Plant Knowledge Journal
ISSN: 2200-5390
EISSN: 2200-5404
EISSN: 2200-5404

Southern Cross Publishing Group
AUSTRALIA

5(1):1-8 (2016)

Cloning and analysis of the NBS-LRR gene family in finger millet (*Eleusine coracana* L.) (Gaertn.)

Dipnarayan Saha[1, 2,*] and Rajeev Singh Rana[1, 2]

[1]Division of Genomic Resources, National Bureau of Plant Genetic Resources, Pusa Campus, New Delhi 110012 INDIA
[2]Present address: Central Research Institute for Jute and Allied Fibres, Barrackpore, Kolkata 700120, INDIA

***Corresponding author: dipsaha72@yahoo.com**

Abstract

Nucleotide binding site leucine rich repeat protein (NBS-LRR) is a multi-member gene family in plants encoding important disease resistance proteins to fend against various pathogen infections. Recent advances in genetics and genomics facilitated discovery and functional analysis of NBS-LRR genes or resistance gene homologs (*RGHs*) in various plant species. In the present work, we have cloned 57 non-identical NBS-LRR sequences from a blast disease resistant finger millet genotype IE1012. We assembled the 57 NBS-LRR sequences to the existing finger millet NBS-LRRs from NCBI Genbank through CAP3 program to obtain a total of 28 sequences. The hidden Markov models (HMMs) analysis further identified 16 NBS-LRRs as *EcRGHs* with uninterrupted open reading frames (ORFs), of which nine *EcRGHs* was added from the present work. The secondary structure analysis of protein sequences revealed characteristic conserved motifs, P-loop, Kinase2 and GLPL and other motifs with significant variations. The phylogenetic classification clustered the 16 EcRGH proteins, mostly in the non-TIR group. BLAST-P analysis showed homology of the 16 EcRGHs to NBS-LRR proteins reported across other grass species. The *EcRGHs* analysed here providing a useful genomic resource for genetic studies of resistance genes in finger millet and exploit them in resistance breeding programs.

Keyword: finger millet; gene family; NBS-LRR; resistance gene homologs.
Abbreviations: BLAST_Basic Local Alignment Search Tool; bp_base pairs; DNA_deoxy ribonucleic acid; HMM_hidden Markov model; NBS-LRR_nucleotide binding site - leucine-rich repeat; NCBI_National Center for Biotechnology Information; ORF_open reading frame; PCR_polymerase chain reaction; RGH_resistance gene homolog; TIR_Toll/interleukin-1 receptor.

Introduction

Finger millet, *Eleusine coracana* L. (Gaertn.), is an important food, feed and industrial crop of the semi-arid tropics. It belongs to the subfamily Cholridoideae of the Poaceae family and was originated in East Africa and India. The cultivated finger millet has an allotetraploid genome (AABB). The grains of finger millet are a rich and balanced source of nutrition. Despite its 'minor crop' tag, this crop is gaining huge world-wide attention in genetics and genomics research owing to its nutritional and agricultural importance (Dida and Devos, 2006; Dida et al., 2007; Babu et al., 2014; Goron and Raizada, 2015). Finger millet blast disease, caused by the blast pathogen, *Magnaporthe oryzae*, is an economically important disease resulting in significant crop loss (Nagaraja et al., 2007; Getachew et al., 2014), particularly the neck and finger blast is most serious (Takan et al., 2012). Since, the majority of the cultivated finger millet has a narrow genetic base due to self-breeding behaviour, the genomic resource development and resistance breeding for blast disease is considered the most efficient approach in genetic improvement of finger millet. In plants, the predominant class of resistance (R) genes encode nucleotide binding site - leucine rich repeat (NBS-LRR) proteins to combat against various pest and pathogen infections and diseases (Jones and Jones, 1997; Meyers et al., 2003). The NBS-LRR protein encoding genes constitutes a large gene family with highly conserved NBS and LRR domains and variable amino and carboxy-terminal domains (McHale et al., 2006). The NBS domain, which is involved in mediating signalling response, is organized in highly conserved motifs, such as P-loop, kinase-2 and a hydrophobic glycine-leucine-proline-leucine (GLPL). Whereas, the LRR domains are highly adaptive structural domain and are involved in regulating macromolecular binding/interaction specificities. The presence or absence of the Toll/interleukin-1 receptor (TIR) and coiled-coil (CC) domains in the variable N-terminal domains are the basis to classify the NBS-LRRs into two groups (McHale et al., 2006; Marone et al., 2013). Members of this large gene family are often located in clusters at the specific chromosomal location, and their expansion in a given plant species is a result of segmental and/or tandem gene-duplication events during evolution (Leister, 2004). In general, there is ample variability within the NBS-LRR genes across plant species in terms of the number of candidate sequences, genomic organizations and functionality. Thus, the NBS-LRR protein coding genes were cloned and analysed at the genome-scale from numerous plant species (Jacob et al., 2013). In grass family, the NBS-LRR genes were characterized in rice (Zhou et al., 2004), sorghum (Cheng et al., 2010), foxtail millet (Weng et al., 2009) and *Brachypodium* (Tan and Wu, 2012). However, the NBS-LRR sequences in finger millet were identified in small numbers (Panwar et al., 2011) compared to other grass species because

of the non-availability of genome sequence. In the present study, we report cloning of NBS-LRR genes from a blast disease resistant finger millet genotype, IE1012 in supplementation to the existing list of finger millet NBS-LRRs of NCBI Genbank database. We also analyse their phylogeny, peptide features and their comparative genetic relationship with the orthologous NBS-LRRs from other grass species.

Results and Discussion

Cloned NBS-LRR sequences from finger millet

The public domain sequence database NCBI, consisted of a small number of resistance gene analogue (NBS-LRR) sequences from finger millet. Mining of the NCBI nucleotide database with 'Eleusine coracana' and 'NBS-LRR' helped us to retrieve 44 sequences (as on 20th April-2015), which consisted of redundant and/or overlapping sequences. In order to PCR amplify, clone and sequence NBS-LRR genes from genomic DNA of a blast disease resistant *E. coracana* genotype, IE1012, we employed eight combinations of degenerate primers targeting conserved NBS domain (Table 1). The PCR amplified products ranged from 250-800 bp, with a majority of them showing a single band. In altogether, 280 recombinant plasmid clones consisting of the PCR products were sequenced using Sanger's sequencing method to obtain 70 good-quality vector sequence-cleaned and putative NB-ARC domain-positive sequences. These sequences were verified from the Pfam and the NCBI BLASTX program showing homology to known NBS-LRR genes with high E-value. The primary processing by eliminating identical sequences from these 70 putative NBS-LRRs using the Galaxy FASTA manipulation tool produced 57 non-identical sequences. The length of these 57 sequences ranged from 239-667 bp with an average of 400 bp per sequence.

Assemblage and ORF predictions in finger millet NBS-LRR sequences

The 44 sequences mined from the NCBI database, and the 57 NBS-LRR sequences cloned in the present study were assembled using the online CAP3 program to eliminate redundant sequences as per the default criterion of the program. The CAP3 assembly program generated a total of 28 sequences, including 19 contigs and nine singletons (Table 2). Twelve of the contigs and four of the singletons were derived from the sequences cloned from the present study. The majority of the NBS-LRRs (30 sequences) mined from the NCBI database was assembled into three contigs (contig 2-4). All these contigs and singletons were further confirmed through the NCBI BLASTX program against the nucleotide collection nr/nt database showing high homology (E-value $>10^{-30}$) to either known NBS-LLR or disease resistance genes. Scanning of NB-ARC domain in the assembled sequences using Pfam and HMM resulted us with 16 sequences, including 14 contigs and two singletons with NB-ARC domain (Pfam00931) and uninterrupted ORFs i.e. without any premature stop codon. Out of the 16 assembled NBS-LRRs of finger millet, the present study contributed nine NBS-LRR sequences, including seven contigs and two singletons. All these 16 assembled NBS-LRR sequences from finger millet are hereafter referred as *Eleusine coracana* resistant gene homologs (EcRGHs) and given arbitrary IDs (EcRGH001-EcRGH016) for the subsequent analysis (Table 2).

Conserved motifs in the EcRGHs

The amino acid motif composition of the NBS-LRR genes is highly conserved across the plant kingdom. However, the relative occurrences of these conserved motifs vary within the members of the gene family (Meyers et al., 2003). The distribution pattern of conserved motifs present in the unaligned 16 EcRGHs was analysed using the online tool MEME suite (Fig. 1). Different motifs were found conserved in the NBS domain across the 16 EcRGHs. The three major conserved motifs were P-loop or Kinase-1, Kinase-2 and GLPL motif, which were found consistent in all the EcRGHs. Interestingly, variations in the distribution pattern of these motifs were detected in a few sequences. For instance, in EcRGH010 no significant conserved motif was found. Likewise, the Kinase2 motif was found absent in EcRGH009. Few EcRGH sequences lacked P-loop motifs. One or more weakly significant motif was also found repeated in a few sequences. Sequence logo along with the amino acid sequences of the three conserved motifs were depicted in Fig. 1.

Multiple sequence alignment and phylogeny of the EcRGHs

A multiple sequence alignment performed with the deduced amino acid sequences of the 16 EcRGHs and three each R-gene representing TIR (N, L6, AtRPP5) and non-TIR (AtRPM1, AtRPS2 and Mi) NBS-LRR protein class showed the degree of conservation in the NBS domain (Fig. 2). The consensus P-loop motif (GMGGVGKT) showed few substitutions in place of V residue; such as I, L, S and M. Variations were also detected in the amino acid sequence of the GLPL motif. The EcRGH001, 008, 011 and 012 sequences consists of one amino acid substitution, GFPL sequence instead of standard GLPL. Similarly, the EcRGH002 and 009 consisted of GHPL and GIPL sequences instead GLPL. In contrast, EcRGH010 consists of entirely different amino acids, QHYL motif in place of GLPL. Variations in the highly conserved Kinase2 (VLDD) motif were too witnessed in the EcRGHs. With few exceptions, the majority of the Kinase2 motif in the EcRGHs consisted of adjacent conserved tryptophan (W) residue, which is characteristic of the TIR-NBS-LRR class (Tarr and Alexander, 2009). Variations observed in other motifs, such as Kinase-3, RNBS-A, B and D were very prominent in comparison to the major conserved motifs. The phylogenetic analysis clustered the 16 EcRGHs and the six known NBS-type R-gene proteins from other plants into two distinct groups, TIR class and non-TIR class (Fig. 3). The confidence of clustering outputs was corroborated with moderate to high bootstrap values in the nodes (data not presented). In the neighbour-joining phylogenetic tree, the majority of the EcRGHs fall into the non-TIR class, whereas only two EcRGHs, namely EcRGH007 and EcRGH010 grouped in TIR class. The results obtained from the current study more or less support the hypothesis that the TIR-NBS-LRR class is rarely present in monocots (Tarr and Alexander, 2009). The exception of two TIR-NBS-LRR type EcRGHs in finger millet, however, corroborates the presence of rare and divergent TIR-NBS-LRR in the rice genome (Bai et al.,

Table 1. Details of degenerate primers used in the present study to clone NBS-LRR genes from finger millet.

Sl. No.	Primer combinations	Sequences (5 '- 3')*	Motifs (amino acids)	Degene racy	~PCR Fragment size (bp)	References
1	Ploop For / Kin2 Rev	GGIGGIRTIGGIAAIACIAC	GG(V/I)GKT	2048	260	Hunger et al., 2003
		CCAIACATCATCMAGSACAA	(L/I/V)LVLDDVW	13824		
2	Ploop For / Kin 2 TIR Rev	GGIGGIRTIGGIAAIACIAC	GG(V/I)GKT	2048	300	-do-
		ATCIACATCATCNAGMACRA	(L/I/V)LVLDDVD	27648		
3	Ploop For / GLPL Rev	GGIGGIRTIGGIAAIACIAC	GG(V/I)GKT	2048	520	-do-
		IAGIGYIARIGGIAGICC	GLPL(T/A)L	13824		
4	NBS-1LR For / NBS-1LR Rev	GGYATGGGNGGYMTHGGNAARAC	GMGG(V/I)GKT	8192	250	Weng et al., 2009
		CCANACATCATCMAGSACAA	L(I/V)VLDD	2304		
5	Ploop2 F / GLPL2R	GGWATGGGWGGWRTHGGWAARACHAC	GGVGKTT	8192	510	Lee et al., 2003
		ARNWYYTTVARDGCVARWGGVARWCC	GLPLAL	13824		
6	Ploop2 F / RNBS-D R	GGWATGGGWGGWRTHGGWAARACHAC	GGVGKTT	8192	800	Bertioli et al., 2003
		GGRAAIARISHRCARTAIVIRAARC	CFLYCALFP	18432		
7	Ploop For / L2	GGIGGIRTIGGIAAIACIAC	GG(V/I)GKT	2048	700	Yaish et al., 2004
		CCKNSAGYMNRTCRTGCAT	MHDV	2048		
8	Ploop For / Q1	GGIGGIRTIGGIAAIACIAC	GG(V/I)GKT	2048	650	-do-
		RAARCAIGCSATRTCIARRAA	FLDIACF	1152		

*International Union of Pure and Applied Chemistry (IUPAC) ambiguity codes for nucleotides: R= A/G; Y= C/T; K= G/T; S= C/G; M=C/A; W= A/T; H=A/C/T; V= A/C/G; D= A/G/T, N= A/G/C/T, I= Inosine base

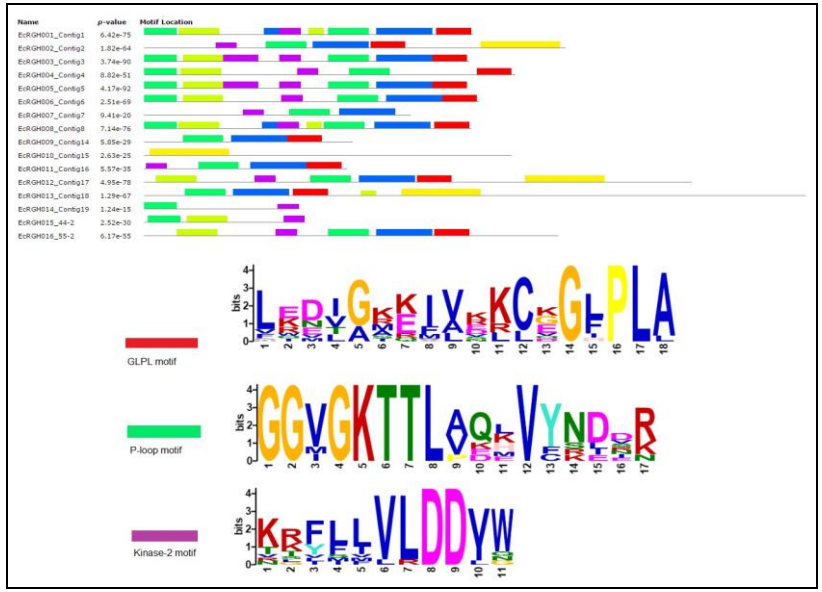

Fig 1. Distribution pattern of conserved motifs in the NBS domain were identified and depicted using MEME suite. Different coloured boxes indicate separate and distinct motifs. The MEME "score" for the overall match of the protein to the motif models is given as a P- value adjacent to the sequence IDs. The sequence logos of the predominant three motifs were depicted below.

Table 2. Details of the NCBI BLASTX and Pfam analysis of the EcRGHs.

EcRGH-IDs	CAP3 analysis		NCBI BLASTX analysis			Pfam analysis		
	Contig / singleton	No of seq in contigs	Description	E value	Accession	Pfam domain	HMM From	HMM To
EcRGH001	Contig1	2	NBS-LRR disease resistance protein [*Eleusine coracana*]	3E-117	ADB12239.1	NB-ARC	28	198
EcRGH002	Contig2	9	NBS-LRR-like disease resistance protein [*Eleusine coracana*]	4.00E-129	ABW04991.1	NB-ARC	86	279
EcRGH003	Contig3	8	NBS-LRR-like disease resistance protein [*Eleusine coracana*]	2.00E-116	ABW04983.1	NB-ARC	28	199
EcRGH004	Contig4	13	NBS-LRR-like disease resistance protein [*Eleusine coracana*]	1.00E-136	ABW04973.1	NB-ARC	28	199
EcRGH005	Contig5	3	NBS-LRR-like disease resistance protein [*Eleusine coracana*]	8.00E-114	ABW04978.1	NB-ARC	28	198
EcRGH006	Contig6	4	NBS-LRR-like disease resistance protein [*Eleusine coracana*]	3.00E-121	ABW04965.1	NB-ARC	28	199
EcRGH007	Contig7	2	Putative disease resistance protein [*Aegilops tauschii*]	2.00E-14	EMT17905.1	NB-ARC	50	188
EcRGH008	Contig8	2	NBS-LRR disease resistance protein [*Eleusine coracana*]	1.00E-67	ADB12239.1	NB-ARC	28	199
EcRGH009	Contig14	3	PREDICTED: putative disease resistance protein RGA4-like [*Setaria italica*]	4.00E-40	XP_004956600.1	NB-ARC	119	206
EcRGH010	Contig15	6	Leucine Rich Repeat family protein, expressed [*Oryza sativa* Japonica Group]	2.00E-93	ABA94704.2	NB-ARC	238	286
EcRGH011	Contig16	4	PREDICTED: disease resistance protein RGA2-like isoform X2 [*Setaria italica*]	5.00E-27	XP_004977426.1	NB-ARC	103	200
EcRGH012	Contig17	2	putative bacterial blight-resistance protein Xa1 [*Oryza sativa* Indica Group]	1.00E-119	BAI39872.1	NB-ARC	43	286
EcRGH013	Contig18	6	PREDICTED: disease resistance protein RPM1-like isoform X1 [*Setaria italica*]	3.00E-170	XP_004956154.1	NB-ARC	104	285
EcRGH014	Contig19	2	NB-ARC domain containing protein, expressed [*Oryza sativa* Japonica Group]	6.00E-24	ABA96699.1	NB-ARC	28	110
EcRGH015	singleton	1	NBS-LRR disease resistance protein [*Musa* AAB Group]	2.00E-26	CAP66296.1	NB-ARC	26	111
EcRGH016	singleton	1	NBS-LRR resistance protein [*Saccharum* hybrid cultivar N11]	2.00E-111	AAQ16581.1	NB-ARC	40	245

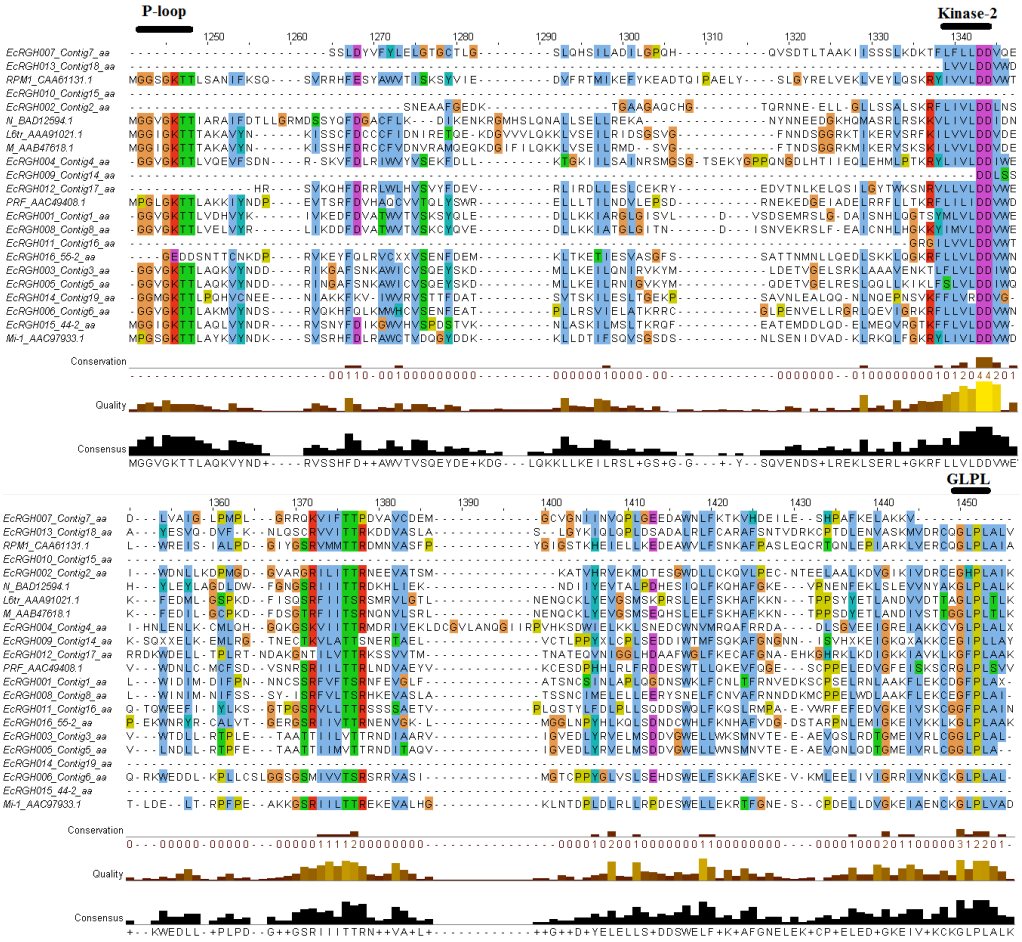

Fig 2. Multiple sequence alignment of the NBS domain of 16 EcRGHs along with six known NBS-LRR type R-genes from other plant species generate by ClustalW2 program and visualized using Jalview. The known predominant conserved motifs were demarcated and shown above the sequence.

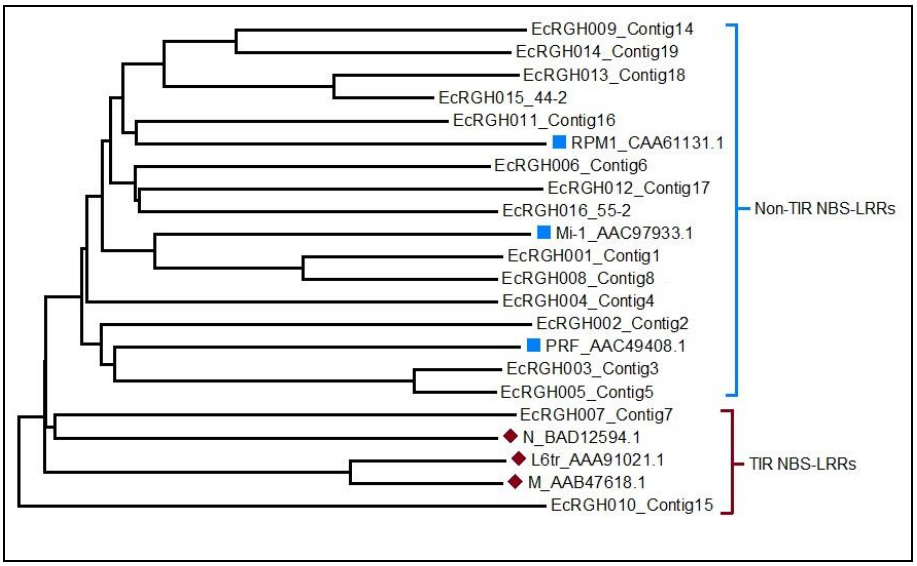

Fig 3. A NJ-phylogenetic tree derived from the 16 EcRGHs and six known NBS-LRR type R-genes from different plant species. The image was generated using MEGA6.0 and Clustal Omega alignment of the NBS domain. The 16 EcRGHs were clustered into non-TIR groups, except two EcRGHs (EcRGH007 and 010) grouped with known TIR NBS-LRR proteins. The reference proteins for TIR and non-TIR classifications were depicted with red rhombus and blue squares respectively.

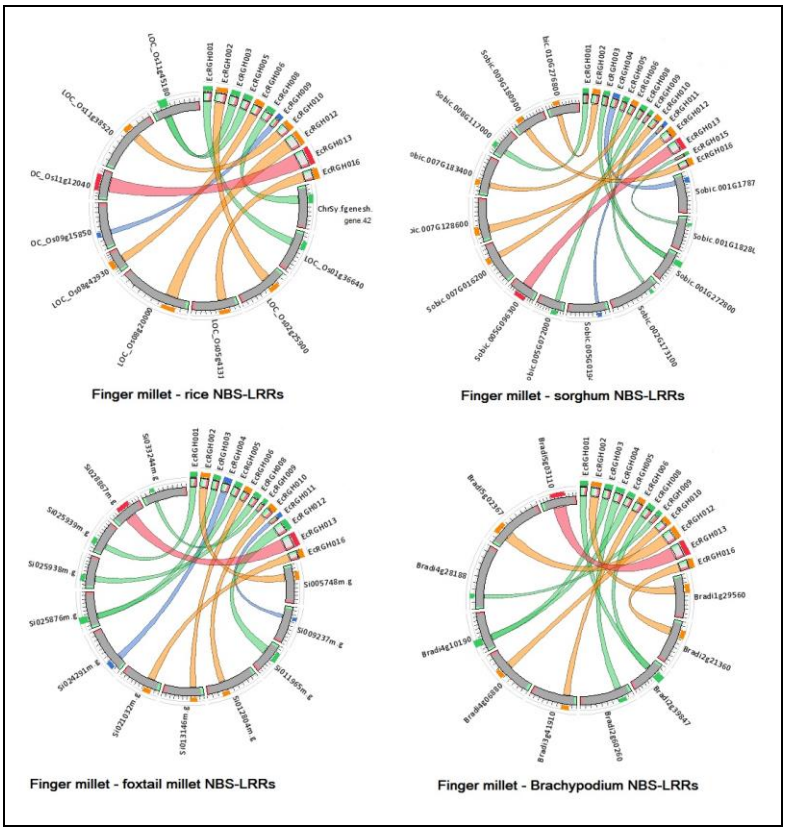

Fig 4. Ideogram showing homologus relationship among EcRGHs and NBS-LRR proteins from other grass species. Each ribbons connecting gene pair represents local alignment by BLAST-P at 10^{-30} E-value cut-off. The color of the ribbons corresponds to the alignment E-value score in four quartile; red indicates highest value, while the orange, green and blue illustrate the relative lower values. Only the single best hits were depicted by keeping off the ribbons untangled.

2002). These exceptional TIR-NBS-LRRs in grass species like rice and finger millet might serve as excellent candidates for studying resistance gene evolution and genome synteny.

Comparative genetic homology of EcRGHs to other NBS-LRRs of grass origin

A BLAST-P based tool combined with the CIRCOS visualization was used to understand comparative relationship (homology) individually to other NBS-LRR gene families from sequenced grass species, such as rice, sorghum, foxtail millet and *Brachypodium* (Fig. 4). It was observed that the majority (14 of 16) EcRGH sequences could find homology at E-value > 10^{-30}. In all the four genomes, the EcRGH013 produced the highest similarity match (E-value >10^{-150} and >60% identity). In overall, nearly the same set of EcRGHs showed high sequence homology to NBS-LRR sequences from other grass species. This indicates that possibly these RGHs shares conserved homologous sequences across the grass family probably much before the evolutionary divergence of the grasses.

Materials and Methods

Plant material and genomic DNA extractions

E. coracana improved cultivar IE1012 (IC432860; GE-669), which is highly resistant to finger millet blast disease was obtained from NBPGR, New Delhi. The seeds were grown in pots and watered with tap water for one month, before the fresh leaves were harvested for genomic DNA extractions. The genomic DNA was extracted using the cetyl trimethyl ammonium bromide (CTAB) method from Doyle and Doyle (1990) with minor changes. The quality and quantity of the DNA was checked by agarose gel (1%) electrophoresis and UV-Vis spectrophotometer (Nanodrop 1000, Thermo Fisher Scientific Inc. USA), respectively. The working concentration of genomic DNA for PCR amplification was reconstituted to 50 ng/µl using Tris (10 mM) -EDTA (1 mM) buffer.

Polymerase chain reaction (PCR) and clone sequencing

Eight sets of oligonucleotide degenerate primers consisting of base pairs with ambiguous nucleotide codes as per the International Union of Pure and Applied Chemistry (IUPAC) were used to target amplification of NBS domain (Table 1). Degenerate nucleotide codes are based on 'degeneracy of codon' that specify an amino acid. It acts as a mixture of similar primer sequences with a number of possible bases in the respective coded nucleotides (Linhart and Shamir, 2005). These primers were derived from Bertioli et al. (2003), Hunger et al. (2003), Lee et al. (2003), Yaish et al. (2004) and Weng et al. (2009). The PCR reactions were performed in a total reaction volume of 50 µl containing as follows: 200 ng of genomic DNA as template, 1.0 x PCR buffer with $(NH_4)SO_4$ (Fermentas Inc, USA), 2.0 mM $MgCl_2$, 0.2 mM dNTP mix (Fermentas Inc, USA), 0.8 µM each forward and reverse degenerate primers and 1.0 unit of *Taq* DNA polymerase (Fermentas Inc, USA). The PCR amplifications were performed in a C1000 Touch^{TM} Thermal Cycler (Bio-Rad, USA) as followed: Initial denaturation of template DNA

at the 94 ^{o}C for 10 min, followed by 38 cycles of denaturation at the 94 ^{o}C for 45 Sec, primer annealing at 48-52 ^{o}C for 1 min and strand extension at the 72 ^{o}C for 2 min. The final extension of PCR products was carried out at the 72 ^{o}C for 5 min. The PCR products were resolved on Ethidium bromide-stained 1.0 % agarose gel and electrophoresed in 1 x TBE buffer at 120 V until the bands were clearly separated. The clear and resolved bands (ranging from ~250-700 bp) were individually gel-purified using the MinElute gel extraction kit (Qiagen Inc.) and ligated into pGEM-T Easy vector (Promega) as per the manufacturers' instruction. The PCR-product ligated vectors were transformed into chemically competent *E. coli* cells and blue-white colony selected on ampicillin resistant *Luria Bertani* agar plates. The positive clones (~190) with PCR product inserts were further confirmed through *Eco*RI restriction digestion of the pGEM-T Easy plasmid and sequenced from Chromous Biotech Pvt. Ltd., India.

DNA sequence analysis and data retrieval from NCBI database

The raw DNA sequence of the cloned PCR products was cleaned of vector sequences using NCBI VecScreen (http://www.ncbi.nlm.nih.gov/tools/vecscreen/) and were confirmed in the presence of NB-ARC domain using the NCBI BLASTX tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastHome) against the near database. The sequences were processed and duplicates were eliminated using the WWW-based Galaxy (https://usegalaxy.org/) FASTA manipulation tool. The available *E. coracana* NBS-LRR sequences were retrieved from the NCBI nucleotide database (during Apr 2015). The combined sequences were assembled using the online CAP3 program (http://doua.prabi.fr/software/cap3) (Huang and Madan, 1999) to get rid of the redundant sequences. The reference R-genes with NBS-LRR domains for phylogenetic study was derived from NCBI (http://www.ncbi.nlm.nih.gov/). The NBS-LRR sequences of other grass species for comparative genetics study were derived from Phytozomev10.1 (http://phytozome.jgi.doe.gov/pz/portal.html).

Analysis of multiple sequence alignment and conserved amino acid domains

The amino acid sequences of the cloned and downloaded NBS-LRR genes were translated into six frame amino acid sequences using the software tool EMBOSS Transeq (http://www.ebi.ac.uk/Tools/st/emboss_transeq/) and sorted the functional peptides by manual comparison with NCBI BLASTX (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastx&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) results and Pfam 27.0 search (http://pfam.xfam.org/) (Finn et al., 2014). The selected peptide sequences with NB-ARC domain and six reference R-gene amino acid sequences (N: BAD12594.1; L6tr: AAA91021.1; M: AAB47618.1; PRF: AAC49408.1; RPM1:CAA61131.1; Mi-1: AAC97933.1) were used for multiple sequence alignment using the software tool ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) (McWilliam et al., 2013). These sequences were also used for phylogenetic tree reconstruction using amino acid substitution, P-distance model, 1000 bootstrap replicates and NJ method in the software program MEGA 6.06 (Tamura et al., 2013). The conserved motifs in the NBS-LRR peptide sequences were analysed suing the online program MEME suite 4.10.0 (http://meme.nbcr.net/meme/doc/cite.html) (Bailey et al., 2009).

Comparative genetic analysis

A comparative genetic approach to identify orthologs of *E. coracana* NBS-LRR proteins from other grass species, such as rice, sorghum, foxtail millet and *Brachypodium* was carried out through BLASTP 2.2.25 [Feb-01-2011] using the software tool Circoletto 01.11.13 (http://tools.bat.infspire.org/circoletto/#info) (Darzentas, 2010). The *E. coracana* NBS-LRR peptide sequences were used as a query sequence to search rice, sorghum, foxtail millet and *Brachypodium* NBS-LRRs as individual databases at strict E-value (10^{-30}) cut-off and only the best hits were visualized in Circos like figure.

Conclusion

In the present study, nine new RGHs including seven contigs and two singletons were supplemented to the existing seven functional contigs resulted from the individual NBS-LRR sequences from NCBI. Significant variations in amino acid substitution were noticed in the conserved regions or motifs of the EcRGHs studied. Altogether, the 16 EcRGH sequences were phylogenetically classified to reveal a majority of them falling into non-TIR class with the exceptions of two EcRGHs. The comparative genetic relationship with homologous NBS-LRR proteins from the grass family identified putative conserved RGHs from the grass family owing to high and consistent sequence similarity. The present study, provides an important information on analysis of this gene family in finger millet that can aid in identification of disease resistance genes and/or functional markers for resistance breeding.

Acknowledgements

Author gratefully acknowledges NBPGR for providing sources and supporting necessary research facilities.

References

Babu BK, Dinesh P, Agrawal PK, Sood S, Chandrashekara C, Bhatt JC, Kumar A (2014) Comparative genomics and association mapping approaches for blast resistant genes in finger millet using SSRs. PLoS ONE. 9: e99182. doi:10.1371/journal.pone.0099182

Bai J, Pennill LA, Ning J, Lee SW, Ramalingam J, Webb CA, Zhao B, Sun Q, Nelson JC, Leach JE, Hulbert SH (2002) Diversity in nucleotide binding site-leucine-rich repeat genes in cereals. Genome Res. 12: 1871-1884.

Bailey TL, Bodén M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acid Res. 37: W202-W208. doi: 10.1093/nar/gkp335.

Bertioli DJ, Leal-Bertioli SCM, Lion MB, Santos VL, Pappas Jr G, Cannon SB, Guimaraes PM (2003) A large scale analysis of resistance gene homologues in Arachis. Mol Gen Genomics. 270: 34-45.

Cheng X, Jiang H, Zhao Y, Qian Y, Zhu S, Cheng B (2010) A genomic analysis of disease-resistance genes encoding nucleotide binding sites in *Sorghum bicolor*. Genet Mol Biol. 33: 292-297.

Darzentas N (2010) Circoletto: visualizing sequence similarity with Circos. Bioinformatics. 26: 2620-2621.

Dida M, Devos K (2006) Finger millet. In: Kole C (ed) Genome mapping and molecular breeding in plants, Vol 1,

Cereals and millets, Springer, Heidelberg, Berlin, p.333-343.

Dida MM, Srinivasachary, Ramakrishnan S, Bennetzen JL, Gale MD, Devos KM (2007) The genetic map of finger millet, *Eleusine coracana*. Theor Appl Genet. 114: 321-332

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus. 12: 13-15.

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M (2014) The Pfam protein families database. Nucleic Acid Res. 42: D222-D230. doi: 10.1093/nar/gkt1223

Getachew G, Alemu T, Tesfaye, KJ (2014) Disease incidence, severity and yield loss of finger millet varieties and invitro mycelial growth inhibition of *P. grisea* using biological antagonists and fungicides. J Appl Biosci. 73: 5883-5901.

Goron TL, Raizada MN (2015) Genetic diversity and genomic resources available for the small millet crops to accelerate a new green revolution. Front Plant Sci. 6: 157. doi: 10.3389/fpls.2015.00157

Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. Genome Res. 9: 868-877.

Hunger S, Di Gaspero G, Mohring S, Bellin D, Schafer Pregl R, Borchardt DC, Durel CE, Werber M, Weisshaar B, Salamini F, Schneider K (2003) Isolation and linkage analysis of expressed disease-resistance gene analogues of sugar beet (Beta vulgaris L.). Genome. 46: 70-82.

Jacob F, Vernaldi S, Maekawa T (2013) Evolution and conservation of plant NLR functions. Front Immunol. 4: 1-16. doi: 10.3389/fimmu.2013.00297

Jones DA, Jones JDG (1997) The role of leucine-rich repeat proteins in plant defenses. Adv Bot Res. 24: 90-167.

Lee SY, Seo JS, Rodriguez-Lanetty M, Lee DH (2003) Comparative analysis of superfamilies of NBS-encoding disease resistance gene analogs in cultivated and wild apple species. Mol Genet Genome. 269: 101-108.

Leister D (2004) Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance gene. Trend Genet. 20: 116-122.

Linhart C, Shamir R (2005) The degenerate primer design problem: theory and applications. J Comput Biol. 12: 431-456.

McHale L, Tan X, Koehl P, Michelmore RW (2006) Plant NBS-LRR proteins: adaptable guards. Genome Biol. 7: 212. doi:10.1186/gb-2006-7-4-212

McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, Cowley AP, Lopez R (2013) Analysis tool web services from the EMBL-EBI. Nucleic Acid Res. 41: W597-W600. doi: 10.1093/nar/gkt376.

Marone D, Russo MA, Laidò G, De Leonardis AM, Mastrangelo AM (2013) Plant nucleotide binding site–leucine-rich repeat (NBS-LRR) genes: active guardians in host defense responses. Int J Mol Sci. 14: 7302-7326.

Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW (2003) Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. Plant Cell. 15: 809-834.

Nagaraja A, Jagadish PS, Ashok EG, Krishne Gowda KT (2007) Avoidance of finger millet blast by ideal sowing time and assessment of varietal performance under rainfed production situations in Karnataka. J Mycopathathol Res. 45: 237-240.

Panwar P, Jha AK, Pandey PK, Gupta AK, Kumar A (2011) Functional markers based molecular characterization and cloning of resistance gene analogs encoding NBS-LRR disease resistance proteins in finger millet (*Eleusine coracana*). Mol Biol Rep. 38: 3427-3436.

Takan JP, Chipili J, Muthumeenakshi S, Talbot NJ, Manyasa EO, Bandyopadhyay R, Sere Y, Nutsugah SK, Talhinhas P, Hossain M, Brown AE, Sreenivasaprasad S (2012) *Magnaporthe oryzae* populations adapted to finger millet and rice exhibit distinctive patterns of genetic diversity, sexuality and host interaction. Mol Biotechnol. 50: 145-158.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 30: 2725-2729.

Tan S, Wu S (2012) Genome wide analysis of nucleotide-binding site disease resistance genes in *Brachypodium distachyon*. Comp Funct Genomics. 2012: 1-12. doi:10.1155/2012/418208

Tarr DE, Alexander HM (2009) TIR-NBS-LRR genes are rare in monocots: evidence from diverse monocot orders. BMC Res Notes. 2: 197. doi:10.1186/1756-0500-2-197

Yaish MWF, Saenz de Miera LE, Perez de la Vega M (2004) Isolation of a family of resistance gene analogue sequences of the nucleotide binding site (NBS) type from Lens species. Genome. 47: 650-659.

Weng Q, Li Z, Xing J, Dong Z, Dong J (2009) Cloning and structure analysis of an NBS-LRR disease-resistant gene from *Setaria italica* Beauv. Front Agri China. 3: 240-246.

Zhou T, Wang Y, Chen JQ, Araki H, Jing Z, Jiang K, Shen J, Tian D (2004) Genome-wide identification of NBS genes in japonica rice reveals significant expansion of divergent non-TIR NBS-LRR genes. Mol Genet Genomics. 271: 402-415.