Allele Mining Strategies: Principles and Utilisation for Blast Resistance Genes in Rice (*Oryza sativa* L.)

Sadegh Ashkani[1,2]*, Mohd Rafii Yusop[1]*, Mahmoodreza Shabanimofrad[3], Amin Azadi[2], Ali Ghasemzadeh[3], Parisa Azizi[1] and Mohammad Abdul Latif [3,4]

[1]Institute of Tropical Agriculture, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
[2]Department of Agronomy and Plant Breeding, College of Agriculture, Yadegar-e-Imam Khomeini (RAH) Shahr-e-Rey Branch, Islamic Azad University, Tehran, Iran
[3]Department of Crop Science, Faculty of Agriculture, Universiti Putra Malaysia, Serdang, Selangor, Malaysia
[4]Bangladesh Rice Research Institute (BRRI), Gazipur 1701, Bangladesh
*Correspondence: ashkani.sadegh@upm.edu.my; mrafii@.upm.edu.my
DOI: http://dx.doi.org/10.21775/cimb.017.057

Abstract
Allele mining is a promising way to dissect naturally occurring allelic variants of candidate genes with essential agronomic qualities. With the identification, isolation and characterisation of blast resistance genes in rice, it is now possible to dissect the actual allelic variants of these genes within an array of rice cultivars via allele mining. Multiple alleles from the complex locus serve as a reservoir of variation to generate functional genes. The routine sequence exchange is one of the main mechanisms of *R* gene evolution and development. Allele mining for resistance genes can be an important method to identify additional resistance alleles and new haplotypes along with the development of allele-specific markers for use in marker-assisted selection. Allele mining can be visualised as a vital link between effective utilisation of genetic and genomic resources in genomics-driven modern plant breeding. This review studies the actual concepts and potential of mining approaches for the discovery of alleles and their utilisation for blast resistance genes in rice. The details provided here will be important to provide the rice breeder with a worthwhile introduction to allele mining and its methodology for breakthrough discovery of fresh alleles hidden in hereditary diversity, which is vital for crop improvement.

Introduction

Rice (*Oryza sativa* L.), which serves as a model monocot species in cereal genome studies, is one of the most critical crops in the developing world. Rice is a significant staple food and supplies the main energy resource for almost 50 % of the world's population (Yu et al., 2002). With the world's population increasing at an alarming rate, estimated to reach 7.5 billion by 2020, rice production ought to rise by 50 % to match the actual growing desire of the ever-increasing figures. Rice varieties with increased yield potential, improved with a top-quality level of resistance to biotic and abiotic stresses, are needed. Disease caused by various fungal, bacterial and viral pathogens is one of the major limiting factors affecting rice production. Rice blast caused by the *Magnaporthe oryzae* pathogen is the most devastating disease of rice (Skamnioti and Gurr, 2009; Helliwell et al., 2013). This disease has been common problem in most rice growing areas and causes considerably huge yield losses in every part of the world, even more than 30 % in large rice producing areas under conducive environmental conditions (Ou, 1985; Correa-Victoria and Zeigler, 1993; Talbot, 2003; Skamnioti and Gurr, 2009). To control this fungal disease, genetic engineering and breeding for resistant varieties are effective and economical methods. Gene pyramiding may provide broad spectrum and durable resistance (Tabien et al., 2002). In general host resistance to *M. oryzae* has been classified into complete resistance conditioned by a single gene and incomplete resistance (partial resistance) conditioned by multiple genes of partial effect (Wang et al., 1994). A variety of terms also have been used to refer to this perceived dichotomy, including horizontal versus vertical, narrow-spectrum versus broad-spectrum and so on. Complete resistance is a qualitative character and race specific controlled by a major gene or resistance (R) genes. Meanwhile, partial resistance is a quantitative character and non-race specific, which is polygenic and controlled by many genes known as quantitative resistance loci (Koizumi, 2007). This resistance which is conferred by quantitative trait loci (QTLs) reduces pathogen reproduction in compatible interactions (Koizumi, 2007). Partial resistance is stable against different pathogenic races of the rice blast fungus; therefore, the use of partial resistance is one of the most promising blast control measures (Zenbayashi et al., 2002). Broad-spectrum resistance is defined as resistance that is effective against two or more pathogen species, and/or many different races within one pathogen species (Wisser et al., 2005; Kou and Wang, 2010). However, if the resistance is highly partial, it can also be controlled by a major gene and is race specific. R gene mediated resistance frequently leads to a rapid and complete inhibition of the pathogen colonisation; however, this resistance is narrow spectrum, meaning that each R gene only recognises pathogen races that carry the corresponding a- virulence (Avr) gene. As a result, R gene-mediated resistance is prone to breakdown due to point mutations, deletion and/or recombination of Avr genes in the pathogen, which leads to disease susceptibility (Bonman, 1992; Dai et al., 2010).

Because of the high level of variability in the blast pathogen population, isolation of many resistance genes as resources to defend against various pathogens is necessary to build long-lasting blast resistance rice varieties. R genes that confer resistance to a broad spectrum of isolate, promise to offer rice with durable blast

resistance (Datta et al., 2002). However, in the practice of resistance breeding, using a single R gene which has a broad resistance spectrum is more effective. Sequencing of the entire rice genome will be used to identify genes conferring important traits in the rice germplasm and to determine the function of unidentified genes in general. Molecular breeding approaches continue to be utilised in many countries to improve resistant and high-yielding commercial rice cultivars. The resistance and valuable alleles identified in wild relatives and landraces of rice plants can be employed to connect rice blast disease and the development of agronomically remarkable cultivars through gene pyramiding. The actual identification of candidate DNA markers linked to resistance genes using fine mapping may well allow rice breeders to efficiently transfer these genes from donor cultivars into new, elite rice cultivars using marker-assisted selection. There is also a need to connect knowledge about genes and gene function to create new productive varieties that are a necessary element of a sustainable food supply for the future. Similar to all living organisms, plants such as rice did accumulate many useful alleles for various agronomic traits. Allele mining is a strategy to identify superior and/or novel alleles based on the available information about the genes. Allele mining offers the avenue for the validation of specific gene(s) responsible for a specific characteristic and mining of good alleles from the rice gene bank. The natural mutation is a great contributor for evolution of new alleles. Mutations like transitions, transvertions and InDels are the basis for this evolution. During the post-genomics era, allele mining in numerous accessions may contribute to genomics exploration for crop enhancement (Varshney et al., 2005). Polymerase chain reaction (PCR) strategies are used to isolate superior alleles of genes from numerous species. Therefore, allele mining requires PCR-based amplification and sequencing of different versions of the genes established in inbred lines, varieties, landraces and wild relatives. Variation in gene sequences that correlates with the phenotypic quality or performance of the accession allows the identification of constructive alleles for future tests. The entire strategy requires the combination of genetic searches and analyses of a carefully selected list of genotypes as well as putative candidate genes that could have particular functional significance (Kumari et al., 2013). Generally, the purpose of this review is to provide an overview of the allele mining strategy and progress in applying this method in rice for *M. oryzae* genes in resistance breeding programs.

Allele Mining and Blast Resistance Genes

The process of identifying alleles of a known gene/locus that are involved in a particular mechanism for any given trait and their variants within other genotypes or identifying novel, superior and beneficial alleles from the germplasm or natural population is known as allele mining. Allele mining is a recently developed technique that promises to dissect naturally occurring allelic variation in candidate genes that regulate important agronomic traits and has potential usage in crop improvement programs. Steps for allele mining are given in Figure 1. Allele mining allows the validation of specific gene(s) responsible for an individual trait and mining of many favourable alleles from the rice

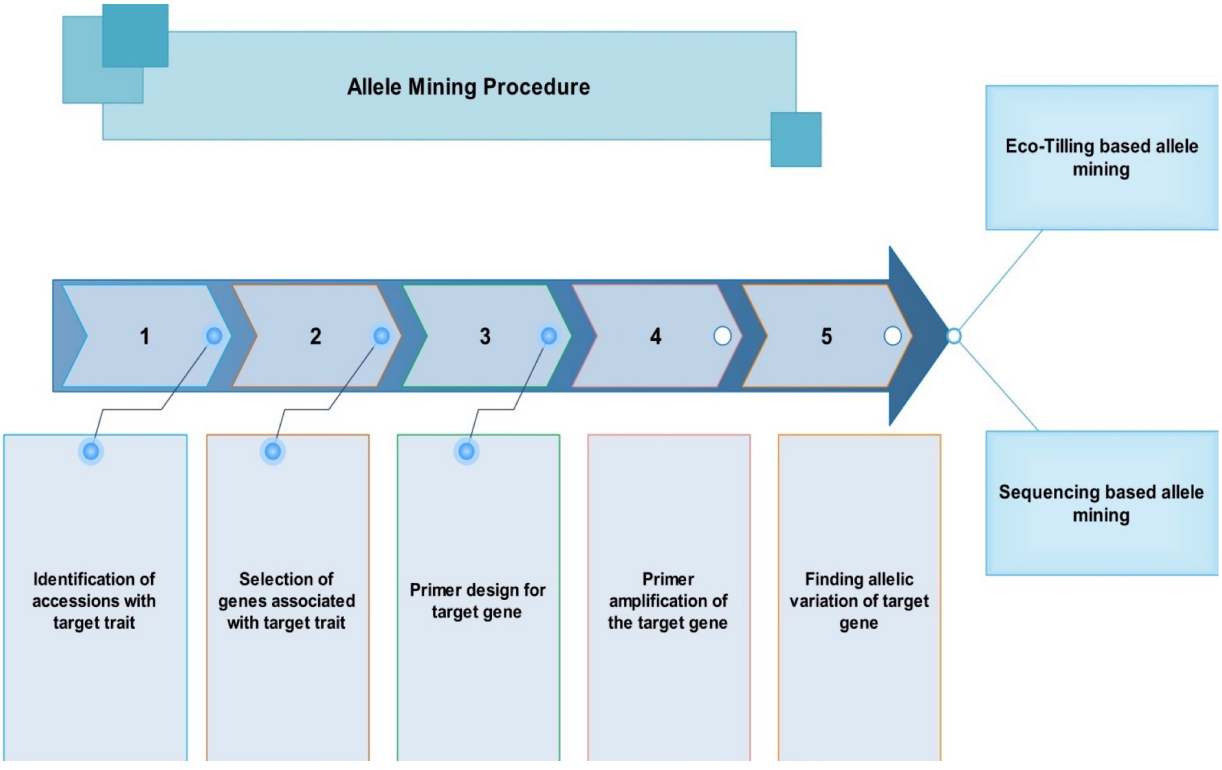


Figure 1. Diagrammatic step involved in allele mining.

gene bank. The DNA bank could be developed to utilise allele mining to determine unique germplasm that contains novel alleles as well as allele combinations. Allele mining has been performed in unique family genes that are involved in a particular mechanism of stress tolerance exhibited by identified accessions (Latha et al., 2004).

To date, approximately 100 blast resistance genes on the rice genome have been mapped (Sharma et al., 2012; Ashkani et al., 2014). From the identified rice blast resistance genes, 22 R-genes, including *Pi-37* (Lin et al., 2007), *Pit* (Hayashi and Yoshida, 2009) and *Pish* (Takahashi et al., 2010) (Chr. 1); *Pib* (Wang et al., 1999) (Chr. 2); *pi21* (Fukuoka et al., 2009) (Chr. 4); *Pi-9* (Qu et al., 2006), *Pi2* (Zhou et al., 2006), *Piz-t* (Zhou et al., 2006), *Pid2* (Chen et al., 2006), *Pid3* (Shang et al., 2009); *Pi25* (Chen et al., 2011) (Chr. 6); *Pi-36* (Liu et al., 2007) (Chr. 8); *Pi5* (Lee et al., 2009) (Chr. 9); *Pb1* (Hayashi et al., 2010); *Pi54 (Pik-h)* (Sharma et al., 2005; Sharma et al., 2010); *Pikm* (Ashikawa et al., 2008), *Pi-k* (Zhai et al., 2011), *Pik-p* (Yuan et al., 2011), *Pia* (Okuyama et al., 2011), *NLS1* (Tang et al., 2011) and *Pi54rh* (Das et al., 2012) (Chr. 11); and *Pita* (Bryan et al., 2000) (Chr. 12) have been successfully cloned and characterised. Some genes very well characterized and many of their alleles have been validated or orthologs have revealed various blast resistance spectra. These types of allelic or orthologous genes might be present with ample allelic variations in rice resources due to the coevolution of rice and the blast pathogen in several rice-developing environments; additionally, these allelic or orthologous genes might even be utilised as more efficient and essential gene resources in the enhancement of rice blast resistance (Hulbert et al., 2001). It is now possible to dissect the naturally occurring allelic variation in these genes from a wide range of germplasm through allele mining. Allele mining of genes from wild and cultivated rice species aims to detect superior alleles for blast resistance (Kumari et al., 2013). R-gene loci have practical alleles, and a series of them provide much stronger and broader resistance against rice blast disease (Yang et al., 2007; Das et al., 2012). Allele mining approaches have been intended to identify superior alleles of rice blast resistance genes such as *Pita* (Yang et al., 2007; Huang et al., 2008; Wang et al., 2008a; Ramkumar et al., 2010), *Pikh* (Ramkumar et al., 2010), *Pi54* (Kumari et al., 2013), and *Pi-2* (Hittalmani et al., 2013) from different cultivated rice varieties and wild species. The blast resistance genes *Pi9, Pi2* and *Piz-t* tend to be alleles from different rice resources while physically on the same gene locus on rice chromosome 6, but their level of resistance spectra can be quite different (Zhou et al., 2006; Zhu et al., 2012). A similar situation was also seen in *Pik, Pik-m, Pik-p, Pi1, Pi54 and Pi54rh* (Ashikawa et al., 2008; Rai et al., 2011b; Yuan et al., 2011; Zhai et al., 2011; Das et al., 2012; Hua et al., 2012), which all reside on the same chromosome locus but are cloned independently from various varieties and show a differential level of resistance spectra to a collection of *M. oryzae* strains (Wang et al., 2009) . Functional orthologs of the rice blast resistance gene *Pid3* have been found primarily in most screened indica types and wild rice varieties (Shang et al., 2009). Lv et al., (2013) isolated an ortholog of *Pid3* from a common wild rice accession A4 (*O. rufipogon*) using sequencing-based allele mining, called *Pid3*-A4, with confirmation of different amount of resistance spectra. Thereafter, Xu et al. (2014) isolated 26 orthologs of *Pid3* from cultivated and wild rice accessions and evaluate their respective blast resistance by gene transformation and blast inoculation. Another study performed to help formulate a good allele-mining set has a heuristic program with SSR genotype information from 4046 rice accessions, and the allele-mining the set established is being tested and verified (Zhao et al., 2010). Examples of allele mining utilisation within rice for blast resistance genes are given in Table 1. Transferring these beneficial, cultivated and wild species-derived alleles to elite genetic backgrounds will improve the characteristic performance of elite varieties in rice (Jinhua et al., 1996; McCouch et al., 2007).

Allelic Variations: Exactly How New Alleles Are Generally Produced?

An important goal of modern crop science is to use nucleotide sequence variation to improve crops. Variation can either be natural, from divergent populations, or induced through treatment with mutagens (Till et al., 2007a). Mutation, among the fundamental forces of development, is usually a sudden and heritable change in a gene that causes existing allelic diversity among organisms and/or any crop species. The generation of mutations in specific genes that can then be assayed for phenotypes (reverse genetics) is a powerful strategy for elucidating gene function and for creating new varieties (Till et al., 2007b). Mutational effects could be beneficial, dangerous, or neutral, depending on the context or even location. Occasionally, the addition of a new allele to a population causes that population to become more competent to survive; sometimes, the addition of a new allele to a population makes it less capable. Generally, for the creation of new alleles, mutations occur in the genic regions of the genome. Mutations may take the form of insertions or deletions, which might be together called InDels, or may alter a single nucleotide inside the genome sequence, resulting in a single-nucleotide polymorphism (SNP). The mutations in coding regions could have an excellent effect on the phenotype through changing the particular encoded healthy protein structure as well as function unlike mutations in noncoding regions of a gene. Allelic variants at presented loci inside a population are usually selected and fixed within a new individual or line to produce new and important cultivars. Allele mining is expected to further increase the speed of the detection of mutations and allows rapid identification and characterisation of allelic variants. Allele mining for blast resistance genes is valuable because it can determine a much greater amount of resistance alleles.

Approaches for Allele Mining

In general, there are two approaches available for allele mining and/or identification of sequence polymorphisms for a given gene in a naturally developing population: (i) modified TILLING (Targeting Induced Local Lesions in Genomes) (Comai et al., 2004), called Eco-TILLING and (ii): Re-sequencing (Huang et al., 2009) or sequencing-based allele mining. A short description of these two methods is given in this section.

Table 1. Examples of Allele Mining in Rice Plants for Blast Resistance Genes

S. No.	Genes /Locus	Alleles	Chr.	Trait	References
1	*Pi-k*	*Pi-k, Pi-k^m, Pi-k^s, Pi-k^k, Pi-k^p, Pi-k^h*	11	Blast resistance	(Kiyosawa and Ando, 1990; Joshi et al., 2009)
2	*Pid3*		6	Blast resistance	(Joshi et al., 2009; Shang et al., 2009; Sharma et al., 2009; Rai et al., 2011a; Lv et al., 2013; Xu et al., 2014)
3	*Pi-ta*(+)		12	Blast resistance	(Yang et al., 2007)
4	*Pi-2*		6	Blast resistance	(Hittalmani et al., 2013)
5	*Pi54*		11	Blast resistance	(Kumari et al., 2013)
6	*Pi-z(t)*		6	Blast resistance	(Sharma et al., 2009)
7	AC134922		11	Blast resistance	(Wang et al., 2014)
8	*Pi-2(t)*	*Pi-z, Piz-5,Pi-zt, Pi-9(t);*	9	Blast resistance	(Kiyosawa and Ando, 1990; Joshi et al., 2009)
9	*Pi-4*	*Pi-4^a(t), Pi-4^b(t), Pi-ta;*	12	Blast resistance	(Kiyosawa and Ando, 1990; Joshi et al., 2009)
10	*Pi-5(t)*	*Pi-3(t), Pi- i(t);*	9	Blast resistance	(Kiyosawa and Ando, 1990; Joshi et al., 2009)
11	*Pi-14(t)*	*Pi-16(t), Pi-d(t);*	12	Blast resistance	(Kiyosawa and Ando, 1990; Joshi et al., 2009)
13	*Pi-ta*	*Pi-ta , Pi-ta^2*	12	Blast resistance	(Kiyosawa and Ando, 1990; Yang et al., 2007; Geng et al., 2008; Huang et al., 2008; Wang et al., 2008a; Joshi et al., 2009; Sharma et al., 2009; Ramkumar et al., 2010)
14	*Pi-b*	*Pi-s*	2	Blast resistance	(Kiyosawa and Ando, 1990; Joshi et al., 2009)
15	*Pi-a*		11	Blast resistance	(Kiyosawa and Ando, 1990)
16	*Pi-i*		6	Blast resistance	(Kiyosawa and Ando, 1990)
17	*Pi-z*	*Pi-z and Pi-zt*	6	Blast resistance	(Kiyosawa and Ando, 1990)
18	*Pi-t*		1	Blast resistance	(Kiyosawa and Ando, 1990)
19	*Pi-sh*		1	Blast resistance	(Kiyosawa and Ando, 1990)
			11	Blast resistance	
20	*Pik^h*		-	Blast resistance	(Sharma et al., 2009; Ramkumar et al., 2010)
21	*NBS-LRR class R-genes*		11	Resistance of the plant	(Yang et al., 2008)

TILLING and Eco-TILLING

Many unique molecular biology techniques have been considered to determine mutations in genes to identify the function of these genes (reverse genetics). TILLING is a valuable and non-transgenic reverse genetic strategy to study gene function that allows screening for mutations in genes with known sequences in a plant mutant population or allows rapid mutational screening to obtain induced lesions in a gene of interest (McCallum et al., 2000).

TILLING helps in the direct identification of induced point mutations in a gene by heteroduplex evaluation (Till et al., 2003). TILLING is suitable for most plants and enables the identification of single-base-pair (bp) allelic variation in a target gene in a high-throughput manner. TILLING has several benefits over other techniques used to detect single-bp polymorphisms. Gel-based assays, such as single-strand conformation polymorphism (SSCP) and denaturing gradient gel electrophoresis (DGGE), do not pinpoint the location of the polymorphism or type of polymorphism contained in the DNA fragment (DeFrancesco and Perkel, 2001). Array hybridisation techniques are only effective in identifying approximately 50% of simple nucleotide polymorphisms (SNPs) (Borevitz et al., 2003).

TILLING can help us to focus on specific genes or genomic regions, avoiding difficulties with other functional genomic methods (such seeing that T-DNA knock-outs or perhaps RNAi-based gene silencing) that need the generation of transgenic plants (Sabetta et al., 2011). TILLING strategy, with developed software, will work for productive reverse genetics in the near future. (Henikoff et al., 2004). TILLING technology has much potential to examine natural as well as induced variation. Attributes and advantages of TILLING have been described in many publications and reviews (e.g. Gilchrist and Haughn, (2005); Stemple, (2004)). In this technique, artificially induced mutations are subjected to phenotypic and genotypic analysis. This technique employs mismatch-specific endonuclease to detect induced or natural DNA polymorphisms. The strategy is based on a PCR screening process within a recognised sequence region of interest for variants of heteroduplex analysis or mismatch cleavage to help detect mutations. Gene regions are targeted for mutation discovery using PCR and standard SNP discovery methods. The general protocol for the creation of a TILLING platform in plants consists of three main steps: 1) Development of a mutagenized (mutated) population, 2) DNA extraction from individual and Creation of DNA pools, and 3) mutation discovery (Till et al., 2007b) or screening the population for induced mutations using different

procedures, e.g. cleavage by specific endonuclease, denaturing high-performance liquid chromatography (DHPLC) or high-throughput sequencing.

The most typical technique used for TILLING is enzymatic mismatch cleavage followed by fluorescence detection using the LI-COR DNA analyser (Till et al., 2006b). In contrast to methods such as denaturing HPLC which can detect a mismatch but does not identify where it lies in the sequence, in the mismatch cleavage system the location of each mutation is determined within a few nucleotides. This method allows for confident identification of each mutation, whether heterozygous or homozygous, with a single sequencing run, priming with the nearer of the amplifying primers (Till et al., 2007b) and has been used for the discovery of thousands of induced mutations (Comai et al., 2004; Till et al., 2009).

For TILLING, DNA is extracted from a mutagenized population (each individual mutant). DNA from up to eight individuals (mutagenized lines) is pooled. After extraction and pooling, samples are typically arrayed into a 96-well format. The target region (~1.5 kb in a gene of interest) is amplified by PCR with gene-specific primers that are end-labeled with fluorescent dyes. The amplified products are denatured by heating and then allowed to cool slowly so that they randomly re-anneal. Heteroduplex molecules form, when a pool includes at least one plant that has a mutation in the amplified region, and become the substrate for enzymatic mismatch cleavage. Heteroduplexes are cleaved using a crude protein extract from celery containing the nuclease CEL I. The cleaved pieces (resultant products) are generally visualized on polyacrylamide denaturing gels electrophoresis and a gel readout platform such as the LI-COR DNA analyser, to identify individuals that may have a mutation or polymorphisms in the gene of interest. Use of a capillary detection system because of potential savings in time and gel costs makes the development of this possibility a desirable goal for the future (Gilchrist and Haughn, 2005). Mutations are identified by the presence of cleaved PCR products in a given pool. Once mutations are discovered, they are sequenced to determine the precise base change. The position with the mutation inside the amplicon can be estimated from the length of the fragments carrying the 30 and 50 fluorescent tags. Plants with mutations predicted to affect protein function can be carefully analysed for phenotypic abnormalities. Individuals or lines exhibiting the desired characteristics can be incorporated into a breeding program. Figure 2 shows a summary of the TILLING procedure. Application of TILLING has increased rapidly over the last 5 years for developing the number of crop species, indicating its attraction as a tool for breeding. TILLING has already been successfully applied in plants, including Arabidopsis (Till et al., 2003), lotus (Perry et al., 2003), maize (Till et al., 2004), wheat (Slade et al., 2004), indica rice (Wu et al., 2005; Till et al., 2007b), barley (Caldwell et al., 2004; Talamè et al., 2008), soybean

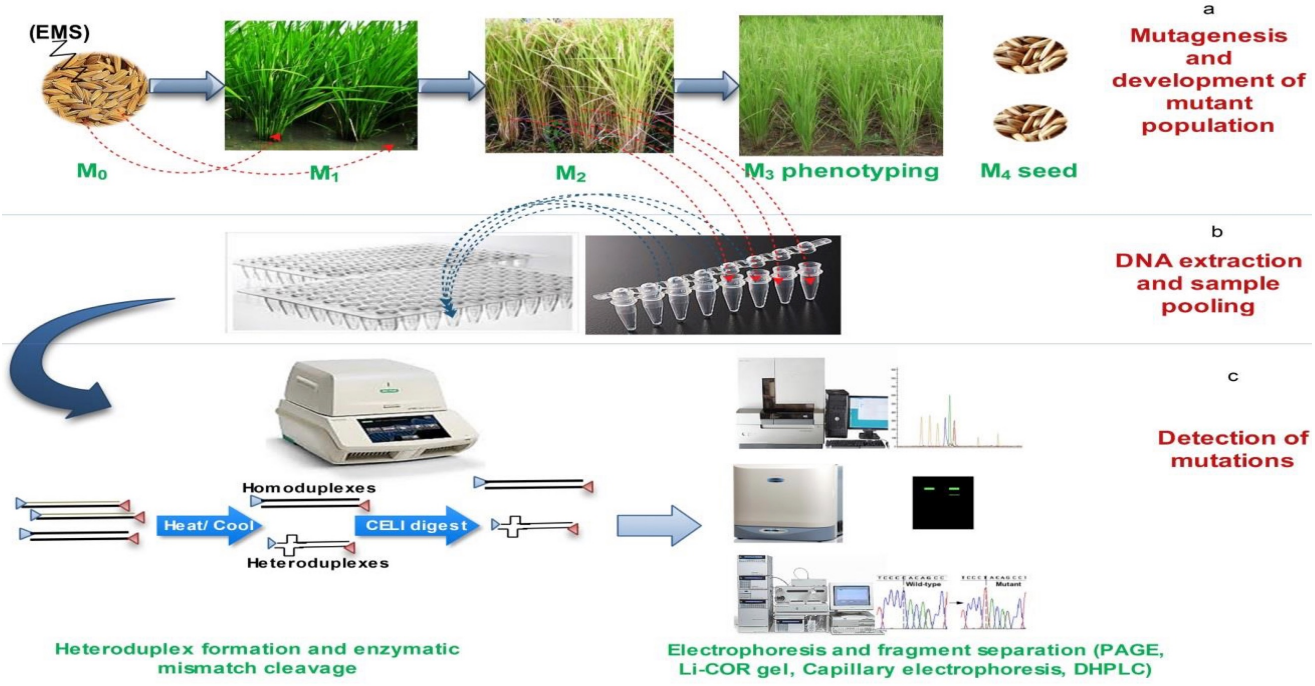


Figure 2. A summary of the TILLING Procedure; (a) EMS Mutagenesis to obtain mutagenized population (b) DNA extraction from each mutated plants (M2) and Pooling individual DNAs (c) Mutation discovery(PCR amplification with fluorescent labelled primers, heteroduplex formation, Enzymatic mismatch cleavage (Cutting the annealed products with a single-strand specific nuclease, typically with an extract containing CEL I enzyme), gel electrophoresis and mutation discovery(Purifying the products and run on gels or capillaries to detect the mismatched products, fragments are denatured and separated typically on a LI-COR DNA analyser).

(Cooper et al., 2008), *Pisum sativum* (Dalmais et al., 2008) and some animal species (Hurlstone et al., 2003; Wienholds et al., 2003; Smits et al., 2004; Stemple, 2004). The modified process for high-throughput TILLING utilising a LI-COR DNA analyser and fluorescently labelled PCR primers has been described for application in rice (Cooper et al., 2013). TILLING achievement with rice has resulted in the development of public TILLING support (http://tilling.ucdavis.edu/index.php/RiceTilling). TILLING promises to be generally applicable, the main limitation for TILLING is that the species is capable of being mutagenized (Barkley and Wang, 2008).

The first modifications of the TILLING technique called Eco-TILLING, was proposed by Comai et al. (2004). This technology is used to survey natural mutation throughout genes/ and or in germplasm. Although the process is generally the same (both techniques use mismatches produced by heteroduplexes of alleles of a gene), Eco-TILLING is ideal for the identification of natural variance within populations or even natural mutations within germplasm without using mutagenesis (Comai et al., 2004). DNA polymorphisms that can be discovered via Eco-TILLING include SNPs, small insertions and deletions (InDels); additionally, the delineation of haplotypes at loci of interest and variation in microsatellite (SSR) repeat number can be determined by Eco-TILLING (Comai et al., 2004; Till et al., 2006a). A modified Eco-TILLING method for the discovery of mutations in a multi-gene family, called Self-EcoTILLING, was reported by Wang et al., (2008b). In Eco-TILLING only one individual for each haplotype needs to be sequenced, therefore it is cost-effective for detection of variations (Kurowska et al., 2011). The technology is applicable to any organism, heterozygous or polyploidy and has practical application in the searching for resistance to new viruses or to create genetic diversity (Kurowska et al., 2011). The Eco-TILLING method has been used as an efficient SNP detection tool in wild and cultivated *Oryza* species (Rakshit et al., 2007). Natural alleles in the wild relatives of cultivars, which have limited genetic diversity, are also detectable through this technique (Kurowska et al., 2011).

Among the strengths of Eco-TILLING is its ability to significantly reduce the amount of sequencing required to determine the nucleotide sequence of a gene among many different individuals in the population (Gilchrist and Haughn, 2005).

So far, many important EcoTILLed genes have been screened in natural populations of different species (Kadaru et al., 2006; Wang et al., 2007; Wang et al., 2008c; Wang et al., 2010). Another valuable application of Eco-TILLING is searching for resistant alleles in natural populations (Mejlhede et al., 2006; Nieto et al., 2007; Elias et al., 2009; Ibiza et al., 2010; Till et al., 2010).

These approaches have been used to recognise alleles of important blast resistance genes from wild rice species such as *Pi-ta* from *O. ufipogon* (Griff) (Geng et al., 2008; Huang et al., 2008) and *O. rufipogon* (ETOR) (Yang et al., 2007), and *Pid3* from *O. rufipogon* (Joshi et al., 2009; Shang et al., 2009; Sharma et al., 2009; Rai et al., 2011a). Huang et al. (2008) and Shang et al. (2009) used PCR based approach to amplify and sequence alleles of 36 accessions of wild rice, *O. rufipogon*. These approaches have been employed in the discovery of almost 10,000 mutagen-caused mutations (Till et al., 2009). The steps involved in Eco-TILLING are illustrated in Figure 3.

Mutagenesis and Developing TILLING Populations

A mutagenized population is prepared by using a mutagen which typically causes single nucleotide polymorphisms, or insertions/deletions randomly throughout the genome. Almost all TILLING populations were developed using chemical mutagens, which have been successfully employed in mutation breeding programs. Many mutagenic treatments produce a chimeric plant in the first generation. Chimeras are dissolved and a structured population is typically developed. In TILLING project, chemical mutagens because of having developed protocol for many species are attractive. Chemical mutagens produce predictable heritable nucleotide changes at high densities. The chemical mutagenic treatment providing an easy and cost-effective way to saturate a genome is coupled with PCR-based mutation detection (Sabetta et al., 2011). Chemical mutagenic agents, such as Ethylmethane sulphonate (EMS), methylmethane sulphonate (MMS), hydrogen fluoride (HF), sodium azide, N-methyl-N-nitrosourea or methylnitrosourea (MNU), and hydroxylamine have been used to generate mutant populations for many years for forward genetics in various plant species (Parry et al., 2009). Among them, the alkylating agent (EMS) was most often applied and has been the mutagen of choice for most plant TILLING projects (Till et al., 2007b). EMS and MNU have both been successfully utilised in rice (Till et al., 2007b; Suzuki et al., 2008). Types of mutagens used and types of nucleotide substitutions detected in different TILLING populations in a variety of plant species has been previously addressed by Kurowska et al., (2011). For most of plants, seed mutagenesis is most practical. For some species, pollen mutagenesis can be considered (Till et al., 2004). The density of induced mutations is a major factor in the efficiency and cost of mutation discovery. Differences in mutation density between organisms may result from differences in the uptake of, cytotoxic response to, and repair of lesions induced by treatment with the mutagen. Following established protocols are a good route to success in mutagenesis. A mutated population becomes a TILLING platform when the DNA samples and seeds collected from a large M2 population are archived and put into databases. Usually, platforms of 3,000–5,000 M2 individuals are created, although larger populations that include 10,000 plants have also been reported. In many cases, phenotypic observations of the M2 generation are performed and large M3 generations are generated for forward screening (Talamè et al., 2008; Wang et al., 2008a; Himelblau et al., 2009; Minoia et al., 2010). TILLING databases are created and often made publicly available for developed platform.

Bioinformatics Tools in the TILLING Strategy

Bioinformatics facilitates both the analysis of genomic and post-genomic data. In addition, it enables integration of data from the related fields of transcriptomics, proteomics, metabolomics and phenomics. So far many bioinformatics tools have been developed for the TILLING strategy (Kurowska et al., 2011). In TILLING the sequence of the

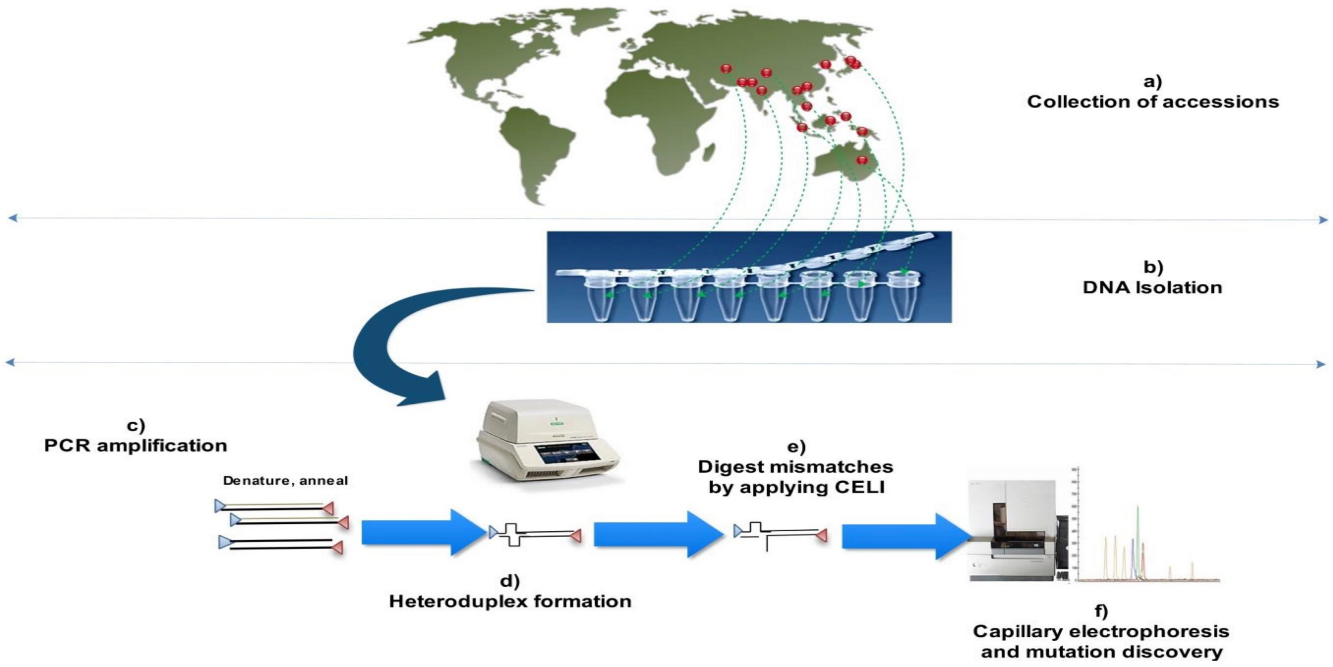


Figure 3. Outline of the Basic Steps for Typical Eco-TILLING to determine natural polymorphisms; (a) Collection of plant material (b) DNA Extraction from Natural population (c) PCR Amplification (d) Heteroduplexes Formation (e) Enzymatic Mismatch Cleavage (f) Mutation Discovery (Electrophoresis and fragment separation by capillary electrophoresis, identification of variation).

gene of interest can be obtained from databases, such as GenBank (http://www.ncbi.nlm.nih.gov/genbank/). The web based program such as CODDLE (Codons Optimized to Discover Deleterious LEsions; http://www.proweb.org/coddle/coddle_help.html) has been designed for use in TILLING analysis. CODDLE performs BLAST alignment to identify a conserved region with the use of SIFT (Sort Intolerant From Tolerant) and PSSM (Position-Specific Scoring Matrix). CODDLE enables users to input genomic, cDNA, or protein sequences and evaluates the probable effect of induced or natural polymorphisms on gene function (Gilchrist and Haughn, 2005). CODDLE analyses the alignment of genomic and amino acid sequences from closely related species with the use of BLASTN and ClustalW2 (http://www.ebi.ac.uk/Tools/ msa/clustalw2/), respectively. The web based program PARSESNP (Project Aligned Related Sequences and Evaluate SNPs; http://www.proweb.org/parsesnp/ can be used to display the locations of the polymorphisms in a gene/ genes in a graphical format (Taylor and Greene, 2003). PARSESNP determines the effect of single nucleotide polymorphisms (SNPs) on protein. It is also possible to predict the putative effect of mutation on the secondary structure of proteins, bioinformatics tools such as SAS (Sequence Annotated by Structure (Milburn et al., 1998) and SOPMA (Self-OPtimized Method for secondary structure prediction with Alignment) (Geourjon and Deleage, 1995) are used for this purpose.

Sequencing-Based Allele Mining

Sequencing-based allele mining involves PCR-based amplification of alleles of a gene in varied genotypes and then DNA sequencing to recognise nucleotide variance in the alleles (Figure 4). Various alleles among the cultivars through this approach can be identified. The method would help to analyse individuals for haplotype structure and diversity to infer genetic association studies in plants. Using this approach to recognise the effect of mutations on gene structure, the sequences are analysed for the location of point mutations or SNPs and insertions or deletions (InDels) to construct haplotypes. Sequencing-based allele mining is a very effective method to expand the rice blast R gene source and manage damaging blast disease. Additionally, this method is applicable to the management of additional crop diseases. Sequencing-based allele mining does not require much sophisticated equipment or involve tedious steps (Ramkumar et al., 2010), but the sequencing cost of the targeted genes for this technique is actually high, which will be the major cost for this type of allele mining program. However, with the advanced sequencing methods and techniques like massively parallel sequencer using supported oligonucleotide ligation and detection system (SOLiD) the sequencing cost is being reduced (Hutchison, 2007). TILLING and Eco-TILLING have been proposed to be economical approaches for haplotyping and SNP discovery. These types of techniques require more elegance and entail several steps, for example the creation of DNA pools, particular conditions for

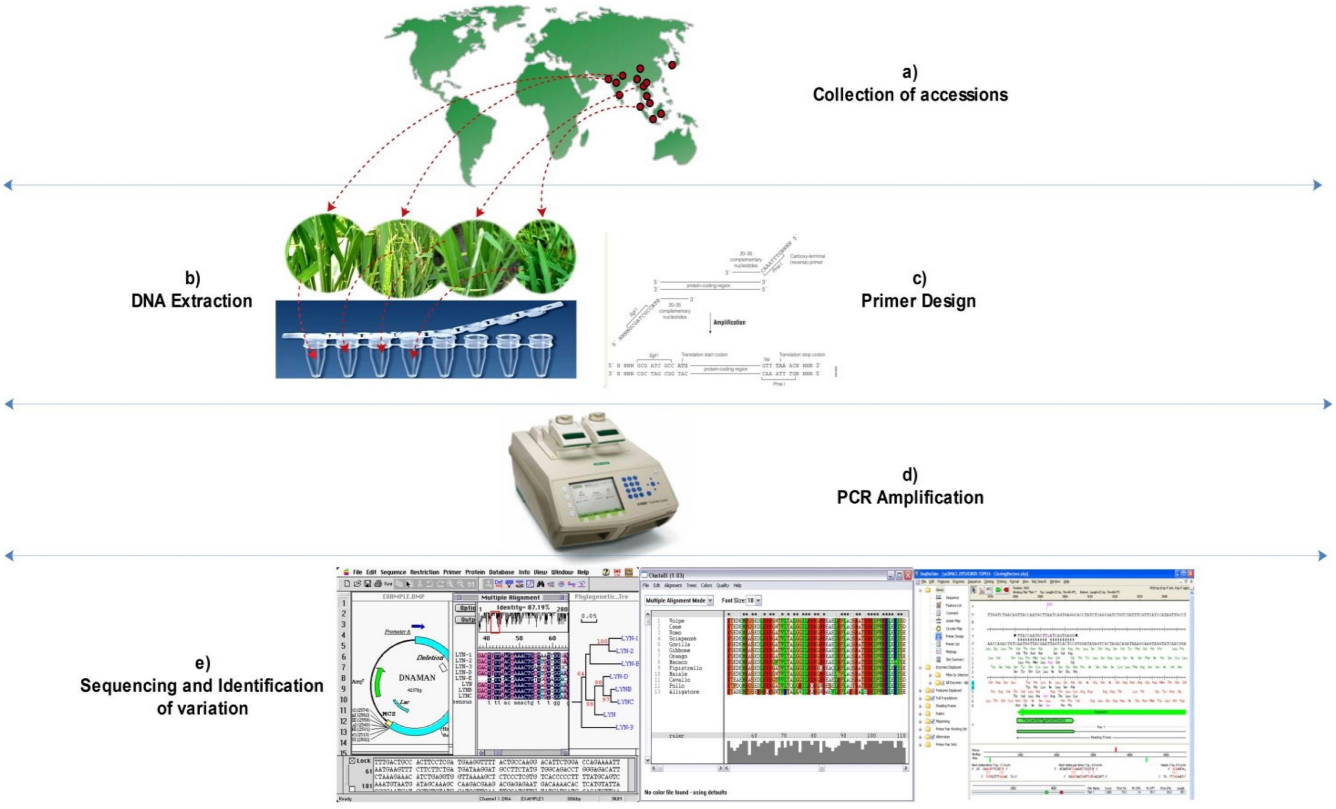


Figure 4. Sequencing-Based Allele Mining Procedure; (a) Collection (b) DNA extraction (c) Primer design (d) PCR amplification (e) Sequencing and identification of allelic variation.

nuclease cleavage, and polyacrylamide gel detection using the LI-COR genotyper. A short comparison between, Eco-TILLING and Sequencing-based allele mining, is provided in Table 2.

Next-generation Sequencing and Allele Mining

Sequencing is a method that allows knowledge of the arrangement using a precise sequence of nucleotides in a sample of DNA. High-throughput (HT) sequencing, also termed as next-generation sequencing (NGS), provides a new tool to add substantial value to gene bank collections (Kilian and Graner, 2012). NGS technologies can generate DNA sequence data inexpensively and faster compared with traditional technologies. Progress in sequencing technologies has increased the sequence capacity and has made whole-genome re-sequencing by individual laboratories feasible (Gupta, 2008; Hudson, 2008; Mardis, 2008). Using this technique, it is possible to re-sequence candidate genes for thousands of individuals, entire transcriptomes or entire plant genomes more efficiently, in greater depth and more economically than ever before (Varshney et al., 2009; Kilian and Graner, 2012). Identification of genetic variations has become so useful and exact that thousands of candidate genes can be tracked inside large gene bank collections (Varshney et al., 2009). Information on candidate genes and candidate regions can be harnessed for thousands of individuals to sample genetic diversity within and between germplasm pools, map Quantitative Trait Loci (QTLs), identify individual genes and determine their functional diversity (Kilian and Graner, 2012). Therefore, these technologies are becoming a standard choice for generating genetic data in fields such as population genetics, conservation genetics and molecular ecology. Large-scale NGS to screen germplasm collections with high accuracy is now possible using several different modern sequencing platforms, including Illumina/GA (Illumina; http://www.illumina.com), Roche/GS FLX (Roche Applied Science; http://454.com/products/index.asp), Applied Biosystems/ SOLiD (Applied Biosystems; https://products.appliedbiosystems.com) and cPAL sequencing (Shendure, 2008; Metzker, 2010) and and the HeliScope Single Molecule Sequencer (Helicos; http://www.helicosbio.com). For example, Illumina/Solexa genome analyser can sequence one billion bases (1 GB) of 30–40 base sequence reads in a single run in a short time period of 3–4 days. This approach will deliver a platform for the large-scale development of SNPs. Genotyping through sequencing is going to be utilized to obtain information on SNP and haplotype patterns. Also, accuracy and detection power of NGS-based TILLING has been compared favourably with other established methods (Tsai et al.,

Table 2. Comparison of allele mining techniques for identification of sequence polymorphisms for a given gene in the naturally occurring populations

Parameters	Eco-TILLING	Sequencing-based allele mining
DNA pooling	Yes	No
PCR	Yes	Yes
Heteroduplex formation	Yes	No
Primer labeling by fluorescent dyes	Yes	No
Primer cost	High	Low
Enzymatic mismatch cleavage	Yes	No
Electrophoresis	Fragments are denatured and separated typically on a LI-COR DNA analyser	Simple Agarose Gel
Cloning and Sequencing	Yes	Yes
Cost of Sequencing	Medium-to-high	High
Samples size for sequencing	Less	More
Time required	More(especially for sample preparation)	Less
Technical expertise	High (especially for DNA Pooling and detection of cleavage of heteroduplexes)	Low
Lab facilities	High Sophisticated	Low Sophisticated
Nucleotide detection	Effective in detection of SNPs rather than InDels	Effective in detection of any type of nucleotide
Throughput	Medium-to-high	High
Complexity	More	Less

2011), considering the NGS capacity and costs, to be a promising approach to TILLING and identification of rare alleles in populations in the future. Nevertheless, NGS will soon be a cost-effective measure for allele mining and useful for the rapid identification of abiotic stress tolerance genes (Hardin, 2008; Barh et al., 2013).

Applications of Allele Mining

Allele mining can be successfully employed for the detection of additional alleles from the gene of interest in diverse genetic resources. Allele mining can offer insight into the molecular basis of variations for any trait to recognise the nucleotide sequence changes related to superior alleles. In this method, allelic similarity/dissimilarity at a candidate gene and the allelic synteny with other members of the family can be studied additionally. Allele mining helps to distinguish molecular variation among related species. It assists in the monitoring the development of alleles and identification of new haplotypes at candidate genes of interest, as well as examining the haplotype variety. The information generated from the haplotype frequencies and their changes within the populations might form the foundation for organisation of association mapping research. Haplotypes of tightly related single-nucleotide polymorphisms (SNPs) probably provide increased power compared with individual SNPs to detect the association between genetic variants and disease. Using sequence information from allele mining studies, syntenic relationships can be analysed for the identified loci/genes over the species. Allele mining exploits the particular DNA sequence of a single genotype to isolate useful alleles from related genotypes (Latha et al., 2004). The accessibility of the sequences of the candidate genes that underlie agronomic traits may help the practical gene-based use of genetic resources (Graner, 2006). The evaluation of genes/QTL conferring the level of resistance to rice blast pathogen has showed high similarity between rice and barley regarding sequence and position, showing one common evolutionary origin for this resistance gene (Chen et al., 2003). Allele mining of the blast resistance gene *Pi-ta* showed 91 nucleotide polymorphisms and 18 InDel variations among wild rice accessions of *O. rufipogon* and *O. barthii* (Huang et al., 2008). Moreover, Wang et al. (2008a) reported 16 different sequences having various insertion and deletions in the exon and intron regions of the *Pita* gene in cultivated and wild rice relatives. Using the recognition of sequence variation, allele-specific molecular markers can be developed via allele mining techniques. Most of these markers are applied for accurate introgression of novel alleles in genomics-assisted breeding programs such as MAS. Common applications involving allele mining are shown in Figure 5.

Bioinformatics Tools Employed for Allele Mining

Using genomic technologies with genetic bioinformatics tools can be successful in identifying allelic variations to dissect trait-gene associations. Presently, a suite of strong bioinformatics tools and computer software programs are freely available for the identification of sequence polymorphisms, comparison with other DNA sequence, prediction of amino acid changes, and identification of transcription factor binding sites (TFBSs), transcriptional regulatory elements and a database of plant promoter sequences. Allele mining entails various bioinformatics resources, including PLACE, plantCARE, TRANSFAC, JASPAR, MEME, Plantprom DB, DCPD, SCPD, ClustalW and BioEdit. Clustal-W and Bioedit are widely used to compare and align multiple DNA sequences. A list of bioinformatics tools utilised to accomplish allele mining evaluation is provided in Table 3.

Figure 5. Applications of Allele Mining.

Conclusions

The rice genome has been completely sequenced and is available to the public (www.ncbi.nlm.nih.gov). These specific routine data have many applications and aid in the discovery of genes and alleles in rice and other related plant species. Identifying useful genes and alleles and their introduction to susceptible genotypes has potential applications in crop improvement programs. There are many methods to identify diversity that may not exist in the germplasm pool of existing breeding lines, including allele mining, transformation, mutation breeding, use of landraces or synthetic polyploids and wide crossing (Able et al., 2007). Allele mining, which can be essential to determine and utilise novel alleles that are hidden in genetic diversity, is discussed here. This review was composed in the perspective of any biologist. We focused primarily on the general concept, and principle of allele mining; thereafter, its methodology and usage in the discovery of novel alleles in rice germplasm and benefits of this strategy in crop sciences were also described. The use of genome sequencing information available in open data sources for the identification and isolation of novel and superior alleles of agronomically important genes is vital for the development of improved cultivars. Allele mining is significant for presenting the rice breeder with direct access to crucial alleles conferring resistance to biotic and abiotic stresses. Allele mining helps in the tracing of the evolution of alleles, identification of new haplotypes with the development of allele-specific markers for use in marker-assisted selection. Therefore, to date, many global crop research institutes are adopting with allele mining strategies. Identification of allelic variants among germplasm collections provides new germplasm for the novel alleles to target trait improvement but additionally categorises the germplasm entries because of their conservation (Latha et al., 2004). The international project to sequence the genome of *O. sativa* L cv. Nipponbare will likely make allele mining possible for all genes of rice, even similar cereals (Latha et al., 2004). Natural mutations are an excellent contributor for the evolution of new alleles. Mutations such as transitions, transversions and Insertions/ Deletions (InDels) are the basis for this advancement. In this review, TILLING and Eco-TILLING techniques, beneficial applications and some limitations for these techniques in plant were discussed. “TILLING is currently being used for the detection of both induced and natural DNA polymorphisms in several plant species. Its advantages over other reverse genetic techniques include its applicability to virtually any organism, its facility for high-throughput and its independence of genome size, reproductive system or generation time” (Gilchrist and Haughn, 2005). TILLING and Eco-TILLING are closely related methods. The Eco-TILLING technology is applicable to any organism including those that are heterozygous and polyploidy, to survey natural variation in genes. Next-generation sequencing as new strategies for sequencing were discussed in this review. NGS has been applied to a limited number of sequenced individuals using Sanger’s method. Techniques can be used for direct mutation determination without any pre- screening (Kurowska et al., 2011). An NGS platform has dramatically increased the speed of DNA sequencing and has also reduced the cost of sequencing by more than two orders of magnitude (Deschamps and Campbell, 2010). Hence, NGS approaches are becoming available to screen major gene bank collections.

Recent progress on the application of mining strategies for blast resistance genes was also discussed in this review. With application of allele mining more functional R genes/or alleles could be identified from rice varieties. This allelic diversity is the essential base for rapid evolution of R genes and could be used as new resources in the improvement of rice varieties. The locus-based blast R gene map should be very informative for the identification of new alleles in germplasm by referring to the sequence diversity in known haplotypes. Several services and bioinformatics tools currently are available for allele mining, some open-source software useful for allele and promoter mining and were addressed in this review. While new technologies are available it is important to note that the current state is great in terms of saving time or cost and will facilitate streamlining the process of functional genomics for all researchers. Improved software and logistics, will serve for efficient reverse genetics over the next few years.

Prospects and Future Consideration for Allele Mining

Allele mining possesses good potential to be applied in molecular plant breeding. With the ever-increasing sequence data in GenBank and ever-expanding crop gene banks, it is highly essential to develop novel and efficient mining strategies to screen GenBank collections more efficiently for DNA sequence variation and the management of genome resources. The success of allele mining mainly depends on the type of genetic materials used for screening and should be as diverse as possible. To this end, wild relatives and local landraces are used because they are reservoirs of useful alleles hidden in their phenotype (Tanksley et al., 1996). Hence, we need to construct more polished sampling methods for the efficient selection of genotypes to distinguish valuable and novel alleles underlying particular traits. Several factors, including efficient and reliable phenotyping techniques, genomic

Table 3. List of some Open-Source Software and Bioinformatics Tools Useful for Allele and Promoter Mining

S.No.	Description or Utility	Database Name	Uniform Resource Locator (URL)	Reference
1	Multiple EM for Motif Elicitation. Analysing DNA and protein sequence motifs for similarities among them. Motif discovery tool	MEME	meme.nbcr.net/meme/cgi-bin/meme.cgi	(Bailey et al., 2006)
2	Database of Transcription Factor Binding Site (TFBS)	JASPAR	http://jaspar.genereg.net/	(Bryne et al., 2008)
3	Promoter analysis based on transcription factor binding sites	MATinspector and beyond	http://www.genomatix.de/products/MatInspector/	(Cartharius et al., 2005)
4	Motif finding program. Effective tool for discovering TF binding sites from gene expression or ChIP-chip data	W-AlignACE	http://www1.spms.ntu.edu.sg/~chenxin/W-AlignACE/	(Chen et al., 2008)
5	Arabidopsis gene regulatory information server. A new information resource of Arabidopsis cis- promoter sequences, transcription factors (TF) and their target genes. Transcription factors and Regulatory Elements (RE) databases	AGRIS	http://arabidopsis.med.ohio-state.edu	(Davuluri et al., 2003)
6	Object-Oriented Transcription Factor Database: an object-oriented successor to TFD. The original Transcription Factors Database	OOTFD	http://www.ifti.org/ootfd/	(Ghosh, 1998)
7	Database of motifs found in plant cis-acting regulatory DNA elements. The PLACE database also contains a brief description of each motif and relevant literature with PubMed ID numbers.	PLACE	http://www.dna.affrc.go.jp/PLACE/index.html	(Higo et al., 1999)
8	Alternative tools for linking gene expression modules with promoter sequences motifs to uncover gene regulation mechanisms in plants. ModuleFinder provided a subset of co-expressed gene modules that are more logically related to biological functions than did subsets derived from traditional hierarchical clustering techniques	ModuleFinder and CoReg	http://www.plantmethods.com/content/2/1/8/	(Holt et al., 2006)
9	A transcriptional regulatory element database. Collection of mammalian regulatory elements	TRED	http://rulai.cshl.edu/cgi-in/TRED/tred.cgi?process=home	
10	FastPCR software for PCR primers or probes design and in silicoPCR, oligonucleotide assembly and analyses, alignment and repeat searching for a single sequence or for comparisons of two sequences	FastPCR	http://en.bio-soft.net/pcr/FastPCR.html http://www.fastpcr.software.informer.com/ www.biocenter.helsinki.fi/bi/Programs/fastpcr.htm	(Kalendar et al., 2009)
11	Transcription Regulatory Region Database. Description of Regulatory Elements and TFBS. An entry of the database corresponds to a gene and contains the data on localization and functions of the transcription regulatory regions as well as gene expression patterns.	TRRD	http://wwwmgs.bionet.nsc.ru/mgs/gnw/trrd/	(Kolchanov et al., 2000)
12	Drosophila Core Promoter Database. The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters	DCPD	http://www.biology.ucsd.edu/labs/Kadonaga/DCPD.html	(Kutach and Kadonaga, 2000)
13	A database of plant cis-acting regulatory elements	PlantCARE	http://bioinformatics.psb.ugent.be/webtools/plantcare/html The PlantCARE relational database is available at: http://sphinx.rug.ac.be:8080/PlantCARE/	(Lescot et al., 2002)
14	Data base of transcription factors. Eukaryotic Transcriptional regulation, comprising data on transcription factors, their target genes and regulatory binding sites	TRANSFAC	http://www.gene-regulation.com/pub/programs.html The TRANSFAC® public release is available at: http://www.gene-regulation.com. For yeast an additional release including the latest data was made available separately as TRANSFAC® *Saccharomyces* Module (TSM) at: http://transfac.gbf.de. For CYTOMER®free download versions are available at http://www.biobase.de:8080/index.html.	(Matys et al., 2003)
15	Primer design. A computer program that suggests PCR primers for a variety of applications, for example to create STSs (sequence tagged sites) for radiation hybrid mapping, or to amplify sequences for single nucleotide polymorphism discovery. Primer3 can also select single primers for sequencing reactions and can design oligonucleotide hybridization probes	Primer 3	http://frodo.wi.mit.edu/primer3/	(Rozen and Skaletsky, 1999)
16	The Eukaryotic Promoter Database. An annotated non-redundant collection of eukaryotic POL II promoters, experimentally defined by a transcription start site (TSS)	EPD	http://www.epd.isb-sib.ch/	(Schmid et al., 2004)
17	A database of plant promoter sequences. An annotated, non-redundant collection of proximal promoter sequences for RNA polymerase II with experimentally determined transcription start site(s), TSS, from various plant species	Plantprom DB	http://mendel.cs.rhul.ac.uk/mendel.php?topic=plantprom http://linux1.softberry.com/berry.phtml?topic=plantprom&group=data&subgroup=plantprom http://mendel.cs.rhul.ac.uk/ http://www.softberry.com/	(Shahmuradov et al., 2003)
18	A promoter database of the yeast *Saccharomyces cerevisiae*. Facilitate a systematic study of the promoters and transcriptionally regulatory cis-elements of the yeast Saccharomyces cerevisiae on a genomic scale. SCPD also provides some simple but useful tools for promoter sequence analysis	SCPD (*Saccharomyces cerevisiae* promoter database)	http://rulai.cshl.edu/SCPD/ http://cgsigma.cshl.org/jian	(Zhu and Zhang, 1999)
19	A biological sequence alignment editor. Nucleotide sequence analysis and manipulation	BioEdit	www.mbio.ncsu.edu/BioEdit/BioEdit.html http://jwbrown.mbio.ncsu.edu/Bioedit/BioDoc.pdf	(Hall, 1999)
20	Multiple alignment of DNA and protein sequences/Multiple sequence alignment programs. The current Clustal programs all derive from Clustal W, These programs have been amended and added to many times since 1994 in order to increase functionality and to increase sensitivity.	Clustal W and Clustal X, T-Coffee, MAFFT and MUSCLE	www.ebi.ac.uk www.clustal.org/clustal2/ http://bioinformatics.oxfordjournals.org/content/23/21/2947.full.pdf&embedded=true	(Thompson et al., 1994), (Thompson et al., 1997), (Larkin et al., 2007)
21	DNA Sequence Polymorphism analysis, to identify SNPs and InDels and Phylogeny analysis based on the dis-similarity of DNA sequences	DnaSP	http://www.ub.edu/dnasp http://bioinformatics.oxfordjournals.org/content/25/11/1451	(Librado and Rozas, 2009)
22	Integrated Software Molecular Evolutionary Genetics Analysis and Sequence Alignment. An integrated tool for conducting sequence alignment, inferring phylogenetic trees, estimating divergence times, mining online databases, estimating rates of molecular evolution, inferring ancestral sequences, and testing evolutionary hypotheses	MEGA	http://www.megasoftware.net/ www.ncbi.nlm.nih.gov/pubmed/24132122	(Tamura et al., 2007), (Tamura et al., 2013)
23	A complex software tool that allows users to extract a common framework of elements from a set of DNA sequences. Motif discovery tool	FrameWorker	http://www.genomatix.de/online_help/help_gems/FrameWorker.html	(Cartharius et al., 2005)
24	DiAlign TF displays transcription factor (TF) binding site matches within a multiple alignment. Data base of promoter analysis	DiAlign TF: Multiple alignment plus TF sites	http://www.genomatix.de/online_help/help_dialign/dialign_TF.html	(Cartharius et al., 2005)

resources, and high-throughput techniques for quick generation of allelic data points, cost-effective sequencing platforms and efficient bioinformatics tools to identify nucleotide variation and molecular marker construction technologies for marker-assisted selection (MAS), are important in allele mining. Considering these factors, allele mining can be initiated for any crop with sequence information and for a related family member. Successful and efficient allele mining activity would depend mainly on the existing genetic base and availability of genome and gene sequence information of a particular crop species. Because, efficient, cost-effective and high-throughput sequencing platforms are available, routine allele mining via sequencing would certainly be a wide-spread tool to score variations at candidate genes between individuals and can be a highly effective and economically feasible method for the development of novel genetic resources for deployment in plant breeding. Sequencing-based allele mining would emerge as a method of choice to reveal natural variations and provide novel and effective alleles and would take center stage for all crop improvement activities (Ramkumar et al., 2010). Using these methods, identification of functional nucleotide polymorphisms in superior alleles in different crop species and from related species with a desired number of functions and comparison of sequence variation in a range of alleles for a given gene could be more possible. TILLING and Eco-TILLING are high-throughput and low-cost methods for the discovery of induced mutations and natural polymorphisms. TILLING methodology makes it possible to focus on specific genes or genomic regions. Progresses in genomic technology, such as developed SNP detection or DNA sequencing, will

make TILLING easier and simpler. It is expected that high-throughput TILLING method will become progressively popular, especially for agriculture, where there is so much useful knowledge to be gained from functional genomics and where no- transgenic methods are especially desirable (Chen et al., 2014). Eco-TILLING has also grown in popularity. Both methods can provide higher efficiencies in breeding program and will therefore be useful in meeting expected demands. Additionally, while several bioinformatics tools are currently available, advanced allele mining strategies and additional efforts are needed to develop computational tools for efficient allele identification. NGS (next-generation sequencing)-based approaches e.g., GBS (genotyping by sequencing), MPS (massively parallel sequencing), or low coverage WGS (whole genome sequencing) have manifested the power for quick and cost-effective identification of massive sequence polymorphisms in diverse germplasm. NGS technologies are expected not only to revolutionise genomics research but also to fundamentally change the management of genetic resource collections and services provided by gene banks to the user community (Kilian and Graner, 2012; McCouch et al., 2012). The challenge to build the association between sequence polymorphisms and putative function is of importance for the application of allele mining for R (resistant) genes. GWAS (genome-wide association study) analysis and high-throughput functional validation, like RNAi approaches, could facilitate the identification of new R alleles. Overall, to improve food security, to promote fundamental research and practical breeding new technologies must be developed and advanced methods for identification of novel alleles for functional genomics and crop improvement may become an effective tool for helping in feeding of the twenty−first century world.

Acknowledgements

We thank our university colleagues for their helpful suggestions, insightful comments and discussions on this valuable review. This work was supported by the Long-term Research Grants Scheme (LRGS), Rice Food Security Project from the Ministry of Education, Malaysia.

References

Able, J.A., Langridge, P. and Milligan, A.S., 2007. Capturing diversity in the cereals: many options but little promiscuity. Trends in plant science 12, 71-79.

Ashikawa, I., Hayashi, N., Yamane, H., Kanamori, H., Wu, J., Matsumoto, T., Ono, K. and Yano, M., 2008. Two adjacent nucleotide-binding site-leucine-rich repeat class genes are required to confer *Pikm*-specific rice blast resistance. Genetics 180, 2267-2276.

Ashkani, S., Rafii, M., Shabanimofrad, M., Ghasemzadeh, A., Ravanfar, S. and Latif, M., 2014. Molecular progress on the mapping and cloning of functional genes for blast disease in rice (*Oryza sativa* L.): current status and future considerations. Critical reviews in biotechnology, 1-15.

Bailey, T.L., Williams, N., Misleh, C. and Li, W.W., 2006. MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Research 34, W369-W373.

Barh, D., Zambare, V. and Azevedo, V., 2013. Omics: applications in biomedical, agricultural, and environmental sciences, CRC Press.

Barkley, N. and Wang, M., 2008. Application of TILLING and EcoTILLING as reverse genetic approaches to elucidate the function of genes in plants and animals. Current genomics 9, 212.

Bonman, J., 1992. Durable resistance to rice blast disease-environmental influences, Breeding for Disease Resistance. Springer, pp. 115-123.

Borevitz, J.O., Liang, D., Plouffe, D., Chang, H.-S., Zhu, T., Weigel, D., Berry, C.C., Winzeler, E. and Chory, J., 2003. Large-scale identification of single-feature polymorphisms in complex genomes. Genome Research 13, 513-523.

Bryan, G.T., Wu, K.-S., Farrall, L., Jia, Y., Hershey, H.P., McAdams, S.A., Faulk, K.N., Donaldson, G.K., Tarchini, R. and Valent, B., 2000. A single amino acid difference distinguishes resistant and susceptible alleles of the rice blast resistance gene *Pi-ta*. The Plant Cell Online 12, 2033-2045.

Bryne, J.C., Valen, E., Tang, M.-H.E., Marstrand, T., Winther, O., da Piedade, I., Krogh, A., Lenhard, B. and Sandelin, A., 2008. JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. Nucleic Acids Research 36, D102-D106.

Caldwell, D.G., McCallum, N., Shaw, P., Muehlbauer, G.J., Marshall, D.F. and Waugh, R., 2004. A structured mutant population for forward and reverse genetics in Barley (*Hordeum vulgare* L.). The Plant Journal 40, 143-150.

Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M. and Werner, T., 2005. MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics 21, 2933-2942.

Chen, H., Wang, S., Xing, Y., Xu, C., Hayes, P.M. and Zhang, Q., 2003. Comparative analyses of genomic locations and race specificities of loci for quantitative resistance to Pyricularia grisea in rice and barley. Proceedings of the National Academy of Sciences 100, 2544-2549.

Chen, J., Shi, Y., Liu, W., Chai, R., Fu, Y., Zhuang, J. and Wu, J., 2011. A *Pid3* allele from rice cultivar Gumei2 confers resistance to *Magnaporthe oryzae*. Journal of Genetics and Genomics 38, 209-216.

Chen, L., Hao, L., Parry, M.A., Phillips, A.L. and Hu, Y., 2014. Progress in TILLING as a tool for functional genomics and improvement of crops. Journal of integrative plant biology 56, 425-443.

Chen, X., Guo, L., Fan, Z. and Jiang, T., 2008. W-AlignACE: an improved Gibbs sampling algorithm based on more accurate position weight matrices learned from sequence and gene expression/ChIP-chip data. Bioinformatics 24, 1121-1128.

Chen, X., Shang, J., Chen, D., Lei, C., Zou, Y., Zhai, W., Liu, G., Xu, J., Ling, Z. and Cao, G., 2006. AB-lectin receptor kinase gene conferring rice blast resistance. The Plant Journal 46, 794-804.

Comai, L., Young, K., Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C. and Odden, A.R., 2004. Efficient discovery of DNA

polymorphisms in natural populations by Ecotilling. The Plant Journal 37, 778-786.

Cooper, J.L., Henikoff, S., Comai, L. and Till, B.J., 2013. TILLING and ecotilling for rice, Rice Protocols. Springer, pp. 39-56.

Cooper, J.L., Till, B.J., Laport, R.G., Darlow, M.C., Kleffner, J.M., Jamai, A., El-Mellouki, T., Liu, S., Ritchie, R. and Nielsen, N., 2008. TILLING to detect induced mutations in soybean. BMC plant biology 8, 9.

Correa-Victoria, F. and Zeigler, R.S., 1993. Pathogenic variability in *Pyricularia grisea* at a rice blast hot spot breeding site in eastern Colombia. Plant Disease 77, 1029-1035.

Dai, Y., Jia, Y., Correll, J., Wang, X. and Wang, Y., 2010. Diversification and evolution of the avirulence gene "*AVR-Pita1*" in field isolates of "*Magnaporthe oryzae*". Fungal Genetics and Biology 47, 973-980.

Dalmais, M., Schmidt, J., Le Signor, C., Moussy, F., Burstin, J., Savois, V., Aubert, G., Brunaud, V., de Oliveira, Y. and Guichard, C., 2008. UTILLdb, a *Pisum sativum* in silico forward and reverse genetics tool. Genome Biol 9, R43.

Das, A., Soubam, D., Singh, P., Thakur, S., Singh, N. and Sharma, T., 2012. A novel blast resistance gene, *Pi54*rh cloned from wild species of rice, *Oryza rhizomatis* confers broad spectrum resistance to *Magnaporthe oryzae*. Functional & integrative genomics 12, 215-228.

Datta, K., Baisakh, N., Thet, K.M., Tu, J. and Datta, S., 2002. Pyramiding transgenes for multiple resistance in rice against bacterial blight, yellow stem borer and sheath blight. Theoretical and applied genetics 106, 1-8.

Davuluri, R.V., Sun, H., Palaniswamy, S.K., Matthews, N., Molina, C., Kurtz, M. and Grotewold, E., 2003. AGRIS: Arabidopsis gene regulatory information server, an information resource of Arabidopsis cis-regulatory elements and transcription factors. BMC bioinformatics 4, 25.

DeFrancesco, L. and Perkel, J.M., 2001. Lab Consumer-In Search of Genomic Variation. Scientist-the Newspaper for the Science Professional 15, 24-27.

Deschamps, S. and Campbell, M.A., 2010. Utilization of next-generation sequencing platforms in plant genomics and genetic variant discovery. Molecular breeding 25, 553-570.

Elias, R., Till, B.J., Mba, C. and Al-Safadi, B., 2009. Optimizing TILLING and Ecotilling techniques for potato (*Solanum tuberosum* L). BMC research notes 2, 141.

Fukuoka, S., Saka, N., Koga, H., Ono, K., Shimizu, T., Ebana, K., Hayashi, N., Takahashi, A., Hirochika, H. and Okuno, K., 2009. Loss of function of a proline-containing protein confers durable disease resistance in rice. Science 325, 998-1001.

Geng, X., Yang, M., Huang, X., Cheng, Z., Fu, J., Sun, T. and Li, J., 2008. Cloning and analyzing of rice blast resistance gene *Pi-ta*+ allele from Jinghong erect type of common wild rice (*Oryza rufipogon* Griff) in Yunnan. Yi chuan= Hereditas/Zhongguo yi chuan xue hui bian ji 30, 109-114.

Geourjon, C. and Deleage, G., 1995. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Computer applications in the biosciences: CABIOS 11, 681-684.

Ghosh, D., 1998. OOTFD (Object-Oriented Transcription Factors Database): an object-oriented successor to TFD. Nucleic Acids Research 26, 360-361.

Gilchrist, E.J. and Haughn, G.W., 2005. TILLING without a plough: a new method with applications for reverse genetics. Current opinion in plant biology 8, 211-215.

Graner, A., 2006. Barley research at IPK. Molecular markers for allele mining, IPGRI, 25.

Gupta, P., 2008. Ultrafast and low-cost DNA sequencing methods for applied genomics research. Proceedings of the Indian National Science Academy-Part B: Biological Sciences 78, 91-102.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucleic acids symposium series. pp. 95-98.

Hardin, S.H., 2008. Real-time DNA sequencing. Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics.

Hayashi, K. and Yoshida, H., 2009. Refunctionalization of the ancient rice blast disease resistance gene *Pit* by the recruitment of a retrotransposon as a promoter. The Plant Journal 57, 413-425.

Hayashi, N., Inoue, H., Kato, T., Funao, T., Shirota, M., Shimizu, T., Kanamori, H., Yamane, H., Hayanoâ€Saito, Y. and Matsumoto, T., 2010. Durable panicle blast-resistance gene *Pb1* encodes an atypical CC-NBS-LRR protein and was generated by acquiring a promoter through local genome duplication. The Plant Journal 64, 498-510.

Helliwell, E.E., Wang, Q. and Yang, Y., 2013. Transgenic rice with inducible ethylene production exhibits broad-spectrum disease resistance to the fungal pathogens *Magnaporthe oryzae* and *Rhizoctonia solani*. Plant biotechnology journal 11, 33-42.

Henikoff, S., Till, B.J. and Comai, L., 2004. TILLING. Traditional mutagenesis meets functional genomics. Plant Physiology 135, 630-636.

Higo, K., Ugawa, Y., Iwamoto, M. and Korenaga, T., 1999. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Research 27, 297-300.

Himelblau, E., Gilchrist, E.J., Buono, K., Bizzell, C., Mentzer, L., Vogelzang, R., Osborn, T., Amasino, R.M., Parkin, I.A. and Haughn, G.W., 2009. Forward and reverse genetics of rapid-cycling *Brassica oleracea*. Theoretical and applied genetics 118, 953-961.

Hittalmani, S., Kahani, F., Dhanagond, S.M. and Rao, A.M., 2013. DNA marker characterization for allele mining of blast and bacterial leaf blight resistant genes and evaluation for grain yield. African Journal of Biotechnology 12, 2331-2340.

Holt, K.E., Millar, A.H. and Whelan, J., 2006. ModuleFinder and CoReg: alternative tools for linking gene expression modules with promoter sequences motifs to uncover gene regulation mechanisms in plants. Plant Methods 2, 8.

Hua, L., Wu, J., Chen, C., Wu, W., He, X., Lin, F., Wang, L., Ashikawa, I., Matsumoto, T. and Wang, L., 2012. The isolation of *Pi1*, an allele at the *Pik* locus which confers broad spectrum resistance to rice blast. Theoretical and applied genetics 125, 1047-1055.

Huang, C.-L., Hwang, S.-Y., Chiang, Y.-C. and Lin, T.-P.,

2008. Molecular evolution of the *Pi-ta* gene resistant to rice blast in wild rice (*Oryza rufipogon*). Genetics 179, 1527-1538.

Huang, X., Feng, Q., Qian, Q., Zhao, Q., Wang, L., Wang, A., Guan, J., Fan, D., Weng, Q. and Huang, T., 2009. High-throughput genotyping by whole-genome resequencing. Genome Research 19, 1068-1076.

Hudson, M.E., 2008. Sequencing breakthroughs for genomic ecology and evolutionary biology. Molecular ecology resources 8, 3-17.

Hulbert, S.H., Webb, C.A., Smith, S.M. and Sun, Q., 2001. Resistance gene complexes: evolution and utilization. Annual review of phytopathology 39, 285-312.

Hurlstone, A.F., Haramis, A.-P.G., Wienholds, E., Begthel, H., Korving, J., van Eeden, F., Cuppen, E., Zivkovic, D., Plasterk, R.H. and Clevers, H., 2003. The Wnt/β-catenin pathway regulates cardiac valve formation. Nature 425, 633-637.

Hutchison, C.A., 2007. DNA sequencing: bench to bedside and beyond. Nucleic Acids Research 35, 6227-6237.

Ibiza, V.P., Cañizares, J. and Nuez, F., 2010. EcoTILLING in Capsicum species: searching for new virus resistances. BMC genomics 11, 631.

Jinhua, X., Grandillo, S., Ahn, S.N., McCouch, S.R., Tanksley, S.D., JiMing, L. and LongPing, Y., 1996. Genes from wild rice improve yield. Nature (London) 384, 223-224.

Joshi, B.K., Bimb, H.P., Parajuli, G. and Chaudhary, B., 2009. Molecular tagging, allele mining and marker aided breeding for blast resistance in rice. BSN e-Bulletin 1, 1-23.

Kadaru, S.B., Yadav, A.S., Fjellstrom, R.G. and Oard, J.H., 2006. Alternative Ecotilling protocol for rapid, cost-effective single-nucleotide polymorphism discovery and genotyping in rice (*Oryza sativa* L.). Plant Molecular Biology Reporter 24, 3-22.

Kalendar, R., Lee, D. and Schulman, A.H., 2009. FastPCR software for PCR primer and probe design and repeat search. Focus on Bioinformatics/Mansour, A.(Ed).

Kilian, B. and Graner, A., 2012. NGS technologies for analyzing germplasm diversity in genebanks. Briefings in functional genomics, elr046.

Kiyosawa, S. and Ando, I., 1990. Blast resistance. Science of rice plant 3, 361-385.

Koizumi, S., 2007. Durability of resistance to rice blast disease. A differential system for blast resistance for stable rice production environment. JIRCAS working report, 1-10.

Kolchanov, N.A., Podkolodnaya, O.A., Ananko, E.A., Ignatieva, E.V., Stepanenko, I.L., Kel-Margoulis, O.V., Kel, A.E., Merkulova, T.I., Goryachkovskaya, T. and Busygina, T., 2000. Transcription regulatory regions database (TRRD): its status in 2000. Nucleic Acids Research 28, 298-301.

Kou, Y. and Wang, S., 2010. Broad-spectrum and durability: understanding of quantitative disease resistance. Current opinion in plant biology 13, 181-185.

Kumari, A., Das, A., Devanna, B., Thakur, S., Singh, P., Singh, N. and Sharma, T., 2013. Mining of rice blast resistance gene *Pi54* shows effect of single nucleotide polymorphisms on phenotypic expression of the alleles. European Journal of Plant Pathology 137, 55-65.

Kurowska, M., Daszkowska-Golec, A., Gruszka, D., Marzec, M., Szurman, M., Szarejko, I. and Maluszynski, M., 2011. TILLING-a shortcut in functional genomics. Journal of applied genetics 52, 371-390.

Kutach, A.K. and Kadonaga, J.T., 2000. The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters. Molecular and cellular biology 20, 4754-4764.

Larkin, M.A., Blackshields, G., Brown, N., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A. and Lopez, R., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947-2948.

Latha, R., Rubia, L., Bennett, J. and Swaminathan, M., 2004. Allele mining for stress tolerance genes in Oryza species and related germplasm. Molecular biotechnology 27, 101-108.

Lee, S.-K., Song, M.-Y., Seo, Y.-S., Kim, H.-K., Ko, S., Cao, P.-J., Suh, J.-P., Yi, G., Roh, J.-H. and Lee, S., 2009. Rice *Pi5*-mediated resistance to *Magnaporthe oryzae* requires the presence of two coiled-coil-nucleotide-binding-leucine-rich repeat genes. Genetics 181, 1627-1638.

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouzé, P. and Rombauts, S., 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Research 30, 325-327.

Librado, P. and Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451-1452.

Lin, F., Chen, S., Que, Z., Wang, L., Liu, X. and Pan, Q., 2007. The blast resistance gene *Pi37* encodes a nucleotide binding site-leucine-rich repeat protein and is a member of a resistance gene cluster on rice chromosome 1. Genetics 177, 1871-1880.

Liu, X., Lin, F., Wang, L. and Pan, Q., 2007. The in silico map-based cloning of *Pi36*, a rice coiled-coil-nucleotide-binding site-leucine-rich repeat gene that confers race-specific resistance to the blast fungus. Genetics 176, 2541-2549.

Lv, Q., Xu, X., Shang, J., Jiang, G., Pang, Z., Zhou, Z., Wang, J., Liu, Y., Li, T. and Li, X., 2013. Functional analysis of *Pid3-A4,* an ortholog of rice blast resistance gene *Pid3* revealed by allele mining in common wild rice. Phytopathology 103, 594-599.

Mardis, E.R., 2008. The impact of next-generation sequencing technology on genetics. Trends in genetics 24, 133-141.

Matys, V., Fricke, E., Geffers, R., Gößling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A.E. and Kel-Margoulis, O.V., 2003. TRANSFAC®: transcriptional regulation, from patterns to profiles. Nucleic Acids Research 31, 374-378.

McCallum, C.M., Comai, L., Greene, E.A. and Henikoff, S., 2000. Targeted screening for induced mutations. Nature biotechnology 18, 455-457.

McCouch, S.R., McNally, K.L., Wang, W. and Hamilton, R.S., 2012. Genomics of gene banks: A case study in rice. American journal of botany 99, 407-423.

McCouch, S.R., Sweeney, M., Li, J., Jiang, H., Thomson, M., Septiningsih, E., Edwards, J., Moncada, P., Xiao, J.

and Garris, A., 2007. Through the genetic bottleneck: *O. rufipogon* as a source of trait-enhancing alleles for *O. sativa*. Euphytica 154, 317-339.

Mejlhede, N., Kyjovska, Z., Backes, G., Burhenne, K., Rasmussen, S. and Jahoor, A., 2006. EcoTILLING for the identification of allelic variation in the powdery mildew resistance genes *mlo* and *Mla* of barley. Plant breeding 125, 461-467.

Milburn, D., Laskowski, R.A. and Thornton, J.M., 1998. Sequences annotated by structure: a tool to facilitate the use of structural information in sequence analysis. Protein engineering 11, 855-859.

Minoia, S., Petrozza, A., D'Onofrio, O., Piron, F., Mosca, G., Sozio, G., Cellini, F., Bendahmane, A. and Carriero, F., 2010. A new mutant genetic resource for tomato crop improvement by TILLING technology. BMC research notes 3, 69.

Nieto, C., Piron, F., Dalmais, M., Marco, C.F., Moriones, E., Gomez-Guillamon, M.L., Truniger, V., Gomez, P., Garcia-Mas, J. and Aranda, M.A., 2007. EcoTILLING for the identification of allelic variants of melon eIF4E, a factor that controls virus susceptibility. BMC plant biology 7, 34.

Okuyama, Y., Kanzaki, H., Abe, A., Yoshida, K., Tamiru, M., Saitoh, H., Fujibe, T., Matsumura, H., Shenton, M. and Galam, D.C., 2011. A multifaceted genomics approach allows the isolation of the rice *Pia*-blast resistance gene consisting of two adjacent NBS-LRR protein genes. The Plant Journal 66, 467-479.

Ou, S.H., 1985. Rice diseases, CABI, Egham.

Parry, M.A., Madgwick, P.J., Bayon, C., Tearall, K., Hernandez-Lopez, A., Baudo, M., Rakszegi, M., Hamada, W., Al-Yassin, A. and Ouabbou, H., 2009. Mutation discovery for crop improvement. Journal of Experimental Botany 60, 2817-2825.

Perry, J.A., Wang, T.L., Welham, T.J., Gardner, S., Pike, J.M., Yoshida, S. and Parniske, M., 2003. A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume Lotus japonicus. Plant Physiology 131, 866-871.

Qu, S., Liu, G., Zhou, B., Bellizzi, M., Zeng, L., Dai, L., Han, B. and Wang, G.-L., 2006. The broad-spectrum blast resistance gene *Pi9* encodes a nucleotide-binding site-leucine-rich repeat protein and is a member of a multigene family in rice. Genetics 172, 1901-1914.

Rai, A., Das, A., Gupta, Y., Thakur, S., Singh, P., Upreti, H., Singh, N. and Sharma, T., 2011a. Allele mining of rice blast resistance gene *Pi54* from wild and cultivated species of rice. Cereal Res Commun.

Rai, A.K., Kumar, S.P., Gupta, S.K., Gautam, N., Singh, N.K. and Sharma, T.R., 2011b. Functional complementation of rice blast resistance gene *Pi-k h* (*Pi54*) conferring resistance to diverse strains of *Magnaporthe oryzae*. Journal of Plant Biochemistry and Biotechnology 20, 55-65.

Rakshit, S., Rakshit, A., Matsumura, H., Takahashi, Y., Hasegawa, Y., Ito, A., Ishii, T., Miyashita, N.T. and Terauchi, R., 2007. Large-scale DNA polymorphism study of *Oryza sativa* and *O. rufipogon* reveals the origin and divergence of Asian rice. Theoretical and applied genetics 114, 731-743.

Ramkumar, G., Biswal, A., Mohan, K.M., Sakthivel, K., Sivaranjani, A., Neeraja, C., Ram, T., Balachandran, S., Sundaram, R. and Prasad, M., 2010. Identifying novel alleles of rice blast resistance genes *Pik h* and *Pita* through allele mining. Plant breeding.

Rozen, S. and Skaletsky, H., 1999. Primer3 on the WWW for general users and for biologist programmers, Bioinformatics methods and protocols. Springer, pp. 365-386.

Sabetta, W., Alba, V., Blanco, A. and Montemurro, C., 2011. sunTILL: a TILLING resource for gene function analysis in sunflower. Plant Methods 7, 20.

Schmid, C.D., Praz, V., Delorenzi, M., Périer, R. and Bucher, P., 2004. The Eukaryotic Promoter Database EPD: the impact of in silico primer extension. Nucleic Acids Research 32, D82-D85.

Shahmuradov, I.A., Gammerman, A.J., Hancock, J.M., Bramley, P.M. and Solovyev, V.V., 2003. PlantProm: a database of plant promoter sequences. Nucleic Acids Research 31, 114-117.

Shang, J., Tao, Y., Chen, X., Zou, Y., Lei, C., Wang, J., Li, X., Zhao, X., Zhang, M. and Lu, Z., 2009. Identification of a new rice blast resistance gene, *Pid3*, by genomewide comparison of paired nucleotide-binding site-leucine-rich repeat genes and their pseudogene alleles between the two sequenced rice genomes. Genetics 182, 1303-1311.

Sharma, T., Gupta, Y., Thakur, S., Singh, P., Upreti, H., Singh, N., Singh, U., Rathour, R., Kapoor, A. and Kaushik, R., 2009. Allele mining for important blast resistance genes from Indian land races of rice, 6 th International Rice Genetics Symposium. Manila, Philipines.

Sharma, T., Madhav, M., Singh, B., Shanker, P., Jana, T., Dalal, V., Pandit, A., Singh, A., Gaikwad, K. and Upreti, H., 2005. High-resolution mapping, cloning and molecular characterization of the *Pi-kh* gene of rice, which confers resistance to *Magnaporthe grisea*. Molecular Genetics and Genomics 274, 569-578.

Sharma, T., Rai, A., Gupta, S. and Singh, N., 2010. Broad-spectrum Blast Resistance Gene *Pi-kh* Cloned from Rice Line Tetep Designated as *Pi54*. Journal of Plant Biochemistry and Biotechnology 19, 87-89.

Sharma, T., Rai, A., Gupta, S., Vijayan, J., Devanna, B. and Ray, S., 2012. Rice blast management through host-plant resistance: retrospect and prospects. Agricultural Research 1, 37-52.

Skamnioti, P. and Gurr, S.J., 2009. Against the grain: safeguarding rice from rice blast disease. Trends in biotechnology 27, 141-150.

Slade, A.J., Fuerstenberg, S.I., Loeffler, D., Steine, M.N. and Facciotti, D., 2004. A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. Nature biotechnology 23, 75-81.

Smits, B.M., Mudde, J., Plasterk, R.H. and Cuppen, E., 2004. Target-selected mutagenesis of the rat. Genomics 83, 332-334.

Stemple, D.L., 2004. TILLING-a high-throughput harvest for functional genomics. Nature Reviews Genetics 5, 145-150.

Suzuki, T., Eiguchi, M., Kumamaru, T., Satoh, H., Matsusaka, H., Moriguchi, K., Nagato, Y. and Kurata, N., 2008. MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. Molecular Genetics and Genomics 279, 213-223.

Tabien, R., Li, Z., Paterson, A., Marchetti, M., Stansel, J. and Pinson, S., 2002. Mapping QTLs for field resistance to the rice blast pathogen and evaluating their individual and combined utility in improved varieties. Theoretical and applied genetics 105, 313-324.

Takahashi, A., Hayashi, N., Miyao, A. and Hirochika, H., 2010. Unique features of the rice blast resistance *Pish* locus revealed by large scale retrotransposon-tagging. BMC plant biology 10, 175.

Talamè, V., Bovina, R., Sanguineti, M.C., Tuberosa, R., Lundqvist, U. and Salvi, S., 2008. TILLMore, a resource for the discovery of chemically induced mutants in barley. Plant biotechnology journal 6, 477-485.

Talbot, N.J., 2003. On the trail of a cereal killer: exploring the biology of *Magnaporthe grisea*. Annual Reviews in Microbiology 57, 177-202.

Tamura, K., Dudley, J., Nei, M. and Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular biology and evolution 24, 1596-1599.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular biology and evolution 30, 2725-2729.

Tang, J., Zhu, X., Wang, Y., Liu, L., Xu, B., Li, F., Fang, J. and Chu, C., 2011. Semi-dominant mutations in the CC-NB-LRR-type R gene, *NLS1*, lead to constitutive activation of defense responses in rice. The Plant Journal 66, 996-1007.

Tanksley, S., Grandillo, S., Fulton, T., Zamir, D., Eshed, Y., Petiard, V., Lopez, J. and Beck-Bunn, T., 1996. Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. Theoretical and applied genetics 92, 213-224.

Taylor, N.E. and Greene, E.A., 2003. PARSESNP: a tool for the analysis of nucleotide polymorphisms. Nucleic Acids Research 31, 3808-3811.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25, 4876-4882.

Thompson, J.D., Higgins, D.G. and Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673-4680.

Till, B., Afza, R., Bado, S., Huynh, O., Jankowicz-Cieslak, J., Matijevic, M. and Mba, C., 2009. Global TILLING projects. Induced plant mutations in the genomics era. Food and Agriculture Organization of the United Nations, Rome, 237-239.

Till, B.J., Burtner, C., Comai, L. and Henikoff, S., 2004. Mismatch cleavage by single-strand specific nucleases. Nucleic Acids Research 32, 2632-2641.

Till, B.J., Colbert, T., Codomo, C., Enns, L., Johnson, J., Reynolds, S.H., Henikoff, J.G., Greene, E.A., Steine, M.N. and Comai, L., 2006a. High-throughput TILLING for Arabidopsis, Arabidopsis Protocols. Springer, pp. 127-135.

Till, B.J., Comai, L. and Henikoff, S., 2007a. TILLING and EcoTILLING for crop improvement, Genomics-assisted crop improvement. Springer, pp. 333-349.

Till, B.J., Cooper, J., Tai, T.H., Colowit, P., Greene, E.A., Henikoff, S. and Comai, L., 2007b. Discovery of chemically induced mutations in rice by TILLING. BMC plant biology 7, 19.

Till, B.J., Jankowicz-Cieslak, J., Sági, L., Huynh, O.A., Utsushi, H., Swennen, R., Terauchi, R. and Mba, C., 2010. Discovery of nucleotide polymorphisms in the Musa gene pool by Ecotilling. Theoretical and applied genetics 121, 1381-1389.

Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R., Young, K. and Taylor, N.E., 2003. Large-scale discovery of induced point mutations with high-throughput TILLING. Genome Research 13, 524-530.

Till, B.J., Zerr, T., Comai, L. and Henikoff, S., 2006b. A protocol for TILLING and Ecotilling in plants and animals. Nature Protocols 1, 2465-2477.

Tsai, H., Howell, T., Nitcher, R., Missirian, V., Watson, B., Ngo, K.J., Lieberman, M., Fass, J., Uauy, C. and Tran, R.K., 2011. Discovery of rare mutations in populations: TILLING by sequencing. Plant Physiology 156, 1257-1268.

Varshney, R.K., Graner, A. and Sorrells, M.E., 2005. Genomics-assisted breeding for crop improvement. Trends in plant science 10, 621-630.

Varshney, R.K., Nayak, S.N., May, G.D. and Jackson, S.A., 2009. Next-generation sequencing technologies and their implications for crop genetics and breeding. Trends in biotechnology 27, 522-530.

Wang, D., Guo, C., Huang, J., Yang, S., Tian, D. and Zhang, X., 2014. Allele-mining of rice blast resistance genes at *AC134922* locus. Biochemical and biophysical research communications 446, 1085-1090.

Wang, G.-L., Mackill, D.J., Bonman, J.M., McCouch, S.R., Champoux, M.C. and Nelson, R.J., 1994. RFLP mapping of genes conferring complete and partial resistance to blast in a durably resistant rice cultivar. Genetics 136, 1421-1434.

Wang, G.-X., Tan, M.-K., Rakshit, S., Saitoh, H., Terauchi, R., Imaizumi, T., Ohsako, T. and Tominaga, T., 2007. Discovery of single-nucleotide mutations in acetolactate synthase genes by Ecotilling. Pesticide biochemistry and physiology 88, 143-148.

Wang, J., Sun, J., Liu, D., Yang, W., Wang, D., Tong, Y. and Zhang, A., 2008a. Analysis of *Pina* and *Pinb* alleles in the micro-core collections of Chinese wheat germplasm by Ecotilling and identification of a novel *Pinb* allele. Journal of Cereal Science 48, 836-842.

Wang, L., Xu, X., Lin, F. and Pan, Q., 2009. Characterization of rice blast resistance genes in the *Pik* cluster and fine mapping of the *Pik-p* locus. Phytopathology 99, 900-905.

Wang, N., Shi, L., Tian, F., Ning, H., Wu, X., Long, Y. and Meng, J., 2010. Assessment of FAE1 polymorphisms in three Brassica species using EcoTILLING and their association with differences in seed erucic acid contents. BMC plant biology 10, 137.

Wang, N., Wang, Y., Tian, F., King, G.J., Zhang, C., Long, Y., Shi, L. and Meng, J., 2008b. A functional genomics resource for *Brassica napus*: development of an EMS

mutagenized population and discovery of FAE1 point mutations by TILLING. New Phytologist 180, 751-765.

Wang, X., Jia, Y., Shu, Q. and Wu, D., 2008c. Haplotype diversity at the *Pi-ta* locus in cultivated rice and its wild relatives. Phytopathology 98, 1305-1311.

Wang, Z.X., Yano, M., Yamanouchi, U., Iwamoto, M., Monna, L., Hayasaka, H., Katayose, Y. and Sasaki, T., 1999. The *Pib* gene for rice blast resistance belongs to the nucleotide binding and leucine-rich repeat class of plant disease resistance genes. The Plant Journal 19, 55-64.

Wienholds, E., van Eeden, F., Kosters, M., Mudde, J., Plasterk, R.H. and Cuppen, E., 2003. Efficient target-selected mutagenesis in zebrafish. Genome Research 13, 2700-2707.

Wisser, R.J., Sun, Q., Hulbert, S.H., Kresovich, S. and Nelson, R.J., 2005. Identification and characterization of regions of the rice genome associated with broad-spectrum, quantitative disease resistance. Genetics 169, 2277-2293.

Wu, J.-L., Wu, C., Lei, C., Baraoidan, M., Bordeos, A., Madamba, M., Suzette, R., Ramos-Pamplona, M., Mauleon, R. and Portugal, A., 2005. Chemical-and irradiation-induced mutants of indica rice IR64 for forward and reverse genetics. Plant molecular biology 59, 85-97.

Xu, X., Lv, Q., Shang, J., Pang, Z., Zhou, Z., Wang, J., Jiang, G., Tao, Y., Xu, Q. and Li, X., 2014. Excavation of *Pid3* Orthologs with Differential Resistance Spectra to *Magnaporthe oryzae* in Rice Resource. PloS one 9, e93275.

Yang, M.-Z., Cheng, Z.-Q., Chen, S.-N., Qian, J., Xu, L.-L. and Huang, X.-Q., 2007. A rice blast-resistance genetic resource from wild rice in Yunnan, China. Journal of plant physiology and molecular biology 33, 589-595.

Yang, S., Gu, T., Pan, C., Feng, Z., Ding, J., Hang, Y., Chen, J.-Q. and Tian, D., 2008. Genetic variation of NBS-LRR class resistance genes in rice lines. Theoretical and applied genetics 116, 165-177.

Yu, J., Hu, S., Wang, J., Wong, G.K.-S., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y. and Zhang, X., 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). Science 296, 79-92.

Yuan, B., Zhai, C., Wang, W., Zeng, X., Xu, X., Hu, H., Lin, F., Wang, L. and Pan, Q., 2011. The *Pik-p* resistance to *Magnaporthe oryzae* in rice is mediated by a pair of closely linked CC-NBS-LRR genes. Theoretical and applied genetics 122, 1017-1028.

Zenbayashi, K., Ashizawa, T., Tani, T. and Koizumi, S., 2002. Mapping of the QTL (quantitative trait locus) conferring partial resistance to leaf blast in rice cultivar Chubu 32. Theoretical and applied genetics 104, 547-552.

Zhai, C., Lin, F., Dong, Z., He, X., Yuan, B., Zeng, X., Wang, L. and Pan, Q., 2011. The isolation and characterization of *Pik*, a rice blast resistance gene which emerged after rice domestication. New Phytologist 189, 321-334.

Zhao, W., Cho, G.-T., Ma, K.-H., Chung, J.-W., Gwag, J.-G. and Park, Y.-J., 2010. Development of an allele-mining set in rice using a heuristic algorithm and SSR genotype data with least redundancy for the post-genomic era. Molecular breeding 26, 639-651.

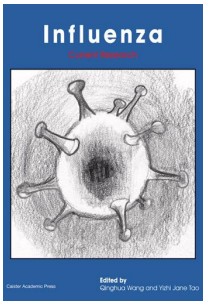
Zhou, B., Qu, S., Liu, G., Dolan, M., Sakai, H., Lu, G., Bellizzi, M. and Wang, G.-L., 2006. The eight amino-acid differences within three leucine-rich repeats between *Pi2* and *Piz-t* resistance proteins determine the resistance specificity to *Magnaporthe grisea*. Molecular plant-microbe interactions 19, 1216-1228.

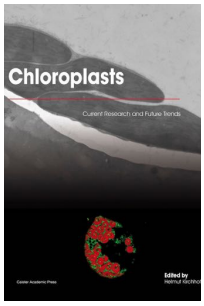
Zhu, J. and Zhang, M.Q., 1999. SCPD: a promoter database of the yeast Saccharomyces cerevisiae. Bioinformatics 15, 607-611.

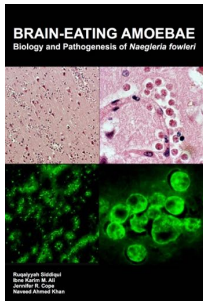
Zhu, X., Chen, S., Yang, J., Zhou, S., Zeng, L., Han, J., Su, J., Wang, L. and Pan, Q., 2012. The identification of *Pi50 (t)*, a new member of the rice blast resistance *Pi2/Pi9* multigene family. Theoretical and applied genetics 124, 1295-1304.

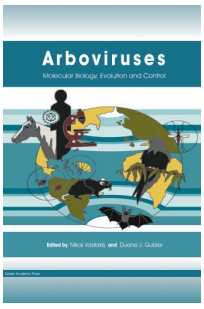
Microbiology / Molecular Biology

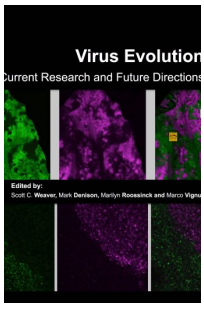
Caister Academic Press is a leading academic publisher of advanced texts in microbiology, molecular biology and medical research. Full details of all our publications at **caister.com**

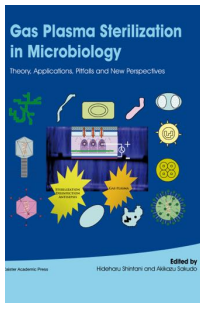




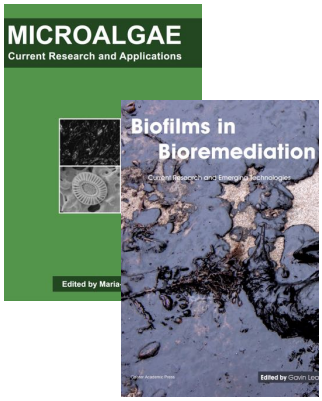


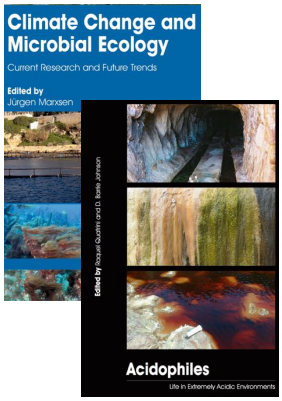






- **Cyanobacteria: Omics and Manipulation**
  Edited by: DA Los (2017)
  www.caister.com/cyano3
- **Brain-eating Amoebae: Biology and Pathogenesis of *Naegleria fowleri***
  Author: R Siddiqui, IKM Ali, JR Cope, et al. (2016)
  "explains the current knowledge and research" (ProtoView)
  www.caister.com/naegleria
- **Foot and Mouth Disease Virus: Current Research and Emerging Trends**
  Edited by: F Sobrino, E Domingo (2017)
  www.caister.com/fmdv
- ***Staphylococcus*: Genetics and Physiology**
  Edited by: GA Somerville (2016)
  www.caister.com/staph2
- **Chloroplasts: Current Research and Future Trends**
  Edited by: H Kirchhoff (2016)
  www.caister.com/chloroplasts
- **Microbial Biodegradation: From Omics to Function and Application**
  Edited by: J Długoński (2016)
  www.caister.com/biodegradation
- **Influenza: Current Research**
  Edited by: Q Wang, YJ Tao (2016)
  www.caister.com/flu3
- **MALDI-TOF Mass Spectrometry in Microbiology**
  Edited by: M Kostrzewa, S Schubert (2016)
  www.caister.com/malditof
- ***Aspergillus* and *Penicillium* in the Post-genomic Era**
  Edited by: RP Vries, IB Gelber, MR Andersen (2016)
  "new and well-presented book" (IMA Fungus)
  www.caister.com/aspergillus2
- **The Bacteriocins: Current Knowledge and Future Prospects**
  Edited by: RL Dorit, SM Roy, MA Riley (2016)
  www.caister.com/bacteriocins
- **Omics in Plant Disease Resistance**
  Edited by: V Bhadauria (2016)
  "essential reading ... highly recommended" (Biotechnol. Agron. Soc. Environ.)
  www.caister.com/opdr
- **Acidophiles: Life in Extremely Acidic Environments**
  Edited by: R Quatrini, DB Johnson (2016)
  "Contributors from a wide range of biological and environmental sciences" (ProtoView)
  www.caister.com/acidophiles





- **Climate Change and Microbial Ecology: Current Research and Future Trends**
  Edited by: J Marxsen (2016)
  "impressive" (ASM: Small Things Considered); "written at a high scientific level" (BioSpektrum)
  www.caister.com/climate
- **Biofilms in Bioremediation: Current Research and Emerging Technologies**
  Edited by: G Lear (2016)
  "describes explicitly the role of biofilms in bioremediation" (Biospektrum); indispensable ... recommended (Biotechnol. Agron. Soc. Environ.)
  www.caister.com/biorem
- **Microalgae: Current Research and Applications**
  Edited by: MN Tsaloglou (2016)
  www.caister.com/microalgae
- **Gas Plasma Sterilization in Microbiology: Theory, Applications, Pitfalls and New Perspectives**
  Edited by: H Shintani, A Sakudo (2016)
  "a nice state of the art compilation" (Doodys)
  www.caister.com/gasplasma
- **Virus Evolution: Current Research and Future Directions**
  Edited by: SC Weaver, M Denison, M Roossinck, et al. (2016)
  "highly informative ... a pleasure to read" (Microbiol. Today)
  www.caister.com/virusevol
- **Arboviruses: Molecular Biology, Evolution and Control**
  Edited by: N Vasilakis, DJ Gubler (2016)
  "a thorough and compelling review ... an outstanding book ... highly recommended" (Am. J. Trop. Med. Hyg.)
  www.caister.com/arbo
- ***Shigella*: Molecular and Cellular Biology**
  Edited by: WD Picking, WL Picking (2016)
  www.caister.com/shigella
- **Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment**
  Edited by: AM Romaní, H Guasch, MD Balaguer (2016)
  "essential reference book" (Biotechnol. Agron. Soc. Environ.)
  www.caister.com/aquaticbiofilms
- **Alphaviruses: Current Biology**
  Edited by: S Mahalingam, L Herrero, B Herring (2016)
  "up-to-date review of the field" (Aus. Vet. J.)
  www.caister.com/alpha

Order from Book Systems Plus caister.com/order +441440706716 bsp2b@aol.com