



Nucleotide variation and identification of novel blast resistance alleles of *Pib* by allele mining strategy

G. Ramkumar · M. S. Madhav · S. J. S. Rama Devi ·
M. S. Prasad · V. Ravindra Babu

Received: 26 November 2014 / Revised: 20 January 2015 / Accepted: 2 March 2015 / Published online: 11 March 2015
© Prof. H.S. Srivastava Foundation for Science and Society 2015

Abstract *Pib* is one of significant rice blast resistant genes, which provides resistance to wide range of isolates of rice blast pathogen, *Magnaporthe oryzae*. Identification and isolation of novel and beneficial alleles help in crop enhancement. Allele mining is one of the best strategies for dissecting the allelic variations at candidate gene and identification of novel alleles. Hence, in the present study, *Pib* was analyzed by allele mining strategy, and coding and non-coding (upstream and intron) regions were examined to identify novel *Pib* alleles. Allelic sequences comparison revealed that nucleotide polymorphisms at coding regions affected the amino acid sequences, while the polymorphism at upstream (non-coding) region affected the motifs arrangements. *Pib* alleles from resistant landraces, Sercher and Kregosa showed better resistance than *Pib* donor variety, might be due to acquired mutations, especially at LRR region. The evolutionary distance, Ka/Ks and phylogenetic analyzes also supported these results.

Electronic supplementary material The online version of this article (doi:10.1007/s12298-015-0284-4) contains supplementary material, which is available to authorized users.

G. Ramkumar · M. S. Madhav (✉) · S. J. S. R. Devi ·
M. S. Prasad · V. R. Babu
Crop Improvement Section, Directorate of Rice Research,
Rajendranagar, Hyderabad 500030, India
e-mail: sheshu24@gmail.com

M. S. Madhav
e-mail: sheshu_24@yahoo.com

G. Ramkumar
e-mail: tgramkumar2000@yahoo.com

S. J. S. R. Devi
e-mail: sagilirama@gmail.com

M. S. Prasad
e-mail: msrinivasprasad@yahoo.co.in

V. R. Babu
e-mail: vrbabu@dricar.org

Transcription factor binding motif analysis revealed that *Pib*^{Sr} had a unique motif (DPBF COREDCDC3), while five different motifs differentiated the resistance and susceptible *Pib* alleles. As the *Pib* is an inducible gene, the identified differential motifs helps to understand the *Pib* expression mechanism. The identified novel *Pib* resistant alleles, which showed high resistance to the rice blast, can be used directly in blast resistance breeding program as alternative *Pib* resistant sources.

Keywords Allele and promoter mining · TFBSs · Rice blast · Land races · *Pib*

Introduction

Rice is one of the most important and staple food crops, which is produced and consumed worldwide. Biotic stresses are the major constraints in rice production, lead to heavy economic loss. Among the biotic stresses, rice blast is one of the devastating diseases, which affects rice yield up to 90 % in susceptible varieties (Khush and Jena 2009). Utilization of resistance sources is considered as one of the best methods to combat this biotic stress problem. Host plant resistance is a complex process, for which more than hundred blast resistance genes and few QTLs were reported (Das et al. 2012). Among them, nineteen genes have been cloned and characterized (Das et al. 2012), which includes *Pib*. *Pib* is a major resistance gene, offers resistance to wide range of isolates of India. It is 5404 bp in size with three exons, encodes for 1251 amino acids, and belongs to NBS-LRR gene family (Acc. No. AB026839) (Wang et al. 1999). As many beneficial alleles have been left behind during the rice domestication, the untapped gene resources can be utilized for the crop enhancement (McCouch et al. 2007; Ramkumar et al. 2010). The availability of the rice genome sequences and bioinformatic

tools enhance the applicability of the allele mining strategy. It is well documented that the mutation in the exonic region affects the protein structure and hence the phenotype of the plant. However, many studies (Ramkumar et al. 2010, 2014) have been reported that the variation at the upstream region and introns of the gene also affects the gene expression and hence leads to the phenotypic variation. Hence the ‘true’ allele mining should include the coding and non-coding regions as well. Based on the above status, the present study was designed to 1. Identify and isolate novel alleles of *Pib*; 2. Analyze nucleotide variations at coding and non-coding regions; 3. Evaluate polymorphism at Transcription factor binding motifs (TFBMs) level among the ecotypes.

Materials and methods

Plant materials and screening with blast isolates

Plant materials of this study include twenty four landraces collected from North Eastern region of India (Supplementary table 1). All the plant materials were screened for the rice blast resistance for twice with three replicates in the consecutive seasons as follows. Spore suspension (approximately 10^5 spores per ml mixed with 0.2 % of Tween-20) of MNP (Manipur)-4 isolate (not compatible with *Pib*) was sprayed on 15 days old plants using glass atomizer. Sprinklers were used to maintain artificial humidity of 80 to 90 %. After 12 days of inoculation, the disease symptoms were scored with 0–9 scale. Engkatek was used as resistance control, while Co-39 and HP2216 were included for screening as susceptible controls along with other test landraces. Nipponbare was also included for the phenotype screening. The dominant molecular marker Nsb (atcaactctgccacaaaatcc/ cccatataccactgttcccc) was also used to confirm the presence/ absence of *Pib* resistance alleles.

PCR based allele mining

Primers were designed to amplify the complete allele, including non-coding regions, based on the *Pib* sequence information (Acc. No: AB026839) using online tool, Primer 3 (Rozen and Skaletsky 2000). As the allele is ~5.5 Kb in size, it was amplified in two segments with 500 bp overlapping region. Each of the 3 Kb allelic fragment is amplified with the PCR mixture contained 250–350 ng template DNA, 5 picomoles of each primer (Supplementary table 2), 10 mM dNTPs, 1 X PCR buffer (10 mM Tris, pH 8.4, 50 mM KCl, 1.5 mM MgCl₂ and 0.01 mg/ml gelatin) and 2 U of High fidelity *Taq* DNA polymerase (Fermentas USA) in a reaction volume of 20 μ l. Thermal profile followed was: Initial denaturation at 94 °C for 5 min and followed by 35 cycles of 60 s denaturation at 94 °C, 60 s annealing at 58 °C and 3 min and 30 s of primer extension at 72 °C and a final extension at 72 °C for 10 min. The

amplicons were eluted from 0.8 % 1X TAE agarose gel using Promega elution kit (Promega USA) and cloned from selected landraces using ZeBaTA cloning vector (Chen et al. 2009). Four clones per amplicon were sequenced and only high quality sequences (Phred score >20 per base) were considered for further sequence analysis. *Pib*^{Nb} allele also was included in this comparative study, which was available at Genbank (Acc. No. NC_008395).

Nucleotide and amino acid sequence analysis

The raw sequences of *Pib* alleles of selected genotypes were compared with *Pib* reference sequence using NCBI sequence alignment tool (<http://blast.ncbi.nlm.nih.gov/>) and MEGA 4.1 (Tamura et al. 2007). Evolutionary distance (θ_{π}) of the sequences in comparison with the reference sequence was also calculated by MEGA. In that analysis, standard error was calculated by bootstrap value with 1000 replicates. Phylogenetic tree was constructed with MEGA - Neighbor-Joining (NJ) method. Gene structure was annotated using online tool softberry- FGENESH (<http://linux1.softberry.com>). Number of SNPs, InDels, and Ka/Ks value of the allelic sequences were calculated using tool, DnaSP version 5.10.01 (Librado and Rozas 2009).

Results and discussion

Screening of plants with isolates

The landraces, collected from North Eastern India are not domesticated well and hence they may harbor untapped beneficial alleles (Choudhury et al. 2013). Hence, landraces were selected for this novel allele identification study. The 24 landraces, collected from North Eastern part of India, were screened with differential isolate MNP (Manipur)-4. In the screening experiment, *Pib* donor, Engkatek showed resistance reaction with mild symptoms that was scored as 1, while the susceptible checks, Co-39 and HP2216 showed high susceptibility (scored as 9). Phenotype reactions of landraces were provided in the supplementary table 1. Based on the phenotypic reaction to the differential isolate, three different landraces were selected for *Pib* allele mining i.e. two landraces, Sercher and Kregosa, which showed complete resistance and were scored as 0, while another landrace, Podumoni Ahu, showed extreme susceptibility. Nipponbare also showed susceptible phenotype. Alleles of *Pib* was amplified, cloned and sequenced from these three landraces (*Pib*^{Kr}, *Pib*^{Sr}, and *Pib*^{Pa}-derived from Kregosa, Sercher and Podumoni Ahu, respectively) and the sequences were used for further analysis. The Genbank accession numbers of all alleles are provided in the Table 1.

Table 1 List of *Pib* alleles, Acc. No. and their polymorphic data in comparison with *Pib* reference allele

Name of allele	GenBank Acc. No.	Phenotype	Allele		ORF length (bp)	No of ORFs	Ka/Ks value	Protein sequence		Amino acid length	Evolutionary distance
			No of SNPs	No of InDels				AA substitutions	No of InDels		
<i>Pib</i> ^{Ek}	AB026839	R	–	–	3789	3	–	–	–	1262	–
<i>Pib</i> ^{Sr}	JN564624	R	65	3	3381	4	1.083	19	2	1126	0.012
<i>Pib</i> ^{Kr}	JN564625	R	27	0	3789	3	0.941	9	–	1262	0.005
<i>Pib</i> ^{Pa}	JN564623	S	40	8	3225	5	1.681	15	3	1074	0.007
<i>Pib</i> ^{Nb}	NC_008395	S	867	95	3039	7	0.789	NM	NM	1012	0.177

NM numerous

Allelic diversity analysis

The derived sequences were compared with the reference sequence (*Pib*^{Ek} derived from Engkatek), which revealed that all the test alleles had more than 90 % similarity. Structural analysis of alleles revealed that *Pib*^{Kr} had three ORFs (3789 bp), matched with the *Pib*^{Ek} ORF where as *Pib*^{Sr} and *Pib*^{Pa} had the allelic lengths of 3225 and 3381 bp with four and five ORFs, respectively. The allele derived from Nipponbare (*Pib*^{Nb}) had the highest number of seven ORFs with the least allelic length of 3039 bp. Sequence analysis with DnaSP revealed that *Pib*^{Sr} had 65 SNPs and three InDels, while *Pib*^{Pa} had 40 SNPs and eight InDels. Interestingly, resistant allele, *Pib*^{Kr}, had the least variations of 27 SNPs without any InDels, while the *Pib*^{Nb} had the highest number of SNPs (867) and InDels (95). Though Sercher and Kregosa are landraces, their genomes are different from each other and might have acquired different level of polymorphisms. Hence, though both are resistant landraces, their similarity with the reference allele differ. Amino acid comparison with *Pib*^{Ek} indicated that *Pib*^{Sr} had the higher number of polymorphism (19 AA substitutions and two InDels), followed by the *Pib*^{Pa} (15 AA substitutions and three InDels) while the least AA substitutions were observed for *Pib*^{Kr} with 9 AA substitutions and no InDels, which indicated that though this allele had relatively more polymorphisms at nucleotide level, it had lesser alterations at the protein level. NBS and LRR regions of reference and test amino acid sequences were compared to assess the polymorphism at conserved region. *Pib*^{Sr} showed 1 and 4 AA substitutions at NBS and LRR region, respectively. In case of *Pib*^{Kr}, 2 AA substitutions were observed at LRR region. In case of *Pib*^{Pa}, 6 AA substitutions and 1 InDel were associated with LRR region, while one AA substitutions was at NBS region. Significantly, the observed InDel at LRR region of susceptible allele, was not observed in resistant alleles. Resistant allele, *Pib*^{Kr} had comparatively lesser divergence from reference allele, which may be under purifying selective pressure. Both

the resistant alleles acquired mutations at LRR region, which is a major determinant of recognition specificity of pathogen avirulence factors (Meyers et al. 1999), might have lead to the positive selection of these alleles.

Surprisingly, Ka/Ks value were significantly higher for all the analyzed *Pib* alleles which indicated presence of more non synonymous polymorphisms over synonymous variations. This analysis also indicated that the susceptible allele *Pib*^{Pa}, had the highest non synonymous mutations, while the resistant allele *Pib*^{Kr} had lower non-synonymous mutations (Table 1). Not surprisingly, *Pib*^{Nb} allele had the highest evolutionary distance value (0.177) with Ka/Ks ratio of 0.789, which indicated that this allele had high deviation from reference allele. This may be due to that the *Pib* allele was originated from indica (Miyamoto et al. 1996), while the Nipponbare belongs to japonica subspecies of *O. sativa*.

Phylogenetic analysis for *Pib* alleles

Phylogeny analysis was performed to know the similarity among the *Pib* alleles. All *Pib* alleles were compared with each other along with the reference *Pib* allele. In this analysis, the resistant allele *Pib*^{Kr} was grouped with reference allele, but another resistant allele *Pib*^{Sr} was grouped with susceptible allele, based on overall sequence similarity among the alleles. Whereas the *Pib*^{Nb} allele was clearly out clustered, due to high nucleotide polymorphisms (Fig. 1).

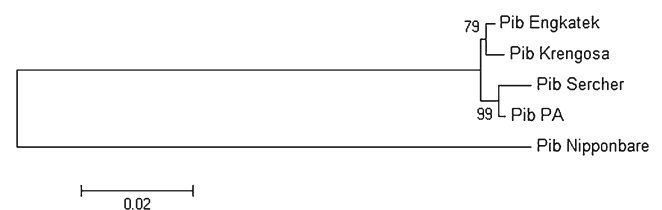


Fig. 1 Phylogenetic tree of *Pib* alleles based on upstream and allelic sequences with bootstrap value

Pib Promoter mining

The transcription factor binding motifs in the promoters of the alleles were compared with each other, which revealed common polymorphic motifs, which differentiated resistant and susceptible promoter alleles. Promoter allele of *Pib*^{Pa} had five additional motifs, viz., MYBST1 (core motif of MybSt1), SREATMSD, TATCCACHVAL21 (TATC CAC box), TATCCAOSAMY (TATCCA element) and TATCCAYMOTIFOSRAMY3D (sugar and amylose repression element) (Higo et al. 1999), and these motifs were present in *Pib*^{Nb} allele also. But, these motifs were absent in *Pib* resistant promoter alleles (*Pib*^{Kr} and *Pib*^{Sr}) and hence these motifs differentiated resistant promoter alleles. The *Pib*^{Sr} promoter allele had a unique motif i.e. DPBFCOREDCDC3 (bZIP transcription factor in carrot), which was absent in all other alleles.

The present study revealed that the landraces had significant diversity at *Pib* allelic region. The nucleotide polymorphisms at alleles corresponded well to the amino acids change and consequently to the phenotypic reactions of the landraces in this study. Interestingly, five *cis* motifs, which could differentiate the resistance and susceptible alleles, were also identified in this study. It is noteworthy that the *Pib* allele is inducible resistance gene, and hence, the identified polymorphic *cis* motifs may help to understand the resistance genes regulation better. The identified *Pib* resistant allele showed significant resistance to rice blast isolates, which can be used as alternative *Pib* resistant allele in the blast resistance breeding programs.

Acknowledgments The Authors thank the Department of Biotechnology, Government of India for funding this research work. S. J. S. Rama Devi thanks Council for Scientific and Industrial Research (CSIR) for providing fellowship.

Conflict of interest The authors declare no conflict of interest.

Authors contribution MSM conceived and designed the experiment; GR and SJSRD performed experiments; GR and MSM prepared manuscript; MSP and VRB analyzed data and improved manuscript.

References

- Chen S, Songkumarn P, Liu J, Wang GL (2009) A versatile zero background T-vector system for gene cloning and functional genomics. *Plant Physiol* 150:1111–1121
- Choudhury B, Khan ML, Dayanandan S (2013) Genetic structure and diversity of indigenous rice (*Oryza sativa*) varieties in the Eastern Himalayan region of Northeast India. *SpringerPlus* 2:228
- Das A, Soubam D, Singh PK, Thakur S, Singh NK, Sharma TR (2012) A novel blast resistance gene, *Pi54th* cloned from wild species of rice, *Oryza rhizomatis* confers broad spectrum resistance to *Magnaporthe oryzae*. *Funct Integr Genomics* 12:215–228
- Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database. *Nucleic Acids Res* 27: 297–300
- Khush GS, Jena KK (2009) Current status and future prospects for research on blast resistance in rice (*Oryza sativa* L.). In: Wang GL (ed) *Advances in genetics, genomics and control of rice blast disease*. Springer, New York, pp 1–10
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452
- McCouch SR, Sweeney M, Li J, Jiang H, Thomson M, Septiningsih E, Edwards J, Moncada P, Xiao J, Garris A, Tai T, Martinez C, Tohme J, Sugiono M, McClung A, Yuan LP, Ahn SN (2007) Through the genetic bottleneck: *O. rufipogon* as a source of trait-enhancing alleles for *O. sativa*. *Euphytica* 154:317–339
- Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral BW, Young ND (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide binding superfamily. *Plant J* 20:317–332
- Miyamoto M, Ando I, Rybka K, Kodama O, Kawasaki S (1996) High resolution mapping of the indica derived rice blast resistance genes I. *Pib*. *Mol Plant-Microbe Interact* 9:6–13
- Ramkumar G, Sakthivel K, Sundaram RM, Neeraja CN, Balachandran SM, Rani NS, Viraktamath BC, Madhav MS (2010) Allele mining in crops: prospects and potentials. *Biotechnol Adv* 28:451–461
- Ramkumar G, Madhav MS, Ramadevi SJS, Manimaran P, Mohan KM, Prasad MS, Balachandran SM, Neeraja CN, Sundaram RM, Viraktamath BC (2014) Nucleotide diversity of Pita, a major blast resistance gene and identification of its minimal promoter. *Gene* 546:250–256
- Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132:365–368
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
- Wang ZX, Yano M, Yamanouchi U, Iwamoto M, Monna L, Hayasaka H, Katayose Y, 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. *Plant J* 19:55–64