex glycemic index quantitative trait in rice (Rathinasabapathi *et al.,* 2015).

References

Abe, A., et al, (2012). Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat. Biotechnol.*, 30: 174-178.

Agarwal, P., Parida, S.K., Raghuvanshi, S., Kapoor, S., Khurana, P., Khurana, J.P., and Tyagi, A.K., (2016). Rice improvement through genome-based functional analysis and molecular breeding in India. *Rice,* 9: 1-17.

Andersen, J.R., and Lübberstedt, T., (2003). Functional markers in plants. *Trends Plant Sci,* 8: 554-560.

Arai-Kichise, Y., Shiwa, Y., Ebana, K., Shibata-Hatta, M., Yoshikawa, H., Yano, M., *et al.* (2014). Genome-wide DNA polymorphisms in seven rice cultivars of temperate and tropical Japonica Groups. *PLoS One,* 9: e86312. doi: 10.1371/journal.pone.0086312.

Ashikari, M., Sasaki, A., Ueguchi-Tanaka, M., Itoh, H., Nishimura, A., *et al.* (2002) Loss-of-function of a rice gibberellin biosynthetic gene, GA20 oxidase (GA20ox-2), led to the rice 'green revolution'. *Breeding Science* 52: 143-150.

Bandillo, N., Raghavan, C., Muyco, P.A., Sevilla, M.A., Lobina, I.T., *et al.* (2013). Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding. *Rice,* 6: 11.

Bernardo, R. (2008). Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Sci,* 48: 1649-1664.

Cheema, K., Grewal, N., Vikal, Y., Sharma, R., Lore, J.S., Das, A., Bhatia, D., Mahajan, R., Gupta, V., Bharaj, T.S., and Singh, K. (2008). A novel bacterial blight resistance gene from *Oryza nivara* mapped to 38 kb region on chromosome 4 L and transferred to *Oryza sativa* L. *Genet Res,* 90: 397-407.

Chen, J.B., Li, X.Y., Cheng, C., Wang, Y.H., Qin, M., *et al.* (2014). Characterization of Epistatic Interaction of *QTLs* LH8 and EH3 Controlling Heading Date in Rice. *Scientific Reports,* doi: 10.1038/srep07387.

Davey, J.W., Paul, A., Hohenlohe, Paul, D.E., Boone, J.Q., Catchen, J.M., and Blaxter, M.L., (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Rev Genet,* 12: 499-510.

Duan, M.J., Sun, Z.Z., Shu, L.P., Tan, Y.N., Yu, D., Sun, X.W., Liu, R.F., Li, Y.J., Gong, S.Y., Yuan, D.Y. (2013). Genetic analysis of an elite super-hybrid rice

parent using high-density SNP markers. *Rice,* 6: 21-27.

Duitama, J., Silva, A., Sanabria, Y., Cruz, D.F., Quintero, C., Ballen, C., *et al.* (2015). Whole Genome Sequencing of Elite Rice Cultivars as a Comprehensive Information Resource for Marker Assisted Selection. *PLoS ONE,* 10(4): e0124617.

Famoso, A.N., Zhao, K., Clark, R.T., Tung, C.W., Wright, M.H., Bustamante, C., et al. (2011). Genetic architecture of aluminum tolerance in rice (*Oryza sativa*) determined through genome-wide association analysis and *QTL* mapping. *PLoS Genet,* 8: e1002221.

Fan, C. et al. (2006) GS3, a major *QTL* for grain length and weight and minor *QTL* for grain width and thickness in rice, encodes a putative transmembrane protein. *Theor Appl Genet,* 112: 1164-1171.

Chuchuan, F., Xing, Y., Mao, H., Lu, T., Han, B., Xu, C., Li, X., and Zhang, Q. (2009). 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,* 118: 465-472.

Feltus F.A. *et al.* (2004). An SNP resource for rice genetics and breeding based on subspecies *indica* and japonica genome alignments. *Genome Res.,* 14: 1812-1819.

Fukuoka, S., Saka, N., Koga, H., Ono, K., Shimizu, T., et al. (2009). Loss of function of a proline-containing protein confers durable disease resistance in rice. Science 325: 998–1001

Ganal MW, Altmann T, Röder MS (2009). SNP identification in crop plants. *Curr Opin Plant Biol,* 12: 211-217.

Gao, Q., Yue, G.D., Li, W.Q., Wang, J.Y., Xu, J.H., and Yin, Y. (2012). Recent progress using high-throughput sequencing technologies in plant molecular breeding. *J Integr plant Biol,* 54: 215-227.

Gao, Z.Y., Zhao, S.C., He, W.M., Guo, L.B., Peng, Y.L., Wang, J.J., Guo, X.S., Zhang, X.M., Rao, Y.C., Zhang, C., Dong, G.J., Zheng, F.Y., Liu, Y., Zhu, L., et al. (2013). Dissecting yield-associated loci in super hybrid rice by resequencing recombinant inbred lines and improving parental genome sequences. *Proc Natl Acad Sci USA,* 110: 14492–14497.

Godfray, H.C. et al. (2010). Food security: the challenge of feeding 9 billion people. Science 327: 812-818.

Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R.L., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley, D., Hutchison, D., Martin, C., Katagiri, F., Lange, B.M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J.P., Miguel, T., Paszkowski, U., Zhang, S.P., Colbert, M., Sun, W.L., Chen, L.L., Cooper, B., Park, S., Wood, T.C., Mao, L., Quail, P., et *al.* (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *Japonica). Science* 296: 92-100.

Gopalakrishnan, S., Sharma, R.K., Anand Rajkumar, K.A., Joseph, M., Singh, V.P., Bha, K.V., Singh, N.K., and Mohapatra, T. (2008). Integrating marker assisted background analysis with foreground selection for identification of superior bacterial blight resistant recombinants in Basmati rice. *Plant Breed,* 127: 131-139.

Guo, L.B., and Ye, G.Y. (2014). Use of major quantitative trait loci to improve grain yield. *Rice,* 7: 1-10.

Gupta, P.K., Rustgi, S., and Mir, R.R. (2008). Array-based high-throughput DNA markers for crop improvement. *Heredity,* 101: 5-18.

Han, B., and Huang, X.H. (2013). Sequencing-based genome-wide association study in rice. *Curr Opin Plant Biol,* 7: 133-138.

Heffner, E.L., Sorrells, M.E., Jannink, J.-L. (2009). Genomic selection for crop improvement. *Crop Sci,* 49: 1-12.

Huang, R., Jiang, L., Zheng, J., Wang, T., Wang, H., Huang, Y., and Hong, Z. (2013). Genetic bases of rice grain shape: so many genes, so little known. *Trends Plant Sci,* 18: 218-226.

Huang, X. et al. (2009). Natural variation at the DEP1 locus enhances grain yield in rice. *Nat Genet,* 41: 494-497.

Huang, X., Kurata, N., Wei, X., Wang, Z., Wang, A., Zhao, Q., Zhao, Y., Liu, L., Lu, H., Li, W., Guo, Y., Lu, Y., Zhou, C., Fan, D., Weng, Q., Zhu, C., Huang, T., Zhang, L., Wang, Y., Feng, L., Furuumi, H., Kubo, T., Miyabayashi, T., Yuan, X., Xu, Q., Dong, G., Zhan, Q., Li, C., Fujiyama, A., and Toyoda, A. (2012). A map of rice genome variation reveals the origin of cultivated rice. *Nature,* 490: 497-501.

Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., Li, C., Zhu, C., Lu, T., Zhang, Z., Li, M., Fan, D., Guo, Y., Wang, A., Wang, L., Deng, L., Li, W., Lu, Y., Weng, Q., Liu, K., Huang, T., Zhou, T., Jing, Y., Lin, Z., Buckler, E.S., Qian, Q., Zhang, Q.F., Li, J., and Han, B. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat Genet,* 42: 961-967.

Huang, X.H., Feng, Q., Qian, Q., Zhao, Q., Wang, L., Wang, A., Guan, J.P., Fan, D.L., Weng, Q.J., Huang, T., Dong, G.J., Sang, T., and Han, B. (2009). High-throughput genotyping by whole-genome resequencing. *Genome Res,* 7: 1068-1076.

Huang, X.H., Lu, T.T., and Han, B. (2013). Resequencing rice genomes: an emerging new era of rice genomics. *Trends Genet,* 29: 225-230.

Huang, X.H., Zhao, Y., Wei, X.H., Li, C.Y., Wang, A., Zhao, Q., Li, W.J., Guo, Y.L., Deng, L.W., Zhu, C.R., Fan, D.L., Lu, Y.Q., Weng, Q.J., Liu, K.Y., Zhou, T.Y., Jing, Y.F.S.L.Z., Dong, G.J., Huang, T., Lu, T.T., Feng, Q., Qian, Q., Li, J.Y., and Han, B. (2012). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat Genet,* 44: 32-41.

Ikeda, M., Miura, K., Aya, K., Kitano, H., and Matsuoka, M. (2013). Genes offering the potential for designing yield-related traits in rice. *Curr Opin Plant Biol,* 16: 213-220.

Imai, I., Kimball, J.A., Conway, B., Yeater, K.M., McCouch, S.R. et al. (2013). Validation of yield-enhancing quantitative trait loci from a low-yielding wild ancestor of rice. *Mol Breed,* 32: 101-120.

IRGSP (2005) The map-based sequence of the rice genome. *Nature,* 7: 793-800.

Jain, M., Moharana, K.C., Shankar, R., Kumari, R., and Garg, R. (2014). Genomewide discovery of DNA polymorphisms in rice cultivars with contrasting drought and salinity stress response and their functional relevance. *Plant Biotechnol J,* 12: 253-264.

Jannink, J.-L., Aaron, J., Lorenz, and Iwata, H. (2010). Genomic selection in plant breeding: from theory to practice. *Briefings in functional genomics,* 9: 166-177.

Jeong, I.S., Yoon, U.H., Lee, G.S., Ji, H.S., Lee, H.J., Han, C.D., Hahn, J.H., An, G.H., and Kim, Y.H. (2013). SNP-based analysis of genetic diversity in anther-derived rice by whole genome sequencing. *Rice,* 7: 6. doi: 10.1186/1939-8433-6-6.

Jiang, Y. et al. (2012). Rice functional genomics research: progress and implications for crop genetic improvement. *Biotechnol Adv,* 30: 1059-1070.

Kanamori, H., Fujisawa, M., Katagiri, S., Oono, Y., Fujisawa, H., Karasawa, W., Kurita, K., Sasaki, H., Mori, S., Hamada, M., Mukai, Y., Yazawa, T., Mizuno, H., Namiki, N., Sasaki, T., Katayose, Y., Matsumoto, T., and Wu, J. (2013). A BAC physical map of *aus* rice cultivar 'Kasalath', and the map-based genomic sequence of 'Kasalath' chromosome 1. *Plant J,* 7: 699-708.

Larkin, P.D., and Park, W.D. (2003). Association of waxy gene single nucleotide polymorphisms with starch characteristics in rice (*Oryza sativa* L.). *Molecular Breeding,* 12: 335-339.

Li, Y., Fan, C., Xing, Y., Jiang, Y., Luo, L., Linag, S., Shao, D., Xu, C., Li, X., Xiao, J., He, Y., Zhang, Q. (2011). Natural variation in GS5 plays an important role in regulating grain size and yield in rice. *Nature Genetics,* 43: 1266-1269.

Li, Z., Shannon, P.R., William, P.D., Paterson, A.H., and Stansel, J.W. (1997). Epistasis for three grain yield components in rice. *Genetics,* 145: 453-465.

Lyu, J., Zhang, S., Dong, Y., He, W., Zhang, J., Deng, X., et al. (2013). Analysis of elite variety tag SNPs reveals an important allele in upland rice. *Nat Commun,* 4: 2138. doi: 10.1038/ncomms3138.

Madhavi, K.R., Prasad, M.S., Laha, G.S., Mohan, K.M., Madhav, M.S., and Viraktamath, B.C. (2011). Combining blast and bacterial blight resistance in rice cultivar, improved Samba Mahsuri. *Indian J Plant Protect,* 39: 124-129.

McCouch, S.R., McNally, K.L., Wang, W., and Sackville, H.R. (2012). Genomics of gene banks: A case study in rice. *Am J Bot,* 99: 407-423.

McCouch, S.R., Zhao, K., Wright, M., Tung, C.W. Ebana, K., Thomson, M., Reynolds, A., Wang, D., DeClerck, G., Ali M.L. (2010). Development of genome-wide SNP assays for rice. *Breed. Sci,* 60: 524-535.

McNally, K.L., Childs, K.L., Bohnert, R., Davidson, R.M., Zhao, K., Ulat, V.J. et al. (2009). Genome wide SNP variation reveals relationships among landraces and modern varieties of rice. *Proc Natl Acad Sci USA,* 106: 12273-12278.

Meuwissen, T.H.E., Hayes, B.J., and Goddard, M.E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157: 1819-1829.

Mishra, K.K., Vikram, P., Yadaw, R.B., Swamy, B.P.M., Dixit, S. et al. (2013). qDTY12.1: a locus with a consistent effect on grain yield under drought in rice. *BMC Genetics* 14: 12-17.

Miura, K., Ashikari, M., and Matsuoka, M. (2011). The role of *QTLs* in breeding of high-yielding rice. *Trends Plant Sci,* 16: 319-326.

Molina, J., Sikora, M., Garud, N., Flowers, J.M., Rubinstein, S., Reynolds, A. et al. (2011). Molecular evidence for a single evolutionary origin of domesticated rice. *Proc Natl Acad Sci USA,* 108: 8351-8356.

Natarajkumar, P., Sujatha, K., Laha, G.S., Rao, K.S., Mishra, B., Viraktamath, B.C., Hari, Y., Reddy, C.S., Balachandran, S.M., Ram, T., Madhav, M.S., Rani, N.S., Neeraja, C.N., Reddy, G.A., Shaik, H., and Sundaram, R.M. (2012). Identification and fine-mapping of *Xa33,* a novel gene for resistance to *Xanthomonas oryzae* pv. *oryzae. Phytopathology,* 102: 222-228.

Pandey, M.K., Rani, N.S., Sundaram, R.M., Laha, G.S., Madhav, M.S., Rao, K.S., Sudharshan, I., Hari, Y., Varaprasad, G.S., Rao, L.V.S., Suneetha, K., Sivaranjani, A.K.P., Viraktamath, B.C. (2013). Improvement of two traditional Basmati rice varieties for bacterial blight resistance and plant stature through morphological and marker-assisted selection. *Mol Breed,* 31: 239-246.

Pérez-de-Castro, A.M., Vilanova, S., Cañizares, J., Pascual, L., Blanca, J., Díez, M.J., Prohens, J., and Picó, B. (2012). Application of Genomic Tools in Plant Breeding. *Curr Genomics,* 13: 179-195.

Pradhan, S.K., Nayak, D.K., Mohanty, S., Behera, L., Barik, S.R., Pandit, E., Lenka, S., and Anandan, A. (2015). Pyramiding of three bacterial blight resistance genes for broad-spectrum resistance in deepwater rice variety, Jalmagna. *Rice,* 8: 19-23.

Rathinasabapathi, P., Purushothaman, N., Ramprasad, V.L., and Parani, M. (2015). Whole genome sequencing and analysis of Swarna, a widely cultivated *indica* rice variety with low glycemic index. *Sci Rep,* 5: 11303.

Scott, A.J. (2016) Rice: The First Crop Genome. *Rice,* 9: 14.

Seck, P.A., Diagne, A., Mohanty, S., and Wopereis, C.S. (2012). Crops that feed the world 7: Rice. *Food Sec,* 4: 7-24.

Septiningsih, E.M., Pamplona, A.M., Sanchez, D.L., Neeraja, C.N., Vergara, G.V., et al. (2009). Development of submergence-tolerant rice cultivars: the *Sub1* locus and beyond. *Ann Bot,* 103: 151-160.

Singh, B.P., Jayaswal, P.K., Singh, B., Singh, P.K., Kumar, V., Mishra, S., Singh, N., Panda, K., and Singh, N.K. (2015). Natural allelic diversity in *OsDREB1F* gene in the Indian wild rice germplasm led to ascertain its association with drought tolerance. *Plant Cell Rep,* 34: 993-1004.

Song, W.Y., Wang, G.L., Chen, L.L., Kim, H.S., Pi, L.Y., Holsten, T., Gardner, J., Wang, B., Zhai, W.X., Zhu, L.H. et al. (1995). A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. *Science,* 270: 1804-1806.

Spindel, J., Begum, H., Akdemir, D., Virk, P., Collard, B., Redoña, E., Atlin, G., Jannink, J.L., and McCouch, S.R. (2015). Genomic selection and association mapping in rice (*Oryza sativa*): effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. *PLoS Genet* 11(2): e1004982.

Subbaiyan, G.K., Waters, D.L., Katiyar, S.K., Sadananda, A.R., Vaddadi, S., and Henry, R.J. (2012). Genome-wide DNA polymorphisms in elite *indica* rice inbreds discovered by whole-genome sequencing. *Plant Biotechnol J,* 10: 623-634.

Sujatha, K., Prasad, M.S., Pandey, M., Ramesha, M.S., Neeraja, C.N., Balachandran, S.M., Shobharani, N., Kemparaju, B., Madhanmohan, K., Sama, V.S.A.K., Hajira, S.K., Baachiranjeevi, C.H., Pranathi, K., Reddy, G.A., Madhav, M.S., and Sundaram, R.M. (2013). Marker-assisted introgression of bacterial blight and blast resistance into IR 58025B, an elite maintainer line of rice. *Plant Breed,* 132: 586-594.

Su, X. et al. (2004). *Xa26,* a gene conferring resistance to *Xanthomonas oryzae* pv. *oryzae* in rice, encodes an LRR receptor kinase-like protein. *Plant J,* 37: 517-527.

Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., Uemura, A., Utsushi, H., Tamiru, M., Takuno, S., Innan, H., Cano, L.M., Kamoun, S., and Terauchi, R. (2013). *QTL*-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J,* 7: 174-183.

Takagi, H., Uemura, A., Yaegashi, H., Tamiru, M., Abe, A. et al. (2013). MutMap-Gap: whole-genome resequencing of mutant F_2 progeny bulk combined with de novo assembly of gap regions identifies the rice blast resistance gene *Pii. New Phytol,* 200: 276-283.

The 3000 rice genomes project (2014). GigaScience 3: 7 doi: 10.1186/2047-217X-3-7

Thomson, M.J., Zhao, K., Wright, M., McNally, K.L., Rey, J., Tung, C.-W., Reynolds, A., Scheffler, B., Eizenga, G., and McClung, A. (2012). High-throughput single nucleotide polymorphism genotyping for breeding applications in rice using the BeadXpress platform. *Mol. Breed,* 29: 875-886.

Uga, Y., Sugimoto, K., Ogawa, S., Rane, J., Ishitani, M. et al. (2013). Control of root system architecture by *DEEPER ROOTING 1* increases rice yield under drought conditions. *Nature Genet,* 45: 1097-1102.

Varshney, R.K., Terauchi, R., and McCouch, S.R. (2014). Harvesting the Promising Fruits of Genomics: Applying Genome Sequencing Technologies to Crop Breeding. *PLOS Biology,* 12: e1001883.

Varshney, R.K., Graner, A., and Sorrells, M.E. (2005). Genomics-assisted breeding for crop improvement. *Trends Plant Sci,* 10: 621-630.

Venuprasad, R., Bool, M.E., Aulatchon, L., Sta Cruz, M.T., Amante, M. et al. (2012). A large-effect *QTL* for rice grain yield under upland drought stress on chromosome 1. *Mol Breeding,* 30: 535-547.

Wei, X. et al. (2010). DTH8 suppresses flowering in rice, influencing plant height and yield potential simultaneously. *Plant Physiol,* 153: 1747-1758.

Weng, J., Gu, S., Wan, X., Gao, H., Guo, T., Su, N., Lei, C., Zhang, X., Cheng, Z., Guo, X., Wang, J., Jiang L., Zhai, H., Wan, J. (2008). Isolation and initial characterization of GW5, a major *QTL* associated with rice grain width and weight. *Cell Res,* 18: 1199-1209.

Xie, W., Feng, Q., Yu, H., Huang, X., Zhao, Q., Xing, Y., Yu, S., Han, B., and Zhang,

Q. (2010). Parent-independent genotyping for constructing an ultrahigh-density linkage map based on population sequencing. *Proc Natl Acad Sci USA*, 107: 10578-10583.

Xu, J.J., Zhao, Q., Du, P.N., Xu, C.W., Wang, B.H., Feng, Q., Liu, Q.Q., Tang, S.Z., Gu, M.H., Han, B., and Liang, G.H. (2010). Developing high throughput genotyped chromosome segment substitution lines based on population whole-genome re-sequencing in rice. *BMC Genomics* 7: 656-669.

Xu, K., Ismail, A., and Ronald, P. (2013). Flood tolerance mediated by the rice Sub1A transcription factor 1. *In*: Jenks, M., and Hasegawa, P. (eds.) Plant abiotic stress, 2nd edition. Wiley-Blackwell, Ames, Iowa.

Xu, K.N., and Mackill, D.J. (1996). A major locus for submergence tolerance mapped on rice chromosome 9. *Molecular Breeding*, 2: 219-224.

Xu, X., Liu, X., Ge, S., Jensen, J.D., Hu, F.Y., Li, X., Dong, Y., Gutenkunst, R.N., Fang, L., Huang, L., Li, J.X., He, W.M., Zhang, G.J., Zheng, X.M., Zhang, F.M., Li, Y.R., Yu, C., Kristiansen, K., Zhang, X.Q., Wang, J., Wright, M., McCouch, S., Nielsen, R., Wang, J., and Wang, W. (2011). Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat Biotech*, 30: 105-111.

Xu, Y., and Crouch, J.H. (2008). Marker-assisted selection in plant breeding: from publications to practice. *Crop Sci*, 48: 391-407.

Xue, W. et al. (2008). Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. *Nat Genet.* 40: 761-767.

Yamamoto, E., Yonemaru, Ji., Yamamoto, T., and Yano, M. (2012), OGRO: The Overview of functionally characterized Genes in Rice online database. *Rice.* 5: 26.

Yang, C.C., Kawahara, Y., Mizuno, H., Wu, J., Matsumoto, T., and Itoh, T. (2012). Independent Domestication of Asian Rice Followed by Gene Flow from japonica to indica. *Mol Biol Evol*, 29: 1471-1479.

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.

Yu, H.H., Xie, W.B., Wang, J., Xing, Y.Z., Xu, C.G., Li, X.H., Xiao, J.H., and Zhang, Q.F. (2011). Grains in *QTL* detection using an ultra-high density SNP map based on population sequencing relative to traditional RFLP/SSR markers. *PLoS One* 7: e17595.

Yu, J., Hu, S.N., Wang, J., Wong, G.K., Li, S.G., Liu, B., Deng, Y.J., Dai, L., Zhou, Y., Zhang, X.Q., Cao, M.L., Liu, J., Sun, J.D., Tang, J.B., Chen, Y.J., Huang, X.B., Lin, W., Ye, C., Tong, W., Cong, L.J., Geng, J.N., Han, Y.J., Li, L., Li, W., Hu, G.Q., Huang, X.G., Li, W.J., Li, J., Liu, Z.W., Li, L., et al. (2002). A draft sequence of the rice genome (*Oryza sativa* L. *ssp. indica*). *Science*, 296: 79-92.

Zhang, Q. (2007). Strategies for developing Green Super Rice. *Proc Natl Acad Sci USA*, 104: 16402-16409.

Zhang, Q., Li, J., Xue, Y., Han, B., and Deng, C.W. (2008). Rice 2020: a call for an international coordinated effort in rice functional genomics. *Mol Plant* 1: 715-719.

Zhang, X., Wang, J., Huang, J., Lan, H., Wang, C., Yin, C., Wu, Y., Tang, H., Qian, Q., Liand, J., and Zhang, H. (2012). Rare allele of *OsPPKL1* associated with grain length causes extra-large grain and a significant yield increase in rice. *Proc Natl Acad Sci USA*, 109: 21534-21539.

Zhao, K., Tung, C.W., Eizenga, G.C., Wright, M.H., Ali, M.L., Price, A.H., et al. (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat Commun*, 2: 467-471.

Zhao, K., Wright, M., Kimball, J., Eizenga, G., McClung, A., Kovach, M., 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.

Zhao K., Tung C.W., Eizenga, G.C., Wright M.H., Ali, M.L., Price, A.H., Norton, G.J.,

Islam, M.R., *et al.* (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat Commun*, 2: 467-471.

Zuo, J., and Li, J. (2014). Molecular genetic dissection of quantitative trait loci regulating rice grain size. *Annu Rev Genet*, 48: 99-118.