

# INTERNATIONAL JOURNAL OF SCIENCE AND NATURE

© 2004 - 2015 Society For Science and Nature(SFSN). All Rights Reserved

www.scienceandnature.org

# IDENTIFICATION AND CHARACTERIZATION OF NOVEL CC-NBS-LRR TYPE DISEASE RESISTANCE (R) GENE HOMOLOGUE SBRGA114 FROM SORGHUM [SORGHUM BICOLOR]

<sup>1</sup>Manju Kittur, <sup>1</sup>B. fakrudin, <sup>2\*</sup>Ayyanagouda Patil, <sup>3</sup>Surya Nandan Meena
<sup>1</sup>Institute of Agri-Biotechnology, University of Agricultural Sciences, Dharwad, Dharwad 580005 Karnataka, India
<sup>2</sup> Department of molecular biology and Agril. Biotechnology, MARS, University of Agricultural Sciences, Raichhur, Karnataka-584 104, INDIA.
<sup>3</sup>Department of Biotechnology, Goa University, Near Taliegao plateau, Panjim, Goa-403206, India
\*Corresponding authors email: ampatil123@gmail.com

#### ABSTRACT

In the present study, the isolation and characterization of full-length CC-NBS-LRR type disease resistance (*SbRGA114*) homologous gene from sorghum (*Sorghum bicolor* (L.) Moench). Based on the short sequence information of previously isolated sorghum resistance gene analogue, *SRGA114*, a 3,220 bp composite cDNA was amplified using 5' and 3' rapid amplification of cDNA ends (RACE) technique. Further, the entire coding region of 2,733 bp was cloned by nested RT-PCR using specific primers designed for 5' and 3' RACE products. It encodes a predicted polypeptide of 910 amino acids with a computed molecular weight of 104.1kDa. The deduced SbRGA114 protein consists of an N-terminal coiled-coil (CC) domain, nucleotide binding site (NBS), three perfect leucine rich repeats (LRR) and nine potential N-linked glycosylation sites. These domains are known to participate in protein-protein interaction and signal transduction during plant response to pathogens. Sequence analysis of *SbRGA114* showed high homology with reported NBS-LRR type resistance genes of maize, barley and rice.

KEYWORDS: CC-NBS-LRR, leucine rich repeats (LRR), RACE, SbRGA114, Sorghum bicolor

### **INTRODUCTION**

Plants have developed complicated defense mechanism during evolution to resist the pathogens that they encountered. Gene for gene concept of disease resistance involves two processes: perception of pathogen attack, followed by response to limit the disease. Perception for pathogen strains, encoded by disease resistance gene (R) involves receptors with a high degree of specificity. Data from the molecular analysis supports the model in which the products of R genes act receptor for the direct or indirect products of pathogen avirulence (Avr) gene<sup>[1]</sup>. Thus, plants can activate a very effective arsenal of inducible defense responses, comprised of the genetically programmed suicide of infected cells (hypersensitive response, HR)<sup>[2]</sup>, as well as tissue reinforcement<sup>[3]</sup>, and phytoalexins production at the site of infection<sup>[4]</sup>. However, programmed cell death in plants is especially prominent in response to abiotic and biotic stimuli<sup>[5]</sup>. These local responses can in turn trigger a long lasting systematic response (systemic acquired resistance, SAR), which highpoints the plant for resistance against a broad spectrum of pathogen<sup>[6]</sup>. Disease resistance genes have been cloned from several plant species and are mainly categorized into eight classes based on the structural similarities of their predicted protein products<sup>[7]</sup>. Most R genes contain a nucleotide binding site (NBS) attached to a C-terminal leucine rich repeat of flexible length. These domains contribute in protein protein interface and signal transduction<sup>[8,9]</sup> such genes are called NBS-LRR genes and

represent the most prevalent class<sup>[10]</sup>. NBS domain usually found in ATP or GTP binding proteins and is essential for catalytic activity of these proteins as it functions directly in ATP and GTP binding<sup>[11]</sup>. NBS-LRR resistance genes are further subdivided into TIR-NBS-LRR and CC-NBS-LRR groups. The TIR homologous domain has been detected in the N- terminal region, suggesting a function similar to Toll receptors that are original discovery of the Drosophila Toll protein and mammalian interleukin-1 receptor and their homologs are related to apoptosis of the cell<sup>[12]</sup>. Hence TIR homologous domains are involved in the resistance response in plants<sup>[13]</sup>, whereas, the CC-NBS-LRR genes contain a leucine zipper coiled-coil domain instead of a TIR domain and involved in<sup>[14,15]</sup>. To date the isolation of resistance genes has required the difficult and complex procedure of map based cloning and transposon tagging. Polymerase chain reaction (PCR) is one method that allows a more efficient means of isolating further resistance genes. PCR degenerate primers based on short stretches of amino acids conserved among NBS-LRR R genes are used to isolate resistance gene analogs (RGAs) and resistance gene like sequences from several plant genotypes (Soybean, Sorghum, Rice, Barley, maize, and Potato).

In the present study, we report the cloning of full-length NBS-LRR class resistance gene homologue using short sequence information from previously characterized RGAs of Sorghum<sup>[16]</sup> by performing Rapid Amplification of cDNA Ends PCR (RACE- PCR).

### MATERIALS & METHODS

#### Plant materials and RNA extraction

Seeds of sorghum [Sorghum bicolor (L.) Moench] M35-1 moderately foliar disease resistant cultivars were grown in the greenhouse. Three week seedlings were inoculated with spores of *Colletotrichum graminicola* (anthracnose causing pathogen) in order to induce plant–pathogen interaction and for salicylic acid (SA) treatment; seedlings were sprayed with 2.5mM SA. The three week old pathogen inoculated seedlings, leaves serve as the starting material for RNA isolation. The sample was powdered in liquid nitrogen with pestle and mortar and the total RNA was extracted using RNeasy <sup>®</sup> plant mini kit (QIAGEN sciences,Maryland, USA) conferring to the manufacturer's instructions. Further poly (A)<sup>+</sup> RNA was separated using a NucleoTrap®MRNAmini kit(Macherey-Nagel, Germany). **3' and 5' RACE-PCR** 

In order to generate unknown 3'fragment of the gene, cDNA synthesis was performed with the 3' RACE system for rapid amplification of cDNA ends using SMART RACE cDNA amplification kit (Clontech, Palo, Alto, CA, USA). Basically an aliquot of isolated poly (A)<sup>+</sup> RNA ( approximately 0.5µg ) was used as the template and the subsequently 3'RACE reaction was performed with a gene specific sense primer MR70F (5'-CTGGATGATGTCTG GGACCAGGAAG-3') designed specifically to elucidated RGA signature sequence as per the manufacturer's instruction. PCR was carried out in a total volume of 50µl, containing 2.5 µl 3'RACE ready cDNA, 10 picomol of MR70F and universal primer mix (UPM, supplied in the kit), 10µ mole dNTP, 1x PCR buffer and 1.5U Taq polymerase. Higher annealing temperature of primer allowed the use of Touch-down PCR program, which significantly reduced the nonspecific amplification, allowing the critical amount of gene product to accumulate. The Reverse transcription (RT-PCR) PCR reaction was performed under the following conditions: Initial five cycles of amplification (94°C -30s, 72°C -3 min) followed by next five cycles (94°C -30, 70°C -30s, 72°C -3 min) and then 25 cycles of amplification (94°C -30s, 68°C -30s, 72°C -3 min) with a final extension for 30 minutes at 72°C. PCR products were separated on 0.8 % agarose gel (Plate1), target DNA band corresponding to 2.0 kb was purified using Perfect prep® Gel Clean up (Eppendorf AG, Hamburg, Germany), ligated in to T/A cloning vector pTZ57R/T (Fermentes, EU) followed by transformation into E. coli strain DH5 . Competent cells were prepared using CaCl<sub>2</sub> double suspension method, while heat shock method was adopted for transformation in terms of the protocol of [40]. After confirmation of the clones through restriction and PCR analysis, positive clones were sequenced using M13 Forward and M13R primers. While to explore 5' end of the gene, 5'RACE-PCR was performed similar 3'RACE, except the 5'RACE ready cDNA and antisense primer MRP70R (5'-CTTGA GCATTTCCCAGAGGCTGGAT-3')designed specifically for above said RGA sequence were made worse. Compared to 3'RACE-PCR, a shorter extension of 2 min was given in the touch-down PCR program, as the

expected size of amplicon from 5'RACE was around 1-1.5 kb. Analysis of PCR product on 0.8 % agarose gel showed the presence of expected 1.4kb amplicon (Plate 2). Elution of the band, ligation, *E. coli* transformation and sequencing was done as mentioned above.

### Generation of *SbRGA114* full-length cDNA sequence

By comparing and aligning the overlapping sequences of 3' and 5' RACE products, the full-length cDNA sequence of the sorghum NBS-LRR type disease resistance gene homologue was deduced using VECTOR NTI 10.0. Further complete CDS was obtained through nested RT-PCR, using a new pair of primers SRGF1 (5'-GGCC TCGACCCCATTTGTGTGTAGT-3') and SRGR1 (5'-TGCG GAAGGAACATATCTGCCC-3') designed specifically to the UTRs of the deduced sequence (Plate 3). PCR was carried out in a total volume of 20ul reaction solution containing 2.5µl 5'RACE ready cDNA, SRGF1 and SRGR1 primers (10pmol), 10mM dNTP, 1x assay buffer 5A (plus Mg<sup>2+</sup>) and 1.5 U of XT-5 polymerase (Bangalore Genei, India) using the following PCR program: 94°C for 3 min, by 36 cycles of amplification (94°C for 40s, 55°C for 40s, 72°C for 4 min) and by 72°C for 45 min. Ligation of the amplified DNA fragment with pTZ57R/T vector, transformation of DH5 and sequencing were carried out as mentioned above.

### **Computer analysis**

Encoded amino acid sequence of *SbRGA114* was deduced using GENSCAN tool. The analysis and comparison of nucleotide and amino acid sequence with those released in NCBI database (www.ncbi.nlm.nih.gov) were done using the BLAST analysis algorithm. While, the structure of the SbRGA114 domain was analyzed with MEME (Multiple expectation maximization for motif elicitation). COILS program was made use to identify putative LZ structure/ domain in the cloned SbRGA114 protein. Cluster analysis was done using MEGA 4.1 software [41]. In the cluster analysis, along with *SbRGA114* protein, we involved RPM1, RPP8, RPS2, L6, RPP5, N and Xa21 sequences.

## RESULTS

# Cloning and sequence analysis of *SbRGA114* full-length cDNA

By performing 3' and 5' RACE-PCR, 3' cDNA end of 2006 bp (Figure 1) and 5' cDNA end of 1374 bp (Figure 2) were obtained respectively. By aligning and assembling these products having overlapping region of 160 bp, complete cDNA designated as SbRGA114 was deduced, obtained by nested RT-PCR (Figure 3) and further verified by sequencing. The full-length deduced SbRGA114 cDNA was 3220 bp (include 28 bp poly (A)<sup>+</sup> tail ) containing a 2733 bp open reading frame (ORF) with 5' un-translated region (UTR) of 371 bp upstream of the start codon and 88 bp 3' UTR downstream of the stop codon. The 5' UTR possessed a moderate G+C content (53.6 %), while 3' UTR region had lower G+C content (43.2 %). Sequence comparison in NCBI database revealed that SbRGA114 had high homology with R-genes reported in other plant species.



**FIGURE 1**: PCR amplified band (2006bp) of 3' cDNA end of *SbRGA114* gene with background amplification due to non- specific priming of UPM primer. M, indicate 1 kb ladder.



**FIGURE 2:** PCR amplified band (1374bp) of 5' cDNA end of *SbRGA114* gene with background amplification due to non- specific priming of UPM primer. M, indicate 100bp ladder.



FIGURE 3: Amplicon of 2808 bp obtained by Nested RT-PCR. Lane 1-5 represents generation of full-length cDNA of *SbRGA114*, which include complete ORF, 48bp, 25bp of 5' and 3' UTRs respectively.

#### **Characterization of predicted SbRGA114 protein**

In silico translation of SbRGA114 starting from the ATG codon yielded a protein of 910 amino acids, with a calculated molecular weight of 104.1kDa. The SbRGA114 protein harbored several regions of similarity to the polypeptide encoded by the known resistance genes. Similarity was mainly in the regions of functional protein domains, which included NBS and perfect LRR motifs. COILS program<sup>[17]</sup> could detect putative coiled-coil (leucine zipper) domain at the N-terminal sequence of SbRGA114. Eight major NBS motifs viz., P-loop, RNBS-A, Kinase2a, RNBS-B, RNBS-C, GLPL, RNBS-D and MHDV, starting from the N-terminus of NBS domain were identified using MEME software (Figure 4). The exact start of NBS domain is not well defined in most of the characterized TNL and CNL proteins; however the NBS- domain always preceded by a pre P-loop conserved sequences. within the NBS domain, kinase-1a (P-loop)

motif has been identified with conserved sequences of GMGGLGKTT (201-209), RNBS-A motif (218-248) and sequence KCLIVLDDVW consensus (277-286)corresponding to kinase-2a motif. Kinase 2a motif contained four consecutive hydrophobic amino acids (LIVL) followed by a conserved aspartate (D) and tryptophan (W) at the end site. RNBS-B motif with consensus SRIIITTRK (304-312) and RNBS-C motif was found with an arginine (R) residue (325-345). Consensus GLPLA (amino acid 369-373), corresponding to Domain2 (hydrophobic domain) and **RNBS-D** domain, CFLYCSLFP (426-434) were also identified. Total 8 major conserved motifs called MHDV, which are highly conserved in maximum CC-NBS-LRR genes, with minor modification in TIR-NBS-LRR protein, was identified with the consensus CTMHDIVRDLALSVA at 493-507 positions.

#### CC-NBS-LRR type disease resistance (R) gene homologue SBRGA114 from sorghum

	10	20	30	40	50	60	70	80	90	
Carabur				CONTRACTOR NO.		MANUTERICS		CHICENERIAN	LITE PATH	
Sorgnum PPM1	MORATUPS (IGRILSULENET									
RPP8	MAEAFVSFGLEKLWOLLSRESERLOGVDEQIDGLKROLRSLOSLLKDADAKKHGSDRVRNFLEDVKDLVFDAEDII									
RPS2	MDFISSLIVGCAQVLCESMNMAERRGHKTDLRQAITDLETAIGDLKAVRDDLTLRIQQDGLEGRSCSNRAREWLSAVQVTETKTALLL									
	:			:	. :			** .*		
	100	110	120	130	140	150	160	170	180	
Sorghum	DRYTYHTVQMEEEW	FLKKYFIKASH	YVLDFTQI	AEEVIKIEKE	REVIELKELW	FQPSQLVAD	LIEMERQ	RSHDNFPL	LIKDED	
RPMI	DEFORMOR ROUGHLINKAFHFF-KIMAKHSIAQALQKYNYN 135ISDSMKKTIHSENTQAALLPFIDDGDAKWYNNISESSLFFSENS ESYVLIKIERGEKG-VKKYURELA FILTDEHKVASDIE GOTKRISOV I GEMOSPETIOOI I DGVEELSIAFDA-DUADET BOTV DESEFED									
RPS2	VRFRREQRTRMRRRCLSCFGCADYKLCKKVSAILKSIGELRERSEAIKTDGGSIOVTCREIPIKS									
	I I .I I .I I									
	190	200	210	220	230	240	250	260	270	
	pre P-loop		P-loo	P	+ RN	BS-A				
Sorghum	LVGIEDNRRMLMGWLYSDELDSTVITVSGMGGLGKTTLVTNVYEREKINFSATAWMVVSQTYTIEALLRKLLMKVGGEQQVPPNI									
RPM1	LVGIDAPKGKLIGRLLSPEPQRIVVAVVGMGGSGKTTLSANIPKSQSVRH-FESYAWVTISKSVVIEDVFRTMIKEFYKEADTQIPAEL									
RFF8	LVGVEQSVKELVGHLVENUVH-QVVSIAGMGGIGKTTLARQVFHHDLVFKH-FUGFARVCVSQQFTQKHVWQRLQELQPHDGDI									
RF32	THE REPORT OF A CONTRACT OF A									
	280	290	300	310	320	330	340	350	360	
		Kina	se-2a		RNBS-	B		RNBS-C		
Sorghum	DKLDVYDLKEKLKQKLKTRKCLIVLDDVWDQEVYLQMSDAFQN-LQSSRIIITTRKNHVAALAHPTRRLDIQPLGNAQAFDLFCRRTF									
RPM1	YSLGYRELVEKLVEYLQSKRYIVVLDDVWTTGLWREISIALPDGIYGSRVMMTTRDMNVASFPYGIGSTKHEIELLKEDEAWVLFSNKAF									
RPP8	LONDEYALQRKLFQLLEAGKYLVVLDDVWKKEDWDVIKAVFPR-KRGWKMLLTSRNEGVGIHADP-TCLTFRASILNPESWKLCERIVF EXETOPNALYVAL FOXFEILI I DOWNEYE DI FEYTGUDEDEDENKVLMFTBSIL ONNG-LEFTTSTEDE EDVELOPTEDEVIC									
RF52	EREIGENPOLETIIN	ALROPART DEDE		LEKIGVPRPDP	CENECE VALUE	* ·	ALIALAV	EF LEANNALL	FCSRVN	
	370	380	390	400	410	420	430	440	450	
			GLPL					+		
Sorghum	YNEKDHACPSDLVEVATSIVDRCQGLPLAIVSIASLLSSRAQTYYIWNQIYKRLRSELSNNDHVRAVLNLSYHDLSGDLR									
RPM1	PASLEOCRTONLEPIARKLVERCOGLPLAIASLGSMMSTKKFES-EWKKVYSTLNWELNNNHELKIVRSIMFLSFNDLPYPLK									
RPP8	PRRDETEVRLDEEMEAMGKEMVTHCGGLPLAVKVLGGLLANKHTVP-EMKRVSDNIGSQIVGGSCLDDNSLNSVYRILSLSYEDLPTHLK									
RPSZ	RKDLLESSSI	RELATIVSKC	GGLPLALI	FLOGAMAHRET	TEE-EN THASE	VLTRF PAEMI	GMNY	VFALLKESIDN	LESDLL	
	460	470	480	490	500	510	520	530	540	
	RNBS-D			+			M	HDV		
Sorghum	N-CFLYCSLFPEDYPIPRESLVRLWVAEGFALSKENNTAEEVAEGNLMELIHRNMLVVMENDEQGRVSTCTMHDIVRDLALSVAKEE									
RPM1	R-CFLYCSLFPVNYRMKRKRLIRMMMAORFVEPIRGVKAEEVADSYLNELVYRNMLOVILWNPFGRPKAFKMHDVIWEIALSVSKLE									
RPP8	H-RFLFLAHFPEYSKISAYDLFNYWAVEGIYDGSTIQDSGEYYLEELVRRNLVIADNRYLSSHSKNCQMHDMMREVCLSKAKEE									
RFSZ	NOUT DIG ALT PERFECT P									
	550	560	570	580	590	600	610	620	630	
								LRR1		
Sorghum	-RFGTANNYRAMIL	MDRDRDVRRLS	SYCWKDSTS	SVVVRLPRLR1	LVSLGTI	SSSPNMLLS	LSESSYLTV	LELQDSEIT	EVPASI	
RPM1	-RFCDVYNDDSDGDDAAETMENYGSRHLCIQKEMTPDSIRATNLHSL-LVCSSAKHKMELLPSLNLLRALDLEDSSISKLPDCL									
RPP8	-NFLQIIKOPTSTSTINAQSPSRSRRLSIHSGKAFHLLGHKNNTKVRSLIV-WDEDFGIRSASVFHNLTLLRVLDLYWVKFEGGKLPSSI									
RP52	KELILVEPSMGHTE	APRAENWRUAL	VISLLUNK	UTLPERLICI	KLITLADQUA	SILKKIPIG	PMMPVLRV	LULSFISII	LIFLOI	
	640	650	660	670	680	690	700	710	720	
	LRI	22		LRR3	222.52					
Sorghum	GNLFNLRYIGLRRTKVRSLPDSVEKLLNLQTLDIKQTKIEKLPRGISKVKKLRHLVADRYADEKQSQFRYFIGMQAPKDLSNLVELQT									
RPM1	VTMFNLKYLNLSKT	OVKELPKNFHR	LVNLETLN	TKHSKIEELPI	LGMWKLKKL	RYLITFRRNI	GHDSNWNYV	LGTRVVPKIWO	LKDLQV	
RPP8	GGLIHLRYLSLFLAGVSHLPSTMRNLKLLLYLNLSVNNKEPIHVPNVLKEMIQLRYLSLPLKKDDKTKLELGDLVNLEF									
RP52	KILVELINLSMSGI	ALSVLPUELGA	LERKLENILDI	DORIGEDUTII	RUAICWLSKL	EVENLII51/	GWG-LQSFQ	EDEVEELGEAU	LEILEN	
	730	740	750	760	770	780	790	800	810	
			+							
Sorghum	LETVEASKDLGEQL	KKLMQL	RTLWID	NISAADCANI	FASLSNMPLLS	NLLLSAKDEN	EPLCFEALK	PRSTGLHRL	IIRGQW	
RPM1	MUCHAREDELIKNLGCMTQLTRISLVMVRREHGROLCDSLNKIKRIRFLSLTSIDEEEPLEIDDLIATASIEKLFLAGKL									
RPP8 RPS2	LEVED 1 VID 3 V 1 V LINE 1 V L KI LAVSLS BEKONE ET LSSSLKE LEKILET LI VLES VE 1 RV UMGEV LDHF I HLKELGLAV									
REDE		I II	i i	i i i i i i i i i i i i i i i i i i i	- Sur an GRALP	* : :	is the virable	ENDRUF DUEV L	insen	
	820	830	840	850	860	870	880	890	900	
				+						
Sorghum	AKG-TLQCPLFRGH	SRHLKYLALSW	CHLSEDPLE	EMLAPQLPNLT	INLRLNNMRSA	STLVLPPGS-	FP	HLKLLVLMHMP	NVKKLV	
RPM1	ERVPSWFN-TLQNLTYLGLRGSQLQEN-AILSIQTLPRLWRLSFYNAYMGPRLRFAQGFQNLKILEIVQMKHLTEVV									
RPS2	RMSKIPDURQLPFRLAQIYICNCKMEEDPMPILEKLLHLKSVKLTFKAFAGRMWVCSKGGFTQLCALEISEQSELEEWI NITRUWCNSVSOCIENIBCINISCINISCINISCINISCINISCINISCINISCI									
				•		:				
	910	920	930	940	950	960	970	980		
					*					
Sorghum	IGKGALRCIEGLYI	SLAELDKVPC	GMESLRTL	KKLSLVSLHR	SFLTEWNKSGM	HDKMQDVI	DIRV	ACCEVUEL CO		
RPPS	LEDGARFE EUGNET VRACHGER IV FRGIERLINUGE INLINUGE INLINUS KUVERINGEGSVURSANI PAIRAT FRTUNGSFIVSLSS									
RPS2	PSRFSFQKVETLVITNCPRVKKLPFQERRTOMLPTVYCEEKWWKALEKDOPNEELCYLPRFVPN									

**FIGURE 4:** Clustal W multiple sequence analysis of deduced SbRGA114 amino acid sequence with CC-NBS-LRR type Arabidopsis RPM1 (AF122982.1), RPP8 (AAC83165.1) and RPS2 (AAM90884.1). Arrows indicate the potential N-linked glycosylation sites. Major domains identified by MEME are underlined.

In the region of LRR, three perfect repeats, each of 23 amino acid long leucine rich repeat motifs were identified (577-645). The conserved aliphatic residue of the -turn-sheet domain (xxLxLxx) was located within the imperfect recurrences (Figure 5). There are 7 possible N-linked glycosylation locations NxS/T in the SbRGA114 protein (amino acid 18, 221, 413, 459, 709, 804 and 894) which may interact with other proteins in defense reaction

signal-transduction pathway. Unlike additional motifs found in several *Arabidopsis* NBS-LRR proteins, no additional domains were identified after a LRR region in the C-terminus of *SbRGA114* protein. Cluster analysis was done using MEGA 4.1 software clearly revealed the grouping of *SbRGA114* with non-TIR type resistance protein (Figure 6).

#### YxxYxYxxxYxxYPxxYxxYxx

577 LTVLELQDSEITEVPASIGNLFN 599600 LRYIGLRRTKVRSLPDSVEKLLN 622623 LQTLDIKQTKIEKLPRGISKVKK 645

**FIGURE 5:** Cluster analysis of *SbRGA114* with known resistance protein using MEGA 4.1 by neighbor joining method. Sequences cited are Arabidopsis RPP8 (AAC83165.1), RPM1 (AF122982.1), RPS2 (AAM90884.1) and RPP5 (AAF08790.1), tobacco N (Q40392), flax L6 (AAA91022.1) and rice (BAD73428.1). Units at branching indicate bootstrap values.



**FIGURE 6:** The consensus in three perfect LRRs of *SbRGA114* is indicated in the line above the alignments of LRRs. Y-represents hydrophobic amino acid (L or I or V) and X- represents any amino acid residue.

## DISCUSSION

Increasing productivity and production of food crops is the need of the hour due to increasing population and decrease in agricultural land. The insect pest and disease continues to cause heavy losses. So far, the best way to tackle any disease is to breed resistant cultivar. Recent developments in molecular biology and biotechnology have provided a number of tools to study the interaction between plants and pathogens. In the present study, we have isolated and characterized a full-length disease resistance gene homologue (SbRGA114) from sorghum encoding a CC-NBS-LRR type protein. Similarity searches in NCBI database revealed that, of the previously characterized 13 analogue sequences, SRGA114 shared extensive homology with NBS-LRR type R genes reported in other cereals, indicating its uniqueness. In several mapping studies, RGAs have been demonstrated to be linked with the recognized R-genes, counting Lr10 gene of wheat that was isolated by RGA screening<sup>[18]</sup>. Hence, it thought appropriate to clone full-length version gene of which SRGA114 was a part.

The SbRGA114 protein shared 29 % identified with the Arabidopsis RPM1 protein encoding bacterial blight resistance<sup>[19]</sup>, While, 24 and 20% identity with Arabidopsis RPP8 and RPS2 proteins <sup>[20,21]</sup>, respectively. It had 17 % identified with the tobacco 'N' protein conferring tobacco mosaic virus (TMV) resistance<sup>[22]</sup>, 14% identity with RPP5 protein of Arabidopsis conferring resistance to fungal pathogen Perunospora parasitica<sup>[23]</sup>, and 13% identity with flax rust resistance protein L6<sup>[24]</sup>. The domains present in SbRGA114 distinguished it from other classes of R-genes eg., tomato Pto<sup>[25]</sup>, tomato Cf-2, Cf-4, Cf-5, Cf-9<sup>[26-29]</sup> and the rice  $Xa-21^{[30]}$  which have a serine/threonine kinase or an extra-cytoplasmic LRR domain or both. The domain of SbRGA114 i.e. NBS reported to show the amino acid motifs conserved among the known resistance proteins<sup>[31]</sup>. Suggesting their possible role in activation of kinases or as G-proteins in signal transduction<sup>[32]</sup>. LRR domain is known to be involved in ligand binding and pathogen recognition, through proteinprotein interaction<sup>[33]</sup>. While, the amino-terminal leucine zipper domain facilitates the formation of coiled-coil structure to promote either dimerization or specific interaction with other proteins. The presence of pre loop consensus on the NBS domain is found similar to Arabidopsis R genes, which was identified in SbRGA114

gene with the consensus DEDLVGIE (amino acid 170-175)<sup>[34]</sup>. The kinase-1a (P-loop) motif involved in interaction with phosphates and Mg2+ ions <sup>[35]</sup>. Kinase-2a motif, known to function in phosphotransfer reaction <sup>[36]</sup>. RNBS-B motif & RNBS-C motif involved in purine or ribose binding<sup>[37]</sup>. It has been observed from the analysis of resistant gene sequences from rice, barley and maize that monocots lack a family of NBS-LRR genes with a TIR motif, instead a non-TIR sub groups of NBS-LRR genes are common<sup>[38]</sup>. Sorghum one of the major cereal crop and member of the grass family, has been considered as non-TIR monocot. However, only few reports on sorghum sequences were used to pull this inference<sup>[39]</sup>. It was substantiated by the analysis of cloned sorghum *SbRGA114* gene.

#### ACKNOWLEDGEMENT

Authors are pleased to acknowledge the Department of Biotechnology, New Delhi for financial support to present study.

#### REFERENCES

- [1]. Cesari, S., Thilliez, G., Ribot, C., Chalvon, V., Michel, C., Jauneau, A. and Kroj, T. (2013) The rice resistance protein pair RGA4/RGA5 recognizes the *Magnaporthe oryzae* effectors AVR-Pia and AVR1-CO39 by direct binding. Plant Cell Online 25, 1463-1481.
- [2]. Greenberg, J.T. (1996) Programmed cell death: a way of life for plants. Proc. Natl. Acad. Sci. 93, 12094-12097.
- [3]. Van-Kan, J.A.L. (2006) Licensed to kill: the lifestyle of necrotrophic plant pathogen. Trends Plant Sci. 11, 247-253.
- [4]. Ahuja, I., Kissen, R. and Bones, A.M. (2012) Phytoalexins in defense against pathogens.*Trends*. Plant Sci.17, 73-90.
- [5]. Williams, B. & Dickman, M. (2008) Plant programmed cell death: can't live with it; can't live without it. *Mol. Plant Pathol.* 9, 531–544.

- [6]. Janus, L., Milczarek, G., Arasimowicz-Jelonek, M., Abramowski, D., Billert, H. & Floryszak, Wieczorek, J. (2013) Normoergic NO-dependent changes, triggered by a SAR inducer in potato, create more potent defense responses to *phytophthora infestans*. Plant Sci. 211, 23-34.
- [7]. Gururani, M.A., Venkatesh, J., Upadhyaya, C.P., Nookaraju, A., Pandey, S.K. and Park, S.W. (2012) Plant disease resistance genes: Current status and future directions. Physiol. Mol. Plant Pathol. 78, 51-65.
- [8]. Dong, J.X., Dong, H.T. and Li, D.B. (2001) Recent advances in plant disease resistance genes. Acta. Phytopathol. Sinica. 31, 1–9.
- [9]. Xu, Q., Wen, X.P. and Deng, X.X. (2005) Isolation of TIR NBS-LRR resistance gene analogues and identification of molecular marker linked to a powdery mildew resistance locus in chestnut rose. Theor. Appl. Genet.111, 819–830.
- [10]. Hulbert, S.H., Webb, C.A., Smith, S.M. and Sun, Q. (2001) Resistance Gene Complexes: Evol. Utilization Annu. Rev. Phytopathol.39, 285-312.
- [11]. Tameling, W.L., Elzinga, S.D.J., Darmin, P.S., Vossen, J.H., Takken, F.L.W., Harling, M.A. and Cornelissen, B. (2002) The tomato R gene products *I*-2 and *Mi-1* are functional ATP binding proteins with ATPase activity. Plant Cell 14, 2929-2939.
- [12]. Lemmers, B., Salmena, L., Bidère, N., Su, H., Matysiak-Zablocki, E., Murakami, K. and Hakem, A. (2007) Essential role for caspase-8 in Toll-like receptors and NF B signaling. J. Biol. Chem. 282, 7416-7423.
- [13]. Bernoux, M., Williams, V.T.S., Warren, C., Hatters, D., Valkov, E. and Dodds, P.N. (2011) Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signaling, and auto-regulation. Cell Host Microbe. 9, 200-211.
- [14]. Pan, Q., Wendel, J., Fluhr, R. (2000) Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. J. Mol. Evol. 50, 203-213.
- [15]. Cannon, S.B., Zhu, H., Baumgarten, A.M., Spangler, R., May, G., Cook, D.R. and Young, N.D. (2002) Diversity, distribution and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies. J. Mol. Evol. 54, 548-562.
- [16]. Totad, A.S., Fakruddin, B. and Kuruvinashetty, M.S. (2005) Isolation and characterization of resistance gene analogs (RGAs) from sorghum (*Sorghum bicolor* L. Moench). Euphytica 143, 179-188.

- [17]. Lupas, A., Van-Dyke, M. and Stock, J. (1991) Predicting coiled-coils from protein sequences. Sci. 252, 1162-1164.
- [18]. Feuillet, C., Travella, S., Stein, N., Albar, L., Nublat, A. and Keller, B. (2003) Map-based isolation of the leaf rust disease resistance gene Lr10 from the hexaploid wheat (*Triticum aestivum* L.) genome. Proc. Natl. Acad. Sci. 100, 15253-15258.
- [19]. Grant, M.R., Godiard, L., Straube, E., Ashfield, T., Lewald, J., Sattler, A., Innes, R., Dangl, J.L. (1995) Structure of the *Arabidopsis RPM1* gene enabling dual specificity disease resistance. Sci. 269, 873-846.
- [20]. [20] McDowell, J.M., Dhandaydham, M., Long, T.A., Aarts, M.G., Goff, S., Holub, E.B., Dangl, J.L. (1998) Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the *RPP8* locus of *Arabidopsis*. Plant Cell, 10, 1861-1874.
- [21]. Martin, G.B., Brommon-Schenkel, S.H., Chunwongse, J., Frary, A., Ganal, M.W., Spivery, R., Wu, T., Earle, E.D. and Tanksley, S.D. (1993a) Map based cloning of a protein kinase gene conferring disease resistance in tomato. Sci. 262, 1432-1436.
- [22]. Whitham, S., Dinesh, S.P., Choi, D., Hichi, R., Corr, C. and Baker, B. (1994) The product of the tobacco mosaic virus resistance genes *N*: similarity to Toll and the interleukin-1 receptor. Cell 78, 1011-1015.
- [23]. Parker, J.E., Coleman, M.J., Szabo, V., Frost, L.N., Schmidt, R., Vander-Biezen, E.A., Moores, T., Dean, C., Daniels, M.J. and Jones, J.D.G. (1997) The *Arabidopsis* downy mildew resistance gene *RPP5* shares similarity to the Toll and interleukin receptor with N and L6. Plant Cell 9, 879-894.
- [24]. Lawrence, G.J., Finnegan, E.J., Agliffe, M.A. and Ellis, J.G. (1995) The *L6* gene for flax rust resistance is related to the *Arabidopsis* bacterial resistance gene *RPS2* and the tobacco viral resistance gene *N*. Plant Cell 7, 1195-1206.
- [25]. Martin, G.B., Frary, A., Wu, T., Bmmmonschenkl, S., Chunwongse, J., Earle, E.D. and Tanksley, S.D. (1993b) A member of the tomato *Pto* gene family confers sensitivity to fenthion resulting in rapid cell death. Plant Cell 6, 1543-1552.
- [26]. Dixon, M.S., Jones, D.A., Keddie, J.S., Thomas, C.M., Harison, K. and Jones, J.D.J. (1996) The tomato *cf-2* disease resistance locus comprises two functional genes encoding leucine rich repeat proteins. Cell 84, 451-459.
- [27]. Dixon, M.S., Hatzixanthis, K., Jones, D.A., Harrison, K. and Jones, J.D.G. (1998) The tomato *cf-5* diseases resistance gene and six homologs show pronounced allelic variation in leucine rich repeat copy number. Plant Cell 10, 1915-1925.

- [28]. Thomas, C.M., Jones, D.A., Parniske, M., Harrison, K., Balintkurti, P.J., Hatzixanthis, K. and Jones, J.D.G. (1997) Characterization of the tomato *cf-4* gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in *cf-4* and *cf-9*. Plant Cell 9, 2209-2224.
- [29]. Jones, D.A., Thomas, C.M., Hammond-Kossack, K.E., Balinkt, P.J. and Jones, J.D.G. (1994) Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by Transposon tagging. Sci 266, 789-793.
- [30]. Song, W.Y., Wang, G.I., Chen, L.L., Kim, H.S., Pi, L.Y., Holsten, T., Gardner, J., Wang, B., Zhai, W.X., Zhu, L.H., Fauquet, C. and Ronald, P. (1995) A receptor kinase like protein encoded by the rice disease resistance gene, *Xa21*. Sci 270, 1804-1806.
- [31]. Collins, N.C., Webb, C.A., Seah, S., Ellis, J.G., Hulbert, S.H. and Pryor, A. (1998) The isolation and mapping of disease resistance analogs in maize. Mol. Plant Microbe. 11, 968-978.
- [32]. Hammond-Kosack, K.E. and Jones, J.D.G. (1997) Plant disease 6 resistance genes. Annu. Rev. Plant. Physiol. Plant Mol. Biol. 48:575-607.
- [33]. Jones, D.A. and Jones, J.D.G. (1997) The role of leucine-rich repeat proteins in plant defenses. Advanced Bot. Res. 24, 90-167.
- [34]. Meyers, B.C., Kozik, A., Griego, A., Kuang, H. and Michelmore, R.W. (2003) Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. Plant Cell Online 15, 809-834.

- [35]. Matte, A., Tari, L.W., Goldie, H. and Delbaere, L.T. (1997) Structure and mechanism of phosphorenolpyruvate carboxykinase. J. Biol. Chem. 272, 8105-8108.
- [36]. Traut, T.W. (1994) The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. Eur. J. Biochem. 222, 9-19.
- [37]. Kar, B., Nanda, S., Nayak, P.K., Nayak, S. and Joshi, R.K. (2013) Molecular characterization and functional analysis of a coiled-coil-nucleotide-binding-siteleucine-rich repeat R gene from *Curcuma zedoaria* Loeb. that confers resistance to *Pythium aphanidermatum*. Physiol. Mol. Plant Pathol. 83, 59-68.
- [38]. Tarr, D.E.K. and Alexander, H.M. (2009) TIR-NBS-LRR genes are rare in monocots: evidence from diverse monocot orders. BMC Res. Notes 2:197.
- [39]. Meyers, B.C., Dickerman, A.W., Michelmore, R.W., Sivaramakrishnan, S., Sobral, B.W. and Young, N.D. (1999) Plant disease resistance genes encodes members of an ancient and diverse protein family within the nucleotide-binding superfamily. Plant J. 20:317-332.
- [40]. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning, Vol. 2, pp: 14-19, Cold spring harbor laboratory press, New York, USA.
- [41]. Kumar, S., Koichiro, T. and Masatoshi, N., (2004) MEGA 3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings Bioinform. 5, 150-163.