Life Science Informatics Publications

Research Journal of Life Sciences, Bioinformatics, Pharmaceutical and Chemical Sciences

Journal Home page http://www.rjlbpcs.com/

RJLBPCS

ISSN 2454-6348

Original Research Article **DOI: 10.26479/2018.0404.36**

IN-SILICO PROMOTER ANALYSIS OF RICE BLAST RESISTANCE GENES IN BR2655 AND HR12 RICE CULTIVARS

R Chandrakanth[1], K Narasimha Murthy[2], N S Devaki[1*]

1. Department of Molecular Biology, Yuvaraja's College, University of Mysore, Mysuru, Karnataka, India.

2. Department of Studies in Biotechnology, University of Mysore, Manasagangotri, Mysuru, Karnataka, India.

ABSTRACT: Rice blast caused by fungus *Magnaporthe oryzae* B.C. Couch, is generally considered as the most devastating diseases of rice (*Oryza sativa* L.) worldwide. Plant resistance to pathogens is a complex process where it relies on the interaction of plant resistance (*R*) genes and pathogen derived Avirulence (*Avr*) genes derived proteins. Most of the resistance proteins contain a central nucleotide binding (NB) domain, also known as the NB-ARC domain and C-terminal leucine rich repeat (LRR) domain. More than 100 genes coding for such resistance proteins are identified where 30 genes are well characterized. In the present study, an attempt was made to analyze promoter regions of the ten resistance genes. Gene sequences of 10 resistance genes were retrieved from Rice Genome Annotation Project (RGAP). 2 kb of upstream sequences of transcription start sites were taken from RAP-DB (The Rice Annotation Project-database). PlantCare and PLACE were used for identification of cis-regulatory elements in promoters of resistance gene sequences. The cis-regulatory elements so obtained were listed. PlantPAN and Cpgplot were used for analysis of CpG islands. Multiple Sequence alignment was carried out by Clustalw and phylogenetic tree was constructed using MEGA X. We identified different classes of cis-regulatory elements and CpG islands which are involved in transcriptional regulation of resistance genes during plant defense mechanism. This analysis will assist in obtaining a deeper insight on the regulation of resistance genes during host pathogen interaction. Outcome of the present investigation indirectly adds to the knowledge required for resistance breeding.

KEYWORDS: Rice, Blast Resistance, Promoter, cis-regulatory element, CpG Island.

Corresponding Author: Dr. N S Devaki* Ph.D.

Department of Molecular Biology, Yuvaraja's College, University of Mysore, Mysuru, Karnataka, India. Email Address:devakineerkaje@gmail.com

1. INTRODUCTION

Rice (*Oryza sativa* L.) is an important staple food crop for more than half of world population which can provide nearly quarter of total energy intake of human population (FAO, 2014). It is cultivated in more than 160 million hectares worldwide and provides the daily energy for over 3.5 billion people [1]. Rice consumption is increasing and demand for rice is also growing with population growth. Production must be increased by >40% by 2030 [2]. To meet this challenge we need rice varieties with higher yield potential and greater yield stability. Although yield potential of rice is 10 tons per hectare, farmers on an average harvest about 5 tons per hectare from irrigated lands. This yield gap is due to the losses caused by biotic and abiotic stresses. Among the biotic stresses, blast disease is the most important one. Rice blast disease caused by the filamentous ascomycetes fungus *Magnaporthe oryzae* B. C. Couch is the most devastating fungal disease in the rice-growing regions of the world thus resulting in huge yield losses [3]. The Fungus is a notorious pathogen among 36 major fungal pathogens reported from rice, having capability to reduce world food grain production by 8% per year [4]. Rice blast is one of the most distributed plant diseases being reported world-wide [5][6]. Blast fungus has a hemibiotrophic phase in its life cycle [7], which can infect many growth stages of rice and causes up to 100% loss of rice production in infected areas [8]. Annual rice harvest amounting to 30% is lost due to rice blast infection, which is enough to feed more than 60 million people for one year [9]. Blast disease is a major constraint for rice cultivation in different agro-climatic regions of India. India is the largest rice growing country accounting for about one third of the world acreage under the crop. India's annual rice production is 103.6 million tons during 2016 [10]. In India, management of rice blast disease is done using chemical fungicides due to low levels of host plant resistance in the cultivated rice varieties. The use of chemicals is neither practical nor environment-friendly for disease control and hence utilisation of host resistance has been the best way to manage the disease, for which identification of sources of resistance genes is necessary. Advances in molecular genetics and completion of the genome sequence of rice paved the way for cloning and characterization of major genes for blast resistance. Further, many major resistance genes have been cloned and characterized. Development of resistant cultivars by introduction of major R genes into elite rice varieties has proven to be the most eco-friendly and sustainable approach for blast control [11]. Characterization of resistance genes from rice plant will help to unravel varied molecular mechanisms underlying the interaction between host and the pathogen. So, various molecular approaches are used to identify and to understand the mechanisms of activation/ expression of resistance genes in rice plants during infection. Resistance genes have been extensively studied in rice with the development of bioinformatics and molecular markers. The first evidence to support this concept was revealed by the direct interaction between the Pita protein and the Avr-Pita effector more than a decade ago [12]. To date, around 100 rice blast

resistance genes have been identified. Resistance genes are members of a very large multigene family and these R genes are distributed throughout the 12 rice chromosomes except chromosome 3 [13][14]. Out of them, 22 have been cloned [15].The identification and characterization of additional host resistance genes and pathogen avirulence genes is now required to deepen understanding of molecular machineries involved in the host-pathogen interaction and strategic deployment of resistance genes in commercial cultivars. Hence we planned to characterize the structure of rice blast resistance genes and planned to check their diversity. The accumulation of information regarding resistance gene mediated host defense mechanisms will facilitate engineering of genes conferring durable resistance to a broad spectrum of pathogens. Rice plants respond to pathogen attack by transcriptionally regulating different blast resistance genes through various types of transcription factors, which also show defense-responsive expression through specific types of cis elements in their promoter regions [16]. Promoters play major role in controlling gene expression. However, very little is known about blast resistance gene promoters. CpG islands are present at promoter regions and they may regulate the tissue-specific gene expression by undergoing modifications. CpG islands are discrete DNA regions in which CpG dinucleotide frequently occurs. Cis-acting regulatory elements (CAREs) associated with specific promoter regions are vital transcriptional gene regulatory units that establish distinct spatiotemporal transcriptional activity [17]. Hence identification and understanding of the cis-acting regulatory regions bound by TFs that control gene expression will put forward the crucial information to elaborate the mechanism of their expression in response to various signals. Various computational methods are employed in this study to analyze the promoter regions with respect to cis-acting regulatory elements and CpG islands of rice blast resistance genes.

2. MATERIALS AND METHODS

M. oryzae sub culturing

M. oryzae (MO36) culture was grown on oat meal agar medium (OMA) medium containing 40mg/l streptomycin sulphate. The culture plates were incubated at 26°C in dark for twelve days. Conidia of *M. oryzae* were collected from culture plates by rinsing with sterile water and10^5 conidia/ml counts were maintained using hemocytometer.

Rice seedling generation and plant disease screening

Seeds of rice cultivars BR2655 and HR12 were surface sterilized, water soaked and allowed to germinate in the Petri dish for five days with appropriate moisture content and then the germinated seedlings were transferred to plastic trays 60x30x30cm, filled with 10-11cm deep puddle sterile soil. Plants were raised in rows, 4-5cm apart with 20-25 plants per row and were grown for 20 days at 26±2°C. 0.1% Tween-20 suspension solution without conidia was used to spray the control plants (mock inoculation) at four-leaf stage. Approximately 10^5 conidia/ml of the pathogen was mixed with 0.1% Tween-20 and this suspension was sprayed on the 20 days old plants using glass

automizer. Relative humidity and the temperature were maintained above 90% and 25±1°C respectively under complete dark for 24h and later transferred to standard growth conditions of 16h of light and 8h of dark. The plants were observed regularly on day to day basis for the development of the disease symptoms. After seven days of inoculation, the disease incidence was assessed by recording the severity of blast by adopting 0 - 9 scale (IRRI, 1996). The rice leaves were collected from both the control and inoculation group and preserved immediately in the ice box and transferred to deep freezer (-20 ^{o}c).

Extraction of RNA and gene expression analysis

Total RNA from rice plant tissues was extracted using RNeasy Plant Mini Kit (Qiagen, Germany) by following manufacturer protocol. The quality was assessed by loading 1µl of total RNA on Agilent 6000 nanochip and to the Agilent 2100 Bio analyzer, USA, and quantification was done using QUBIT RNA HS kit (Thermo fisher Scientific, USA). NEBNext® Ultra™ RNA Library Prep Kit to prepare the mRNA libraries and Paired-end sequencing was performed with the TruSeq SBS Kit (Illumina, Inc. USA) on Illumina NextSeq 500 (Illumina., USA). Differential gene expression analysis was carried out and genes were shortlisted based on their fold change in expression. Rice blast resistance gene sequences of rice cultivars BR2655 and HR12 were retrieved from Rice Genome Annotation Project (RGAP) (www.rice.plantbiology.msu.edu/) [18] Chromosome maps of *O. sativa* blast resistance genes were constructed by Chromosome Map Tool available at Oryza base (www.viewer.shigen.info/oryzavw/maptool), The intron/exon organization of splice variants of resistance genes of *O. sativa* was retrieved from Rice Genome Annotation Project.

Retrieval of promoter regions and analysis of cis-regulatory elements

Promoter sequences (2 kb upstream of translation start site) of each blast resistance gene under consideration were retrieved from the RGAP. The tools PlantCare (www.bioinformatics.psb.ugent.be/webtools/plantcare)[19] and PLACE (www.dna.affrc.go.jp)[20] were used for identification of cis-regulatory elements in promoter of resistance gene sequences. PlantPAN (www.plantpan2.itps.ncku.edu.tw)[21] was used for the analysis of CpG islands. Similarly CpG plot tool was employed to check the CpG islands (www.ebi.ac.uk/Tools). The sequences were aligned for multiple sequence alignment to observe the sequence similarity among resistance gene upstream sequences using Clustalw. Sequence data was analyzed by Molecular Evolutionary Genetic Analysis (MEGA X) software version 7.0 and tree was constructed using Neighbour Joining method with default parameters [22][23][24].

3. RESULTS AND DISCUSSION

Plant Disease screening

Three-week old BR2655 rice plant seedlings inoculated with *M. oryzae* (M036) conidial suspension, showed less infection and scored 2 based on IRRI SES scale. HR12 rice cultivars

showed characteristic blast symptoms and it scored 8 on the scale.

Extraction of RNA and gene expression analysis

Total mRNA was isolated from leaves of BR2655 and HR12 rice cultivars, RNA libraries were prepared and they were sequenced under control conditions using Illumina NextSeq 500 (Illumina, USA). In total, we obtained 75.8 and 69.7 million raw reads for BR2655 134 and HR12 rice cultivars respectively. Reference based assembly were done using of CGWB tool. We identified 7577 differentially expressed genes (DEG) in the BR2655 and 4290 DEGs in HR12 cultivars. Based on fold change, five resistance genes were shortlisted for the promoter analysis from each of BR2655 and HR12 cultivars (Table 1). Blast resistance genes of BR2655 and HR12 rice cultivars were retrieved from RGAP databases given in Table 1. The first five genes are from BR2655 and the next five are from HR12 cultivar. The size of 10 rice blast resistance genes ranged from 2.4 kb (Os10g0161400bp) to around 3.1 kb (Os11g0677000) and this data was also verified in RAP-DB (www.rapdb.dna.affrc.go.jp). *In-silico* chromosome mapping of 10 resistance genes of two rice cultivars is presented in Fig. 1 and Fig. 2 shows the exon-intron and intron phase arrangement of these resistance genes. The resistance genes are distributed in 7 out of 12 chromosomes. Chromosome 11 harbored 3 resistance genes, whereas chromosome 3, 4, 5, 6, 8 showed no resistance genes. Promoter sequences up to 2 kb upstream from the translation start site of each resistance gene of rice were retrieved by RAP-DB (Table1.) and scanned using PlantCare program for the identification of cis-acting regulatory elements (CAREs). The study revealed a total of 10 CAREs in 10 resistance genes. The length of cis-acting regulatory elements varied from 5-11bp in selected rice cultivars. Cis-regulatory elements are listed into different functional categories as shown in Table 2. Os11t0590400 and Os12g0467300 gene promoters show two CAREs each, Os01t0953200, Os07t0294100 and Os11t0591000 gene promoters show three CAREs each, remaining resistance genes show single CARE each. Similarly promoters were scanned using PLACE and this revealed many cis-acting regulatory elements as listed in Table 3. 32 cis elements are recognized across ten blast resistance genes. Maximum of 17 cis elements are identified in Os09g0311600gene and least of four cis elements were recognized in Os10g0161400. Multiple sequence alignment done for these cis-acting regulatory elements showed similarity which is depicted in the phylogenetic tree as shown in Fig. 3. The tree was constructed by Neighbor Joining method. CpG island of length 801, 929, 1157, 1037 and 849 were detected in the promoter regions of Os01t0953200, Os11t0590400, Os12t0467300, Os02t0558400 and Os09t0311600 respectively. Similar CpG islands were also identified in the same set of promoter regions when CpG plot tool was used.

Table 1. Rice Blast Resistance genes of BR2655 and HR12 rice cultivars shortlisted for the analysis

Sr. No	RAP-GB ID	RGAP Locus ID	Length	Description
1	Os01g0953200	LOC_Os01g72390	2574	NBS type disease resistance protein, putative, expressed
2	Os11g0590400	LOC_Os11g37850	2763	Stripe rust resistance protein Yr10, putative, expressed
3	Os07g0294100	LOC_Os07g19320	2874	Stripe rust resistance protein Yr10, putative, expressed
4	Os11g0591000	LOC_Os11g37880	2913	Stripe rust resistance protein Yr10, putative, expressed
5	Os12g0467300	LOC_Os12g28100	2844	NBS-LRR disease resistance protein, putative, expressed
6	Os11g0677000	LOC_Os11g45180	3096	NBS-LRR disease resistance protein, putative, expressed
7	Os12g0215950	LOC_Os12g11370	3045	Verticillium wilt disease resistance protein, putative, expressed
8	Os10g0161400	LOC_Os10g07400	2415	Disease resistance RPP13-like protein 1, putative, expressed
9	Os02g0558400	LOC_Os02g35210	2844	Resistance protein, putative, expressed
10	Os09g0311600	LOC_Os09g14100	2976	Disease resistance protein RPS2, putative, expressed

Sl. no. 1-5: BR2655, 6-10: HR12 rice cultivars

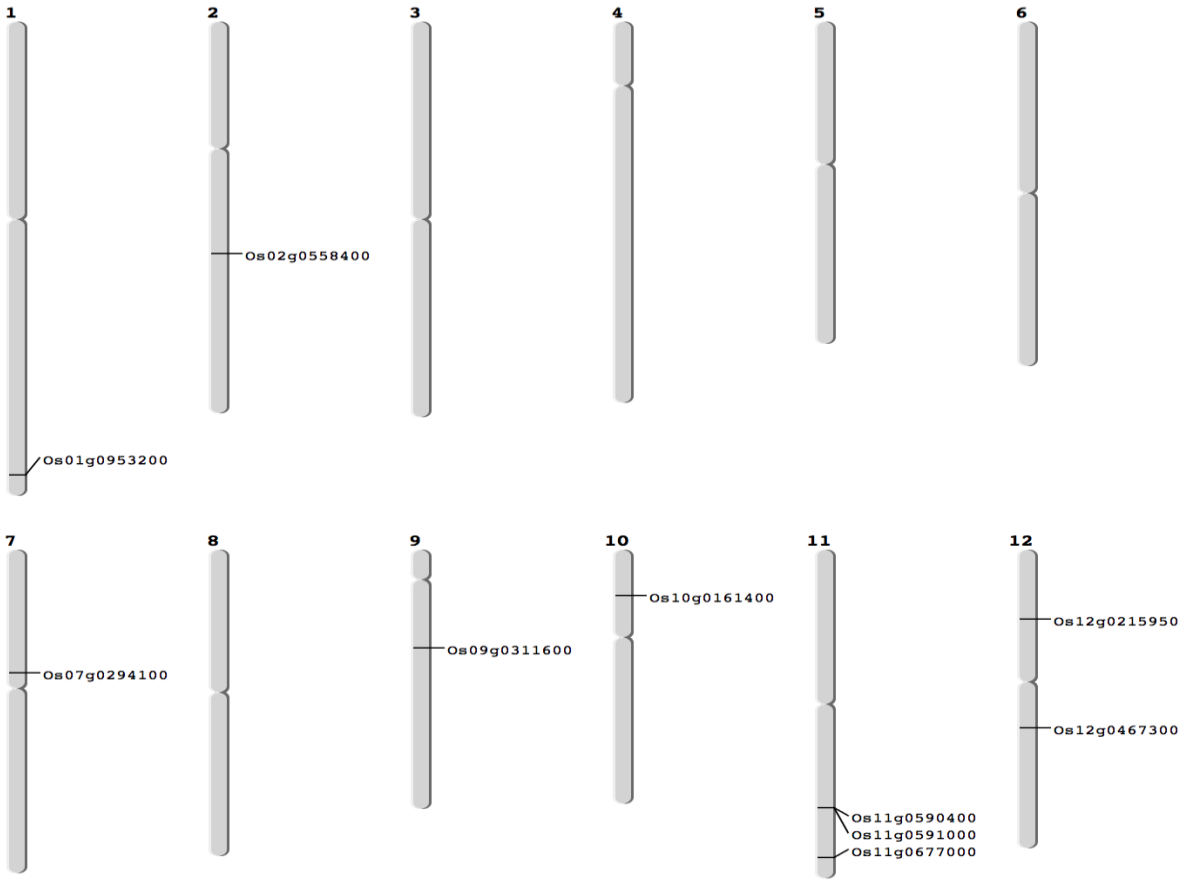

Figure 1. Distribution of resistance genes on rice chromosomes

(Os01g0953200, Os11g0590400, Os07g0294100, Os11g0591000, Os12g0467300 genes localized on BR2655 cultivar. Os11g0677000, Os12g0215950, Os10g0161400, Os02g0558400, Os09g0311 600 genes localized on HR12 cultivar.)

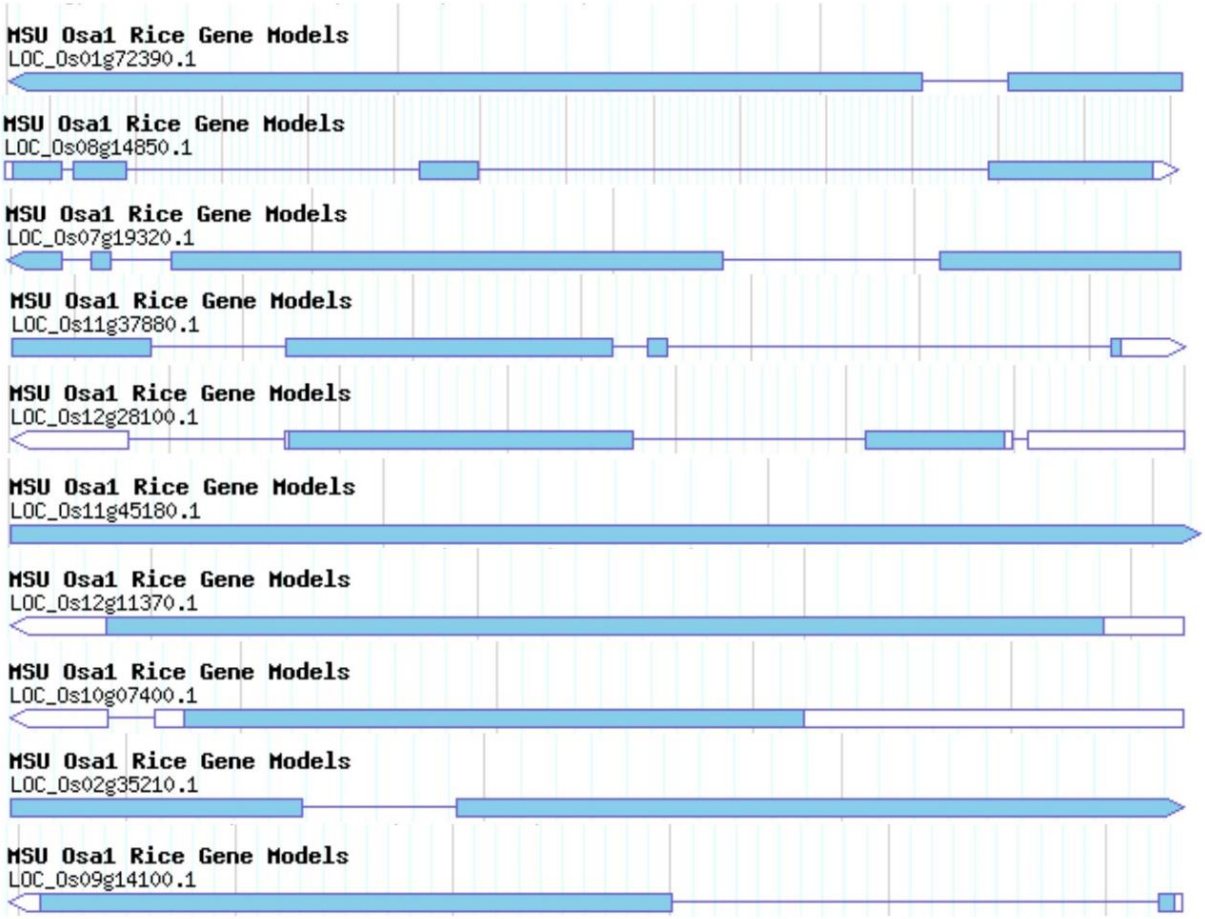

Figure 2. Exon-intron arrangement of of rice blast resistance genes. Exon-intron arrangement among splice forms in BR2655 [first five] and HR12 cultivars [last five]

Table 2: Cis-regulatory elements identified in Resistance gene promoters of BR2655 and HR12 rice cultivars through Plant CARE Database

GENE	ID OF SITE	SEQUENCE	DESCRIPTION
Os01g0953200	OS~Skn-1_motif	GTCAT	cis-regulatory element required for high levels of endosperm expression in cooperative interaction with other motifs (AACA; GCN4; ACGT) [25] (Pubmed ID:1680490)
	OS~GCN4_motif	TGAGTCA	motif highly conserved in glutelin promoters which may be involved in tissue-specific expression [26] (Pubmed ID:2263449)
	OS~GCN4_motif	CAAGCCA	truncated motif highly conserved in 5prime of all glutelin genes cis-regulatory element involved in endosperm expression with other

			motifs (AACA and A
Os11g0590400	OS~TATC-box	TATCCCA	cis-acting element involved in gibberellin-responsiveness but not involved in the expression of GluB-1 [25] (Pubmed ID:1680490)
	OS~motifIIb__1	CCGCCGCGCT	abscisic acid responsive element [27] (Pubmed ID: 14645724)
Os07g0294100	OS~Skn-1_motif	GTCAT	cis-regulatory element required for high levels of endosperm expression in cooperative interaction with other motifs (AACA; GCN4; ACGT) [25] (Pubmed ID:1680490)
	OS~TATA-box	TACAAAA	core promoter element around -30 of transcription start
	OS~TATC-box	TATCCCA	cis-acting element involved in gibberellin- responsiveness but not involved in the expression of GluB-1 [25] (Pubmed ID:1680490)
Os11g0591000	OS~P-box	GCCTTTTGAGT	cis-acting element involved in gibberellin responsiveness; conserved element in the upstream sequence of GA inducible genes in cereal seeds [26] (Pubmed ID:17319974)
	OS~Skn-1_motif	GTCAT	cis-regulatory element required for high levels of endosperm expression in cooperative interaction with other motifs (AACA; GCN4; ACGT) [25] (Pubmed ID:1680490)
	OS~TATA-box	TACAAAA	core promoter element around -30 of transcription start
Os12g0467300	OS~Skn-1_motif	GTCAT	cis-regulatory element required for high levels of endosperm expression in cooperative interaction with other motifs (AACA; GCN4; ACGT) [25] (Pubmed ID:1680490)

	OS~Sp1	GGGCGG	involved in light responsiveness [28] (Pubmed ID:1623185)
Os11g0677000	OS~GCN4_motif	TGAGTCA	motif highly conserved in glutelin promoters which may be involved in tissue-specific expression [26] (Pubmed ID:2263449)
Os12g0215950	OS~Box II -like sequence	TCCGTGTACCA	cis-acting regulatory element; similar DNA sequence and relative position to the motif present in alpha-amylase promoter [26](Pubmed ID:2263449)
Os10g0161400	OS~Skn-1_motif	GTCAT	cis-regulatory element required for high levels of endosperm expression in cooperative interaction with other motifs (AACA; GCN4; ACGT) [25] (Pubmed ID:1680490)
Os02g0558400	OS~Skn-1_motif	GTCAT	cis-regulatory element required for high levels of endosperm expression in cooperative interaction with other motifs (AACA; GCN4; ACGT) [25] (Pubmed ID:1680490)
Os09g0311600	OS~G-box	CACGTG	cis-acting regulatory element [29]

Table 3: Characteristics of cis-regulatory elements identified in resistance gene promoters of BR2655 and HR12 rice cultivars through PLACE Database

Gene	CIS ID	CIS Element name	Sequence	Description
Os01g0953200	S000353	AACACORE OSGLUB1	AACAAAC	Core of AACA motifs found in rice (O.s.);[30]
	S000232	AGCBOXNPGLB	AGCCGCC	"AGC box" repeated twice in a 61 bp enhancer element in tobacco (N.p.) class I beta-1,3-glucanase (GLB) gene; [31]
	S000020	AMYBOX1	TAACARA	"amylase box"; Conserved sequence found in 5'-upstream region of alpha-amylase gene of rice, wheat, barley; [32]
	S000477 (2)	ANAERO1 CONSENSUS	AAACAAA	One of 16 motifs found in silico in promoters of 13 anaerobic genes involved in the fermentative pathway (anaerobic set 1)[33]
	S000478	ANAERO2 CONSENSUS	AGCAGC	One of 16 motifs found in silico in promoters of 13 anaerobic genes involved in the fermentative pathway (anaerobic set 1)[33]
	S000436	BP5OSWX	CAACGTG	OsBP-5 (a MYC protein) binding site in Wx promoter; [34]
	S000277	GCN4OSGLUB1	TGAGTCA	"GCN4 motif" found in GluB-1 gene in rice (O.s.); Required for endosperm-specific expression; [35]
	S000181	MYBGAHV	TAACAAA	Central element of gibberellin (GA) response complex (GARC);[36]
	S000259	PYRIMIDINEBOX OSRAMY1A	CCTTTT	Pyrimidine box found in rice (O.s.) alpha-amylase (RAmy1A) gene; [37]
	S000474(3)	SITEIIATCYTC	TGGGCY	"Site II element" found in the promoter regions of cytochrome; [38]
	S000400	TATABOXOSPAL	TATTTAA	Binding site for OsTBP2, found in the promoter of rice pal; [39]
Os11g0590400	S000353	AACACORE OSGLUB1	AACAAAC	Core of AACA motifs found in rice (O.s.);[30]
	S000477	ANAERO1 CONSENSUS	AGCAGC	One of 16 motifs found in silico in promoters of 13 anaerobic genes involved in the fermentative pathway (anaerobic set 1)[33]
	S000478 (4)	ANAERO2 CONSENSUS	CAACGTG	OsBP-5 (a MYC protein) binding site in Wx promoter; [33]
	S000432	LEAFYATAG	CCAATGT	Target sequence of LEAFY in the intron of AGAMOUS gene;[40]
	S000474(2)	SITEIIATCYTC	TGGGCY	Site II element" found in the promoter regions of cytochrome [38]
Os07g0294100	S000477	ANAERO1 CONSENSUS	AGCAGC	One of 16 motifs found in silico in promoters of 13 anaerobic genes involved in the fermentative pathway (anaerobic set 1);[33]
	S000478	ANAERO2 CONSENSUS	CAACGTG	OsBP-5 (a MYC protein) binding site in Wx promoter;[33]
	S000259	PYRIMIDINEBOX OSRAMY1A	CCTTTT	Pyrimidine box found in rice (O.s.) alpha-amylase (RAmy1A) gene;[37]
	S000474(2)	SITEIIATCYTC	TGGGCY	"Site II element" found in the promoter regions of cytochrome'';[38]
	S000403	TATCCAOSAMY	TATCCA	"TATCCA" element found in alpha-amylase promoters of rice [41]

Os11g0591000	S000477	ANAERO1 CONSENSUS	AGCAGC	One of 16 motifs found in silico in promoters of 13 anaerobic genes involved in the fermentative pathway (anaerobic set 1);[33]
	S000478	ANAERO2 CONSENSUS	CAACGTG	OsBP-5 (a MYC protein) binding site in Wx promoter;[33]
	S000259	PYRIMIDINE BOXOSRAMY1A	CCTTTT	Pyrimidine box found in rice (O.s.) alpha-amylase (RAmy1A) gene;[37]
	S000474	SITEIIATCYTC	TGGGCY	"Site II element" found in the promoter regions of cytochrome'';[38]
	S000400	TATABOX OSPAL	TATTTAA	Binding site for OsTBP2, found in the promoter of rice pal;[39]
	S000403(3)	TATCCA OSAMY	TATCCA	"TATCCA" element found in alpha-amylase promoters of rice;[41]
	S000256	TATCCAYMOTIF OSRAMY3D	TATCCAY	"TATCCAY motif" found in rice (O.s.) RAmy3D alpha-amylase; [42]
	S000433	WUSATAg	TTAATGG	Target sequence of WUS in the intron of AGAMOUS gene [40]
Os12g0467300	S000478 (6)	ACGTOSGLUB1	CAACGTG	OsBP-5 (a MYC protein) binding site in Wx promoter; [33]
	S000259	PYRIMIDINE BOXOSRAMY1A	CCTTTT	Pyrimidine box found in rice (O.s.) alpha-amylase (RAmy1A) gene;[37]
	S000102 (2)	RYREPEATVFLEB4	CATGCATG	"RY repeat motif"; quantitative seed expression; Gene: Viciafaba;[43]
	S000474(7)	SITEIIATCYTC	TGGGCY	"Site II element" found in the promoter regions of cytochrome'';[38]
	S000403	TATCCAOSAMY	TATCCA	"TATCCA" element found in alpha-amylase promoters of rice;[41]
Os11g0677000	S000498 (6)	BIHD1OS	TGTCA	Binding site of OsBIHD1, a rice BELL homeodomain transcription factor;[44]
	S000353	AACACOREOSGLUB1	AACAAAC	Core of AACA motifs found in rice (O.s.);[30]
	S000012	ABREOSRAB21	ACGTSSSC	ABA responsive element (ABRE)";[45]
	S000232	AGCBOXNPGLB	AGCCGCC	"AGC box" repeated twice in a 61 bp enhancer element in tobacco (N.p.) class I beta-1,3-glucanase (GLB) gene;[31]
	S000021	AMYBOX2	TATCCAT	Conserved sequence found in 5'upstream region of alpha-amylase gene of rice, wheat, barley;[33]
	S000478	ANAERO2 CONSENSUS	CAACGTG	OsBP-5 (a MYC protein) binding site in Wx promoter;[33]
	S000421	CAREOSREP1	CAACTC	"CAREs (CAACTC regulatory elements)" found in the promoter region of a cystein proteinase (REP-1) gene in rice;[46]
	S000053	HEXMOTIF TAH3H4	ACGTCA	"hexamer motif" found in promoter of wheat (T.a.) histone genes H3 and H4;[47]
	S000432	LEAFYATAG	CCAATGT	Target sequence of LEAFY in the intron of AGAMOUS gene;[40]
	S000354	PROLAMINBOX OSGLUB1	TGCAAAG	"Prolamine box" found in the rice (O.s.) GluB-1 gene promoter;[30]
	S000102(2)	RYREPEATVFLEB4	CATGCATG	"RY repeat motif"; quantitative seed expression; Gene: Viciafaba;[43]
	S000256	TATCCAYMOTIF OSRAMY3D	TATCCAY	"TATCCAY motif" found in rice (O.s.) RAmy3D alpha-amylase;[42]

	S000353	AACACORE OSGLUB1	AACAAAC	Core of AACA motifs found in rice (O.s.);[30]
	S000020(3)	AMYBOX1	TAACARA	"amylase box"; Conserved sequence found in 5'-upstream region of alpha-amylase gene of rice, wheat, barley;[32]
	S000021	AMYBOX2	TATCCAT	Conserved sequence found in 5'upstream region of alpha-amylase gene of rice, wheat, barley;[33]
	S000477(3)	ANAERO1 CONSENSUS	AAACAAA	One of 16 motifs found in silico in promoters of 13 anaerobic genes involved in the fermentative pathway (anaerobic set 1)[33]
	S000270(2)	ARFAT	TGTCTC	ARF (auxin response factor) binding site found in the promoters of primary/early auxin response genes of Arabidopsis thaliana (A.t.); AuxRE;[48]
Os12g0215950	S000421	CAREOSREP1	CAACTC	"CAREs (CAACTC regulatory elements)" found in the promoter region of a cystein proteinase (REP-1) gene in rice;[46]
	S000419(2)	GARE1OSREP1	TAACAGA	"Gibberellin-responsive element (GARE)" found in the promoter region of a cystein proteinase (REP-1) gene in rice; [46]
	S000053	HEXMOTIF TAH3H4	ACGTCA	"hexamer motif" found in promoter of wheat (T.a.) histone genes H3 and H4;[30]
	S000181	MYBGAHV	TAACAAA	Central element of gibberellin (GA) response complex (GARC);[36]

Table 4: Characteristics of CpG islands on different rice blast resistance gene promoter region through PlantPAN

Gene ID	Start site	End site	Length	G+C frequency	CpG o/e ratio*
Os01g0953200	659	1473	801	0.48	0.6
Os11g0590400	422	1366	929	0.49	0.75
Os07g0294100	-	-	-	-	-
Os11g0591000	-	-	-	-	-
Os12g0467300	858	2033	1157	0.54	0.84
Os11g0677000	-	-	-	-	-
Os12g0215950	-	-	-	-	-
Os10g0161400	-	-	-	-	-
Os02g0558400	43	1096	1037	0.51	1.07
Os09g0311600	1	863	849	0.5	0.92

* Observed/Expected ratio of CpG dinucleotides

Table 5: Characteristics of CpG islands on different rice blast resistance gene promoter regions through CpG plot

Sl. no	GENE	LENGTH	STRAND	FROM	TO
1	Os01g0953200	245	+	1010	1254
		245	+	3010	3254
2	Os11g0590400	230	+	737	966
3	Os07g0294100	-		-	-
4	Os11g0591000	-		-	-
5	Os12g0467300	548	+	1075	1622
6	Os11g0677000	-		-	-
7	Os12g0215950	-		-	-
8	Os10g0161400	-		-	-
9	Os02g0558400	603	+	206	808
10	Os09g0311600	350	+	284	633

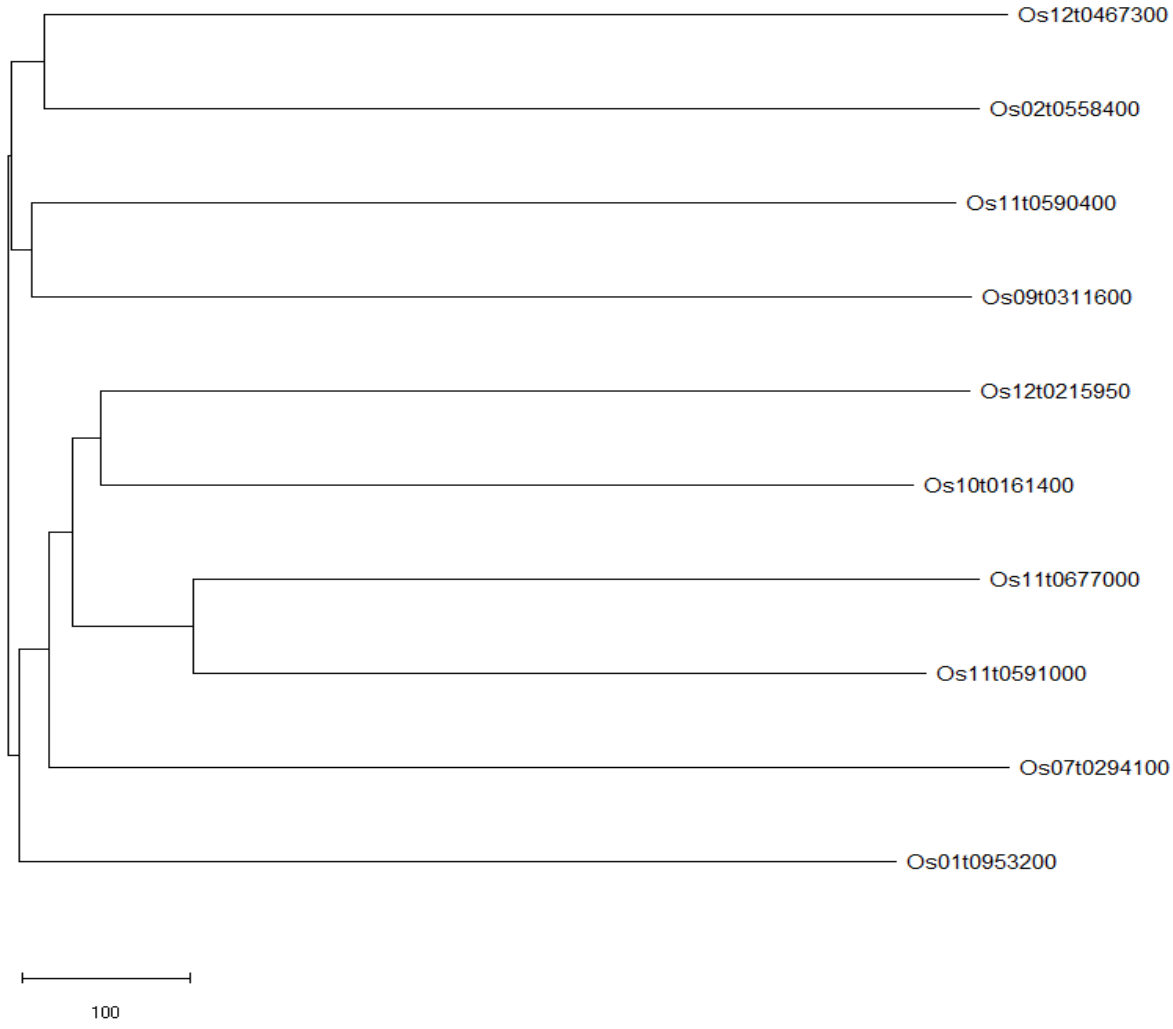

Figure 3: Phylogenetic analysis of rice blast resistance gene promoters of BR2655 and HR12 rice cultivars using MEGA X through Neighbor Joining method

Peer review under responsibility of Life Science Informatics Publications

Resistance genes (R) are members of a very large multigene family and these R genes are distributed throughout the 12 rice chromosomes except chromosome 3 [13][14]. Our study also revealed the absence of any one of the 10 resistance genes of BR2655 and HR12 in the 3rd chromosome in addition to the other chromosomes namely 4, 5, 6 and 8. The selected Resistance genes *viz.*, Os11g0590400, Os07g0294100 and Os11g0591000 show similarity based on their function of involvement in stripe rust resistance. Os01g0953200, Os12g0467300 and Os11g0677000 are specifically involved in encoding NBS-LRR disease resistance protein. The position of exons/introns and intronic phase distribution are important characteristics for gene structure analysis. The number of intron and exon arrangements were different in all the resistance genes. Os11g0677000 and Os12g0215950 genes do not show any introns. To analyse the general features of the promoter regions of the resistance genes, the 2 kb sequences upstream from the start site of the genes were used to search against known cis-regulatory elements in the PlantCARE database using the Search for Care program. The specific elements presented in the promoter regions of the coding strands were counted and are listed in Table 2. The data revealed that these promoter regions contained several cis-regulatory elements such as core promoter element with high level endosperm expression, glutelin promoters which may be involved in tissue-specific expression, Plant hormones responsive and light responsive elements having stress response and cellular functions. This type of investigation assists in better understanding of the functions of resistant genes from a holistic point of view [56]. Multiple sequence alignment and phylogenetic analysis of promoter sequences revealed that there is no conserved signature in the nucleotide sequences of cis-regulatory elements of selected promoter regions. All the 10 promoters are distinct to each other and specific conserved regions were not observed in selected upstream sequences. Five CpG islands were detected for the selected promoter regions in PlantPAN whereas six CpG islands were identified across five gene promotor sequences in CpGplot. This knowledge can be further used to understand the epigenetic changes during plant pathogen interaction [57].

4. CONCLUSION

Promoter analysis of rice blast resistance genes help us to understand the transcriptional regulation of resistance genes which are going to express transcripts during blast infection. The findings of the current research investigation will be helpful to understand the mechanism of expression of these resistance genes during plant defense mechanism. This knowledge will indirectly assist in resistance breeding.

ACKNOWLEDGEMENT

This work was supported by the Indian Council of Medical Research ICMR-SRF (IRIS ID NO. 2014-21680). The authors are grateful to Yuvaraja's College (Autonomous), University of Mysore, Mysuru, Central Instrumentation and Research Facility, Institution of Excellence scheme, Vijnana Bhavana, University of Mysore, Mysuru for providing lab facilities and green house facilities to

carry out this work.

CONFLICT OF INTEREST

The authors have no conflict of interest.

REFERENCES:

1. Muthayya S, Sugimoto JD, Montgomery S, Maberly GF. An overview of global rice production, supply, trade, and consumption. Ann N Y Acad Sci. 2014; 1324(1):7-14.
2. Khush GS. What it will take to feed 5.0 billion rice consumers in 2030. Plant Mol Biol. 2005; 59(1):1-6.
3. Samalova M, Meyer AJ, Gurr SJ, Fricker MD. Robust anti-oxidant defences in the rice blast fungus *Magnaporthe oryzae* confer tolerance to the host oxidative burst. New Phytol. 2014; 201(2):556-73.
4. Wilson RA, Talbot NJ. Under pressure: investigating the biology of plant infection by *Magnaporthe oryzae*. Nat Rev Microbiol. 2009; 7(3):185.
5. Zeigler RS. Recombination in *Magnaporthe grisea*. Annu Rev Phytopathol. 1998; 36(1):249-75.
6. Dean R, Van Kan JA, Pretorius ZA, Hammond-Kosack KE, Di Pietro A, Spanu PD, Rudd JJ, Dickman M, Kahmann R, Ellis J, Foster GD. The Top 10 fungal pathogens in molecular plant pathology. Mol Plant Pathol. 2012; 13(4):414-30.
7. Koeck M, Hardham AR, Dodds PN. The role of effectors of biotrophic and hemibiotrophic fungi in infection. Cell Microbiol. 2011; 13(12):1849-57.
8. Zeigler RS, Leong SA, Teng PS, editors. Rice blast disease. Int Rice Res Inst. 1994.
9. Dagdas YF, Yoshino K, Dagdas G, Ryder LS, Bielska E, Steinberg G, Talbot NJ. Septin-mediated plant cell invasion by the rice blast fungus, *Magnaporthe oryzae*. Science. 2012; 336(6088):1590-5.
10. Anonymous. Statistical database. www.fao.org.2016.
11. Skamnioti P, Gurr SJ. Against the grain: safeguarding rice from rice blast disease. Trends Biotechnol. 2009; 27(3):141-50.
12. Jia Y, McAdams SA, Bryan GT, Hershey HP, Valent B. Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. EMBO J. 2000; 19(15):4004-14.
13. Liu J, Wang X, Mitchell T, Hu Y, Liu X, Dai L, WANG GL. Recent progress and understanding of the molecular mechanisms of the rice-*Magnaporthe oryzae* interaction. Mol Plant Pathol. 2010; 11(3):419-27.
14. Yang JY, Shen CH, Zeng LX, Li YL, Zhen CH, Li CY, Zhu XY. Race specificity of major rice blast resistance genes to *Magnaporthe grisea* isolates collected from indica rice in Guangdong, China. Rice Science. 2008; 15(4):311-8.

15. Singh AK, Singh PK, Arya M, Singh NK, Singh US. Molecular screening of blast resistance genes in Rice using SSR markers. Plant Pathol J. 2015; 31(1):12.
16. Singh KB, Foley RC, Oñate-Sánchez L. Transcription factors in plant defense and stress responses. Curr Opin Plant Biol. 2002; 5(5):430-6.
17. Qiu P. Computational approaches for deciphering the transcriptional regulatory network by promoter analysis. Biosilico. 2003; 1(4):125-33.
18. Kawahara Y, de la Bastide M, Hamilton JP, Kanamori H, McCombie WR, Ouyang S, Schwartz DC, Tanaka T, Wu J, Zhou S, Childs KL. Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. Rice. 2013; 6(1):4.
19. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res.2002; 30(1):325-7.
20. Higo K, Ugawa Y, Iwamoto M, Korenaga T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res. 1999; 27(1):297-300.
21. Chow CN, Zheng HQ, Wu NY, Chien CH, Huang HD, Lee TY, Chiang-Hsieh YF, Hou PF, Yang TY, Chang WC. PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. Nucleic Acids Res. 2015; 44(D1):D1154-60.
22. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987; 4(4):406-25.
23. Nei M, Kumar S. Molecular evolution and phylogenetics. Oxford University Press; 2000.
24. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol. 2018; 35(6):1547-9.
25. Takaiwa F, Oono K, Wing D, Kato A. Sequence of three members and expression of a new major subfamily of glutelin genes from rice. Plant Mol Biol. 1991; 17(4):875-85.
26. Kim SY, Wu R. Multiple protein factors bind to a rice glutelin promoter region. Nucleic Acids Res. 1990; 18(23):6845-52.
27. Rabbani MA, Maruyama K, Abe H, Khan MA, Katsura K, Ito Y, Yoshiwara K, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. Plant Physiol. 2003; 133(4):1755-67.
28. Litts JC, Erdman MB, Huang N, Karrer EE, Noueiry A, Quatrano RS, Rodriguez RL. Nucleotide sequence of the rice (*Oryza sativa*) Em protein gene (Emp1). Plant Mol Biol. 1992; 19(2):335-7.

29. Sakamoto A, Okumura T, Kaminaka H, Sumi K, Tanaka K. Structure and differential response to abscisic acid of two promoters for the cytosolic copper/zinc-superoxide dismutase genes, SodCcl and SodCc2, in rice protoplasts. FEBS Lett. 1995; 358(1):62-6.
30. Wu CY, Washida H, Onodera Y, Harada K, Takaiwa F. Quantitative nature of the Prolamin-box, ACGT and AACA motifs in a rice glutelin gene promoter: minimal cis-element requirements for endosperm-specific gene expression. Plant J. 2000; 23(3):415-21.
31. Zhang H, Huang Z, Xie B, Chen Q, Tian X, Zhang X, Zhang H, Lu X, Huang D, Huang R. The ethylene-, jasmonate-, abscisic acid-and NaCl-responsive tomato transcription factor JERF1 modulates expression of GCC box-containing genes and salt tolerance in tobacco. Planta. 2004; 220(2):262-70.
32. Huang N, Sutliff TD, Litts JC, Rodriguez RL. Classification and characterization of the rice α-amylase multigene family. Plant Mol Biol. 1990; 14(5):655-68.
33. Mohanty B, Krishnan SP, Swarup S, Bajic VB. Detection and preliminary analysis of motifs in promoters of anaerobically induced genes of different plant species. Ann Bot. 2005; 96(4):669-81.
34. Zhu Y, Cai XL, Wang ZY, Hong MM. An interaction between a MYC protein and an EREBP protein is involved in transcriptional regulation of the rice Wx gene. J Biol Chem. 2003; 278(48):47803-11.
35. Onodera Y, Suzuki A, Wu CY, Washida H, Takaiwa F. A rice functional transcriptional activator, RISBZ1, responsible for endosperm-specific expression of storage protein genes through GCN4 motif. J Biol Chem. 2001; 276(17):14139-52.
36. Gubler F, Raventos D, Keys M, Watts R, Mundy J, Jacobsen JV. Target genes and regulatory domains of the GAMYB transcriptional activator in cereal aleurone. Plant J. 1999; 17(1):1-9.
37. Mena M, Cejudo FJ, Isabel-Lamoneda I, Carbonero P. A role for the DOF transcription factor BPBF in the regulation of gibberellin-responsive genes in barley aleurone. Plant Physiol. 2002; 130(1):111-9.
38. Welchen E, Gonzalez DH. Overrepresentation of elements recognized by TCP-domain transcription factors in the upstream regions of nuclear genes encoding components of the mitochondrial oxidative phosphorylation machinery. Plant Physiol. 2006; 141(2):540-5.
39. Zhu Q, Ordiz MI, Dabi T, Beachy RN, Lamb C. Rice TATA binding protein interacts functionally with transcription factor IIB and the RF2a bZIP transcriptional activator in an enhanced plant in vitro transcription system. Plant Cell. 2002; 14(4):795-803.
40. Kamiya N, Nagasaki H, Morikami A, Sato Y, Matsuoka M. Isolation and characterization of a rice WUSCHEL-type homeobox gene that is specifically expressed in the central cells of a quiescent center in the root apical meristem. Plant J. 2003; 35(4):429-41.

41. Chen PW, Chiang CM, Tseng TH, Yu SM. Interaction between rice MYBGA and the gibberellin response element controls tissue-specific sugar sensitivity of α-amylase genes. Plant Cell. 2006; 18(9):2326-40.
42. Rubio-Somoza I, Martinez M, Abraham Z, Diaz I, Carbonero P. Ternary complex formation between HvMYBS3 and other factors involved in transcriptional control in barley seeds. Plant J. 2006; 47(2):269-81.
43. Curaba J, Moritz T, Blervaque R, Parcy F, Raz V, Herzog M, Vachon G. AtGA3ox2, a key gene responsible for bioactive gibberellin biosynthesis, is regulated during embryogenesis by LEAFY COTYLEDON2 and FUSCA3 in Arabidopsis. Plant Physiol. 2004; 136(3):3660-9.
44. Luo H, Song F, Goodman RM, Zheng Z. Up-regulation of OsBIHD1, a rice gene encoding BELL homeodomain transcriptional factor, in disease resistance responses. Plant Biol. 2005; 7(5):459-68.
45. Busk PK, Pagès M. Regulation of abscisic acid-induced transcription. Plant Mol Biol. 1998; 37(3):425-35.
46. Sutoh K, Yamauchi D. Two cis-acting elements necessary and sufficient for gibberellin-upregulated proteinase expression in rice seeds. Plant J. 2003; 34(5):635-45.
47. Shimizu H, Sato K, Berberich T, Miyazaki A, Ozaki R, Imai R, Kusano T. LIP19, a basic region leucine zipper protein, is a Fos-like molecular switch in the cold signaling of rice plants. Plant Cell Physiol. 2005; 46(10):1623-34.
48. Nag R, Maity MK, DasGupta M. Dual DNA binding property of ABA insensitive 3 like factors targeted to promoters responsive to ABA and auxin. Plant Mol Biol. 2005; 59(5):821-38.
49. Vandepoele K, Vlieghe K, Florquin K, Hennig L, Beemster GT, Gruissem W, Van de Peer Y, Inzé D, De Veylder L. Genome-wide identification of potential plant E2F target genes. Plant Physiol. 2005; 139(1):316-28.
50. O'Neill SD, Kumagai MH, Majumdar A, Huang N, Sutliff TD, Rodriguez RL. The α-amylase genes in Oryza sativa: characterization of cDNA clones and mRNA expression during seed germination. Mol Gen Genet. 1990; 221(2):235-44.
51. Skinner JS, von Zitzewitz J, Szűcs P, Marquez-Cedillo L, Filichkin T, Amundsen K, Stockinger EJ, Thomashow MF, Chen TH, Hayes PM. Structural, functional, and phylogenetic characterization of a large CBF gene family in barley. Plant Mol Biol. 2005; 59(4):533-51.
52. Toyofuku K, Umemura TA, Yamaguchi J. Promoter elements required for sugar-repression of the RAmy3D gene for α-amylase in rice. FEBS lett. 1998; 428(3):275-80.
53. Washida H, Wu CY, Suzuki A, Yamanouchi U, Akihama T, Harada K, Takaiwa F. Identification of cis-regulatory elements required for endosperm expression of the rice storage protein glutelin gene GluB-1. Plant Mol Biol. 1999; 40(1):1-2.

54. Hwang YS, Karrer EE, Thomas BR, Chen L, Rodriguez RL. Three cis-elements required for rice α-amylase Amy3D expression during sugar starvation. Plant Mol Biol. 1998; 36(3):331-41.
55. Loke JC, Stahlberg EA, Strenski DG, Haas BJ, Wood PC, Li QQ. Compilation of mRNA polyadenylation signals in Arabidopsis revealed a new signal element and potential secondary structures. Plant Physiol. 2005; 138(3):1457-68.
56. Singh PK, Nag A, Arya P, Kapoor R, Singh A, Jaswal R, Sharma TR. Prospects of Understanding the Molecular Biology of Disease Resistance in Rice. Int J Mol Sci. 2018; 19(4):1141.
57. Kaur A, Pati PK, Pati AM, Nagpal AK. In-silico analysis of cis-acting regulatory elements of pathogenesis-related proteins of *Arabidopsis thaliana* and *Oryza sativa*. PloS one. 2017; 12(9):e0184523.