Biochemical and Biophysical Research Communications 446 (2014) 1085–1090

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Allele-mining of rice blast resistance genes at *AC134922* locus

Dan Wang [1], Changjiang Guo [1], Ju Huang, Sihai Yang, Dacheng Tian *, Xiaohui Zhang *

State Key Laboratory of Pharmaceutical Biotechnology, School of Life Sciences, Nanjing University, 210093 Nanjing, China

ARTICLE INFO

Article history:
Received 1 March 2014
Available online 21 March 2014

Keywords:
Resistance
AC134922
Allele-mining
Gene conversion

ABSTRACT

The *AC134922* locus is one of the most rapidly evolving nucleotide binding site-leucine-rich repeat (*NBS-LRR*) gene family in rice genome. Six rice blast resistance (*R*) genes have been cloned from this locus and other two resistance candidate genes, *Pi34* and *Pi47*, are also mapped to this complex locus. Therefore, it seems that more functional *R* genes could be identified from this locus. In this study, we cloned 22 genes from 12 cultivars based on allele-mining strategy at this locus and identified 6 rice blast *R* genes with 4 of them recognizing more than one isolates. Our result suggests that gene stacking might be the evolutionary strategy for complex gene locus to interact with rapidly evolving pathogens, which might provide a potential way for the cloning of durable resistance genes. Moreover, the mosaic structure and ambiguous ortholog/paralog relationships of these homologous genes, caused by frequent recombination and gene conversion, indicate that multiple alleles of this complex locus may serve as a reservoir for the evolutionary novelty of these *R* genes.

1. Introduction

Rice (*Oryza sativa*) is a staple food consumed by nearly half the world's population. Blast disease, caused by the most devastating pathogen of rice, *Magnaporthe oryzae*, is a recurrent problem in all rice-growing regions of the world. The pathogen infect rice plant at all growth stages from seedling to grain formation, affecting leaves, nodes, collars, panicles and roots resulting total loss of the rice grain [1]. Plants have developed multiples layers of defense against various plant pathogens, including structural, chemical and protein-based defense. There are two lines of protein-based defense in plants [2]. Basal resistance is the first line to protect plants against entire group of pathogens. If a pathogen can suppress the first line, plants may respond with the second one: the hypersensitive response (HR). The HR is typically more specific than the first line and is often triggered when the products of plant *R* genes specially recognize the pathogen avirulence (Avr) effectors. Interactions and co-evolution between *R* genes and pathogen effectors are key to durable resistance of *R* genes [3].

Therefore, breeding for *R* genes is considered one of the best cultivation strategies for disease management [4]. However, host resistance is short-lived due to the rapid speed of evolution and high level of variability in the pathogen population [5]. Thus isolation of plenty of resistance genes as resources to defend against a variety of pathogens is required to develop durable blast resistance rice varieties. Indeed, stacking of *R* genes that confer resistance to a broad spectrum of isolates promises to deliver rice with durable blast resistance [6].

Natural selection drives pathogen effectors and *R* genes to evolve rapidly in an evolutionary "arm race" relationship [7,8]. It is suggested that more *R* genes are detected in complex loci than in simple loci [9]. Rapidly evolving complex families are characterized by a multiple and variable copy number, ambiguous ortholog/paralog relationship and high ratio of *Ka*/*Ks* [10]. Such as *L* in flax and *RPP4*/*RPP5* in *Arabidopsis* [11,12], polymorphic analysis has suggested that these *R* genes have high level of diversity between or even within populations [13]. This allelic diversity is the essential base for rapid evolution of *R* genes so as to expand recognition specificity, and supplies us a huge resource to identify more durable and plenty of functional genes at allelic *R* genes from global rice varieties. Such as the *AC134922* locus, one of the largest highly diversified *R* gene families in rice [10], six rice blast resistance (*R*) genes have been cloned from this locus [14] and other two resistance candidate genes, *Pi34* and *Pi47* [15,16], are also mapped to this complex locus. It seems that more functional *R* genes could be identified from this locus. Therefore, we chose this locus to survey its evolutionary patterns and to clone more rice blast *R* genes.

In the last two decades, lots of plant *R* genes have been cloned and most of the genes encode NBS-LRR proteins [17]. However, it is far from enough for breeding resistant cultivars. Map-based cloning, the traditional method used to identify *R* genes is

* Corresponding authors. Fax: +86 2583592740.
E-mail addresses: dtian@nju.edu.cn (D. Tian), xiaohuizhang@nju.edu.cn (X. Zhang).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.bbrc.2014.03.056
0006-291X/

time-consuming and costly. Even though the method has been improved by combing with other strategies, it is still difficult to clone some complex gene locus. Another cloning method is allele-mining, a high-throughput strategy to isolate *R* gene homologues and detect functional *R* genes. It is widely used in cloning potato late blight *R* genes, such as the homologues of *Rpi-blb1* in *Solanum bulbocastanum*, the resistance genes *Rpi-sto1* and *Rpi-pta1* in *S. stoloniferum* and *S. papita* [18]. Furthermore, the wheat powdery mildew resistance alleles of *Pm3* and the rice blast resistance gene *Pi54* (*Pik^h*) were also identified using this approach [19,20]. In this study, our research based on allele-mining strategy at *AC134922* locus could provide more comprehensive and thorough understanding about the complexity of the *NBS-LRRs* in rice. Meanwhile, the six new *R* genes we identified at this locus demonstrated the effective of the allele-mining method, which will provide guidance for cloning other resistance genes, even other important genes in crops.

2. Materials and methods

2.1. Allele-mining and blast resistance assays

Seeds of the rice lines used were from various sources (Table S1). Genomic DNA was extracted using the cetyltrimethyl ammonium bromide (CTAB) method. Degenerated primers incorporated the putative start and stop codons of candidate *AC134922* gene homologs were designed based on the alignment of all the available rice-derived *AC134922*-like sequences (Table S2). The products of long PCR were inserted into the binary vector (pCAMBIAI300-AC134922), which contained the native promoter and terminator from the *AC134922*. The sequenced and validated clones, were transferred into blast-susceptible rice cultivars (TP309 and Shin2) using Agrobacterium strain EHA105. All transgenic lines (T0) were reproduced to obtain enough seeds (T1 or T2). The presence of the transgenic DNA fragments in the T0 plants was confirmed by PCR.

Blast strains were collected and isolated from different areas of China in 2008–2009 (Table S3). The transgenic line screening and gene expression assays was according to Yang et al. [14]. The transformed line was required to exhibit a consistent *R* phenotype against at least one of the 17 blast strains across all three replicates.

2.2. Sequence alignment and phylogenetic analysis

Homologous sequences at *AC134922* locus were reported before [14,21]. Alignments of coding sequences (CDSs) were toggled using Clustal W [22] inside MEGA version 5.0 [23] with default parameters. Based on the alignment results, phylogenetic trees were generated using the bootstrap neighbor-joining (NJ) method with the Kimura two-parameter model in MEGA v5.0. The stability of internal nodes was assessed by bootstrap analysis with 1000 replicates.

Nucleotide diversity (π) was estimated with the Jukes and Cantor correction using DnaSP v5.0 [24]. GENECONV1.81 was used to investigate sequence exchanges [25]. The default setting of 10,000 permutations was used for the analysis. The statistical significance of gene conversion events was defined as a global permutation *P* value of <0.05.

To detect positive selection, the ratios of nonsynonymous to synonymous nucleotide substitutions (*Ka*/*Ks*) were calculated using DnaSP v5.0. And the HyPhy package with the random effects likelihood (REL) method was used to further detect positively selective sites as implemented on the Datamonkey web server [26].

3. Results and discussion

3.1. Allele-mining and screening NBS-LRRs in the AC134922 locus

The *AC134922* gene family was the most diversified and complex *NBS-LRR* locus featured with multiple and variable copy numbers, low divergence of paralogs and high level of within-species diversity in the rice genome [14]. Previous research has showed six rice blast *R* genes identified from this locus [14]. In this study, we adopted the allele-mining strategy to dig more candidate *R* genes at the *AC134922* locus. A total of 22 genes were cloned from 12 cultivars and 39 transgenic lines with enough first generation (T1) or second generation (T2) seeds were gained for the subsequent screening of rice blast. Seventeen blast isolates from 85 strains collected throughout China were chosen based on their high level of nucleotide diversities [14], their geographical distribution, and their ability to produce large numbers of spores (Table S3). The transformed rice lines were exposed to spores from each of the 17 strains of rice blast disease. Their sensitivity or resistance to each line was classified based on their phenotypes (Fig. S1).

Repeated tests and screening of the transformants showed that 6 new *R* genes from 4 cultivars and 11 lines conferred resistance to one or more rice blast isolates (Table S4). Yang et al. have detected 6 rice blast *R* genes in this locus [14]. However, three of them conferred resistance to only one isolate. The same phenomenon was also detected in our research. Only 4 genes (18.2%) from 3 cultivars and 7 (17.9%) lines were identified as resistant to at least two strains. Furthermore, only *BG1* gene had relatively broad-spectrum resistance, which is resistant to 5 of the 8 screened isolates. These results suggested that the *AC134922* locus might have no or few broadly resistant members; on the contrary, it contained more moderate *R* genes to deal with different isolates. This is similar to the *Rpi* and *R3* loci on chromosome IV and chromosome XI in potato, both of which are *R*-gene clusters with multiple genes recognizing different races of *Phytophthora infestans* [27]. Similarly, thirteen alleles of the *L* locus have been described and each of which confers a different rust-resistance specificity [11]. Stacking of these functional alleles results in a broad spectrum resistance. This may be an evolutionary strategy for complex gene locus to interact with various rapidly evolving pathogens.

In the other hand, the *R* genes we identified in the *AC134922* locus came from various rice lines but not concentrated in one or several lines. In general, the resistant cultivars contained more candidate resistant genes, such as Gumei2 and Tetep (Table S5). Particularly, there are three genes respectively cloned from Tetep and Gumei2 were resistant, although most of the *R* genes only conferred to one or two isolates (Table S4). This phenomenon further indicated that gene stacking or gene pyramiding might be greatly helpful to breeding durable resistance cultivars.

3.2. Phylogenetic analyses of homologous genes in the AC134922 locus

Previous research of neighbor-joining tree of the homologs at the *AC134922* family in three sequenced rice genome, *Nipponbare*, *9311* and *GLA4*, has demonstrated its high divergent feature [21]. Here, we constructed a phylogenetic tree based on the CDS of the cloned genes, together with the homologs downloaded and the genome sequences from *Nipponbare* and *9311* at the *AC134922* locus (Fig. 1). All the homologous sequences in the phylogenetic tree could be divided into eight multi-gene groups based on the following three conditions: the branches at the root of each group were supported by high confident bootstrap values (>80%), the average nucleotide similarity among the gene members exceeded 80% and each group must contain at least one genome sequence. As