Isolation and characterization of nucleotide-binding site and C-terminal leucine-rich repeat-resistance gene candidates in bananas

Y. Lu[1,2], W.H. Xu[3], Y.X. Xie[2], X. Zhang[2], J.J. Pu[2], Y.X. Qi[2] and H.P. Li[1]

[1]College of Natural Resources and Environment,
South China Agriculture University, Wushan, China
[2]Environment and Plant Protection Institute,
Chinese Academy of Tropical Agriculture Sciences, Hainan, China
[3]Shangrao Normal University, Zhimin State in Shangrao, JiangXi, China

Corresponding author: H.P. Li
E-mail: ytluy2010@163.com

ABSTRACT. Commercial banana varieties are highly susceptible to fungal pathogens, as well as bacterial pathogens, nematodes, viruses, and insect pests. The largest known family of plant resistance genes encodes proteins with nucleotide-binding site (NBS) and C-terminal leucine-rich repeat (LRR) domains. Conserved motifs in such genes in diverse plant species offer a means for the isolation of candidate genes in banana that may be involved in plant defense. Six degenerate PCR primers were designed to target NBS and additional domains were tested on commercial banana species *Musa acuminata* subsp *malaccensis* and the *Musa* AAB Group propagated *in vitro* and plants maintained in a greenhouse. Total DNA was isolated by a modified CTAB extraction technique. Four resistance gene analogs were amplified and deposited in GenBank and assigned numbers HQ199833-HQ199836. The predicted amino acid sequences compared to the amino acid sequences of known resistance genes (*MRGL1*, *MRGL2*, *MRGL3*, and *MRGL4*) revealed significant sequence similarity. The presence of consensus do-

mains, namely kinase-1a, kinase-2 and hydrophobic domain, provided evidence that the cloned sequences belong to the typical non-Toll/interleukin-1 receptor-like domain NBS-LRR gene family.

Key words: Resistance genes; NBS-LRR; Banana

INTRODUCTION

Banana is one of the most important fruit crops in the world. This fruit constitutes a major staple food for millions of people in developing countries, and it is also considered the highest selling fruit in the world market, making this crop a very important export commodity for the producing countries (Jones, 2000; FAO, 2005). At the present time, one of the most significant constraints of banana production that causes significant economic losses are numerous environmental challenges, particularly with fungal and bacterial pathogens as well as pests and abiotic stresses (Agrios, 1997; Carlier et al., 2000; Marín et al., 2003). This situation is exacerbated by the limited diversity of cultivars. Moreover, traditional breeding strategies are problematic due to a low female fertility, sterility, ploidy, and poor seed set. As a result, classical genetics is difficult and limited, as well as being extremely time-consuming. Taken together, the difficulty in conventional breeding, limited genetic diversity and poorly controlled diseases all point to the necessity of developing alternative strategies for banana improvement. With the development of an efficient banana transformation system, molecular breeding appears to be the most promising strategy to generate disease-resistant banana. The challenge lies with identifying, isolating and characterizing genes that confer the desired disease resistance traits.

Up to now, more than 50 resistance (R) genes have been cloned and characterized from both mono- and dicotyledonous plants through map-based cloning, transposon tagging and genome homologous analysis (Joshi et al., 2010). Although these R genes were isolated from different plants or through resistance to different pathogens, their encoding amino acid sequences have a high degree of similarity. According to the amino acid sequences of the cloned R genes, they can be divided into the following four classes: Class 1 encodes cytoplasmic receptor-like proteins that contain a leucine-rich repeat (LRR) domain and a nucleotide-binding site (NBS). This class is the largest R gene class, such as *Xa1* (Yoshimura et al., 1998), *Pib* (Wang et al., 1999), *RPM1* (Grant et al., 1995), *L6* (Lawrence et al., 1995), *Prf* (Salmeron et al., 1996), *RPS2* (Bent et al., 1994), *N* (Whitham et al., 1994), *Bs2* (Tai et al., 1999), *M* (Anderson, 1997), *Dm3* (Thevissen et al., 2000), and *I2C-1* (Ori et al., 1997). Class 2 encodes putative transmembrane receptors with an LRR domain, including *cf2* (Dixon et al., 1996), *cf9* (Jones et al., 1994) and the *HS1* gene (Yamanashi et al., 1993). Class 3 encodes a serine-threonine kinase domain (Martín et al., 1993). Class 4, including *Xa2* (Song et al., 1995) and *Lr10* (Feuillet et al., 1997), encodes putative transmembrane receptors with a large extracytoplasmic LRR domain and an intracellular serine-threonine kinase domain. With the increase in R gene sequences in the database, homologous analysis has become a new strategy to unravel R genes (Michelmore, 2000). Although the overall sequence homology among different R genes is poor, short stretches of peptide sequences such as kinase-1a (MGGVGKTT) and domain 2 (GLPLAL) in the NBS-LRR class are highly conserved. Using the conserved motifs in the NBS-LRR class of R genes to design degenerate primers for amplifying similar

sequence fragments may facilitate positional cloning of new functional R genes (Aarts et al., 1998; Xie et al., 1998; Leister et al., 1999; Wang et al., 1999). In *Musa*, progress in resistance gene analog (RGA) characterization has been recently made, and there are only about 10 NBS-LRR disease resistance-like protein sequences currently deposited in GenBank. Most of them are non-TIR (Toll/interleukin-1 receptor-like domain) NBS-LRR genes, amplified in Gongjiao, Xinyiyejiao, Zhongshandajiao, Fenjiao, and Williams (Pei et al., 2007). In other groups, R genes include *cf* orthologs amplified in landrace Zebrina GF (Wiame et al., 2000), *pto* family RGAs amplified in *M. acuminata* (Peraza-Echeverria et al., 2007) and *M. ornate*, *M. schizocarpa*, *M. textilis*, and *M. velutina* (Robert et al., 2008).

Given that R gene sequences so far studied represent only a small fraction of all resistance genes in *Musa*, the objectives of this study were to identify NBS-LRR RGAs and to explore their diversity in *Musa*. We used degenerate oligonucleotide primers based on NBS and LRR domains of resistance genes, and obtained new RGAs from *Musa* by PCR amplification, as well as studying the sequences and diversity of the NBS-LRR class of RGAs obtained from *Musa* by comparison with known resistance gene sequences.

MATERIAL AND METHODS

Plant material and DNA extraction

Musa acuminata subsp *malaccensis* and the *Musa* AAB Group were propagated *in vitro* and plants maintained in a greenhouse.

Total DNA was isolated by a modified cetyl-trimethyl ammonium bromide (CTAB) extraction technique. Approximately 0.2 g fresh leaf tissue was ground in liquid nitrogen and added to 700 µL extraction buffer (2% CTAB, 100 mM Tris-HCl, 20 mM EDTA, 1.4 M NaCl, 0.2% β-mercaptoethanol, pH 8) and incubated at 60°C for 30 min. The aqueous solution was extracted with 600 µL chloroform:isoamyl alcohol (24:1) and centrifuged for 20 min at 12,000 rpm. The extraction was repeated and 1 mL 96% cold ethanol was added to the aqueous phase and left at -20°C for 20 min to precipitate the nucleic acid. The samples were then centrifuged for 10 min at 12,000 rpm. The precipitated nucleic acid was washed with 1 mL 70% ethanol, left to air dry, rehydrated in 100 µL TE buffer and digested with 1 µL RNase (10 mg/mL; Fermentas, Germany) for 45 min at 37°C. The quality of genomic DNA was determined by agarose (1%) gel electrophoresis and quantified spectrophotometrically at 260 and 280 nm. Each sample was diluted to 50 ng/µL in TE buffer and stored at -20°C.

Degenerate primer sets

Six degenerate primers were designed in this study (Table 1). LF1, LF2 and LR1B were designed based on a conserved motif in non-TIR NBS-LRR domain-containing monocotyledon sequences obtained from GenBank. Degenerate primer LR3B targeted the non-TIR NBS-LRR protein motif in dicotyledons, designed following the alignment of resistance proteins *Arabidopsis thaliana* PRS2 (gi:548086) and PRS1 (gi:963017). Primer p-loop/kinase-1a and Not1 targeting non-TIR NBS-LRR R proteins were designed from motifs (Table 1) present in several dicotyledonous plant RGAs (*A. thaliana*, *Solanum lycopersicon*, *Nicotiana glutinosa*).

PCR amplification

All reactions were carried out in a final volume of 25 μL containing 5 μL 10X PCR buffer, 0.2 mM dNTPs, 1.5 mM MgCl$_2$, 0.5 mM each primer, 1.0 U Taq polymerase, and 1.0 μL genomic DNA from banana per reaction. The final volume was adjusted to 25 μL with sterile Milli-Q water. Amplification was performed in thin-walled PCR tubes.

PCR amplification was performed using a PTC200 DNA thermal cycler (TM Research) with modification of the annealing temperature to optimize the reaction mixtures. Samples were denatured at 94°C for 5 min and subjected to 35 cycles of 1 min at 94°C, 1 min at 45°C, 1.5 min at 72°C, and a final extension step of 10 min at 72°C and eventually stored at 4°C.

Amplicons were electrophoresed on 1.5% agarose gels with 1X TAE buffer at 5 V/cm. Band patterns were photographed under UV light (302 nm) after staining with ethidium bromide. The size of amplified fragments was estimated by comparison to DNA marker DL2000 (Takara, Biotechnology Co. Ltd., Danlian, China).

Cloning of PCR products and sequencing

Following electrophoresis, amplicons of expected size were purified using Advantage PCR-Pure kit (Clontech Laboratories, Inc.). Products were ligated into the pGEM-Teasy Vector Systems (Promega) and then transferred into DH-5α competent cells via a standard heat shock protocol. Recombinant plasmid clones were selected and manipulated following standard protocols (Sambrook and Russell, 2001). We typically sequenced one to three clones from each fragment. Sequences were submitted to GenBank with accession numbers HQ199833-HQ199836.

Sequencing analysis

Sequences were processed to remove vector and poor-quality sequences using the Staden sequence analysis software package (Staden, 1996). Contig assembly was performed by manual editing. RGAs were identified on the basis of sequence similarity using the BLASTX program (Altschul et al., 1997). Further confirmation was obtained by checking for the Pfam NB-ARC domain (van Der Biezen and Jones, 1998), which is a protein domain characteristic of plant resistance genes, using the program HMMER (Eddy, 2007). Only intact reading frames between the NBS domain kinase 2 and GLPL motifs (a common sequence to all generated contigs) were retained. Amino acid sequences from resistance genes from other plant species were added to the set of NBS sequences, and cluster analysis was carried out using the CLUSTALX package based on the neighbor-joining method (Saitou and Nei, 1987) and drawn by TREEVIEW (Page, 1996). The reliability of tree topologies was tested by bootstrapping 1000 times, with construction of a final majority rule consensus tree.

RESULTS

Degenerate primer design

Public databases at present contain only very limited numbers of *Musa* RGA sequences.

To enrich the fraction of RGA candidates in *Musa* recoverable by PCR, we designed degenerate primers from monocotyledon sequences and targeted NBS and additional domains. We began with HMMER-based selection of monocotyledon sequences from GenBank, containing a characteristic domain shared by R genes. Following removal of redundant sequences, about 200 RGA candidates were obtained. Based on this subset, a search for conserved sequence motifs was conducted using the MEME program (Bailey and Elkan, 1994). NBS family motifs were observed across the sequence, as well as novel conserved motifs outside the NBS domain. Most of those identified served as candidates for degenerate primer design, with an additional constraint imposed, whereby motifs had to be present in at least 25% of sequences.

PCR amplification of targeted RGA fragment

Primer combinations 1-5 targeted conserved amino acid motifs in non-TIR NBS-LRR sequences in monocotyledon and dicotyledon plants (Table 1). Combination LF1/LR1B, LF2/LR3B and p-loop/kinase-la/Not1 did not produce PCR amplification products. By contrast, primer combinations LF1/LR3B and LF2/LR1B consistently yielded products of approximately 650 bp. Four of 12 sequence fragments showed similarity to NBS sequences and were named *MRGL1*, *MRGL2*, *MRGL3*, and *MRGL4*. These fragments are deposited in GenBank and their accession numbers are HQ199833-HQ199836.

Table 1. Degenerate primer sequences and target motifs used for resistance gene analog isolate in *Musa*.

Degenerate primer	Target motif name/plant origin	Primer sequence (5' to 3')
LF1 (forward)	p-loop-GKTT/monocotyledon	GGCGGGGTGGGCaaracnacnht
LF2 (forward)	Kinase 2/monocotyledon	GAGGTACTTCCTGGTGCTGgaygayrtbtgg
LR3B (reverse)	GLPL/dicotyledon	AIITYIRIIRYIAGIGGIAGICC
LR1B (reverse)	LRR/monocotyledon	CGGCCAAGTCGTGCAyvakrtcrtgca
p-loop/kinase-la (forward)	GVGKTT/dicotyledon	GGIGTIGGIAARACIAC
Not1 (reverse)	CFLYCALFPED/dicotyledon	CYTCIGGRAAIARIGCRCARTAIARRAARC

Degenerate code: I = inosine; R = A/G; Y = C/T; M = A/C; K = G/T; W = A/T; S = C/G; B = C/G/T; D = A/G/T; H = A/C/T; V = A/C/G; N = A/C/G/T. LRR = leucine-rich repeat.

Most sequences that were not RGAs showed similarity to retroelements. These can constitute a large fraction of the plant genome (SanMiguel et al., 1996) and many R gene loci have been reported to contain interspersed transposable elements (Song et al., 1995; Noel et al., 1999).

BLASTp for identical protein sequences

The nucleotide sequences of the cloned fragments were translated into amino acid sequences using the ExPASy (Expert Protein Analysis System) Translate Tool and the amino acid sequences of RGAs were compared with protein sequences deposited in GenBank using the BLASTp algorithm (Table 2). The NBS analog *MRGL1* showed 97% identity (E value = 0.0) with resistance protein gene of *M. acuminata* subsp *malaccensis* (GI: EU163981). *MRGL2* showed 96% identity ((E value = 0.0) with resistance gene candidate NBS-LRR protein of *Musa* AAB and *Musa* ABB Groups, followed by 96% identity (E value = 0.0) with *Musa* ABB Group NBS resistance protein (RGA-I) gene. MRGL3 showed 68% identity (E value = 2e-15) with *Phaseolus vulgaris* NBS-LRR-type putative disease resistance protein gene and 97%

identity (E value = 0.0) with *M. acuminata* subsp *malaccensis* NBS-type resistance protein RGC5. *MRGL4* showed 93% (E value = 6e-123) and 67% (E value = 2e-14) with *Musa* ABB Group NBS resistance protein and *Elaeis guineensis* resistance protein, respectively.

Table 2. Results of the search for similarity between *Musa* NBS sequences and GenBank accessions carried out using the BLASTp algorithm.

Musa NBS	GenBank protein accession showing the highest similarity	GenBank ID	Amino acid identity (%)	E value
MRGL1	*Musa acuminata* subsp *malaccensis* resistance protein gene	EU163981	97	0.0
MRGL2	*Musa* AAB Group partial nbs gene for NBS-LRR disease resistance protein, cultivar Pisang-Mutiara	AM931335	96	0.0
	Musa ABB Group NBS resistance protein (RGA-I) gene	EU855839	96	0.0
	Musa ABB Group partial NBS gene for NBS-LRR disease resistance protein, cultivar Pelipita	AM931366	68	4e-36
MRGL3	*Phaseolus vulgaris* NBS-LRR-type putative disease resistance protein CNL-B16 gene	EU856772	68	2e-15
	Musa acuminata subsp *malaccensis* NBS-type resistance protein RGC5 (RGC5) mRNA	EU616674	97	0.0
MRGL4	*Musa* ABB Group NBS resistance protein (RGA-H) gene, partial cds	EU855838	93	6e-123
	Elaeis guineensis resistance protein (RGA4)	AF197924	67	2e-14

E value refers to the number of matches expected by chance. The lower the E value, the greater the similarity is and the stronger match is. NBS = nucleotide-binding site; LRR = leucine-rich repeat.

Sequence analysis for conserved motif

The amino acid sequences of RGAs cloned in the present investigation were compared with other known R genes, *N*, *RPS2*, *RPM1* of *A. thaliana* in the NBS region using the CLUSTALX multiple alignment program (Figure 1). The various motifs of NBS, i.e., kinase-1a, kinase-2 and kinase-3a, were conserved in the four selected *Musa* RGAs. The hydrophobic region represented by the GLPL domain was also observed in all four RGA sequences. High sequence homology of the RGAs with NBS motifs of *N*, *RPS2* and *RPM1* clearly establishes that the RGAs cloned in the present investigation belong to the NBS-LRR class of R genes. Four *Musa* NBS analogs also have a conserved tryptophan residue at the end of the kinase-2 domain, further confirming their similarity to the non-TIR subfamily of the NBS-LRR class of R genes.

Figure 1. Multiple alignment of deduced amino acid sequence of resistance gene analogs with known NBS class R genes *RPS2* (GenBank accession No. U12860), *RPM1* (GenBank accession No. NM_111584) and *N* (GenBank accession No. U15605), using CLUSTALX. Motif identified are given at the top of the panel.

Phylogenetic analysis

Phylogenetic analysis of *Musa* RGA sequences with already characterized R genes was done by constructing a BLOSUM62 average distance tree in the Jalview Java editor of CLUSTALX. All four *Musa* RGA were clustered together with well-characterized non-TIR-NBS-LRR genes *RPS2* (U12860), *XA1* (AB002266), *RPP8* (AF089710), while the TIR-NBS-LRR genes *N* (U15605) and *L6* (U27081) were grouped as a separate cluster (Figure 2). Another average distance tree was constructed for the RGAs cloned in this study and those isolated from other plants.
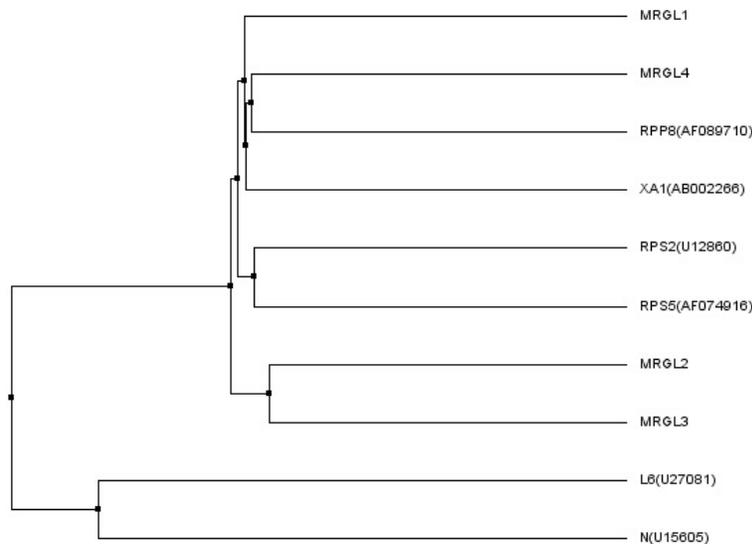

Figure 2. Average distance tree using % identity showing relationship of *Musa* NBS analogs with NBS-LRR class of R genes. For abbreviations, see legend to Table 3.

Table 3. Distribution of NBS sequences by group based on phylogenetic analysis.

TIR-NBS-LRR	
Arabidopsis thaliana	*RPP1*, At.23854, NM001084768, At.22810, NM105280, At.28151, At.27102, At.30211, Ak229329, At.29667, At.29956, At.31047, *RPP5*
Glycine max	AF175395, AF175399, AY308656, AF502079
Helianthus annuus	AY154299
Nicotiana glutinosa	*N*
Solanum tuberosum	AY196151
Non-TIR-NBS-LRR	
Avena sativa	AAC31552, AAC31553
Musa	*MRGL1*, *MRGL2*, *MRGL3*, *MRGL4*
Phaseolus vulgaris	AAK50042
Solanum tuberosum	*Gpa2*, ABO93001
Arabidopsis thaliana	AY057522, *RPM1*, *RFL1*, *RPS2*, *RPS5*, *RPP8*
Oryza sativa	AK110307, AAL79344, *XA1*
Hordeum vulgare	AJ507089
Triticum aestivum	*Lr1*

TIR = Toll/interleukin-1 receptor-like domain; NBS = nucleotide-binding site; LRR = leucine-rich repeat.

Forty-one full- and partial-NBS disease resistance domains were detected in the current GenBank database and shown in Table 3. Two major branches designated as TIR-NBS-LRR and non-TIR-NBS-LRR were revealed. A further significant subdivision can be noticed in Figure 3. The internal stability of these two major branches is strongly supported by the bootstrapping experiment (99% for the two branches). Branch length between group members is generally shorter in non-TIR-NBS-LRR than TIR-NBS-LRR, indicating accelerated diversification in the TIR-NBS-LRR group. *MRGL1*, *MRGL2*, *MRGL3*, and *MRGL4* were grouped in non-TIR-NBS-LRR (Figure 3).
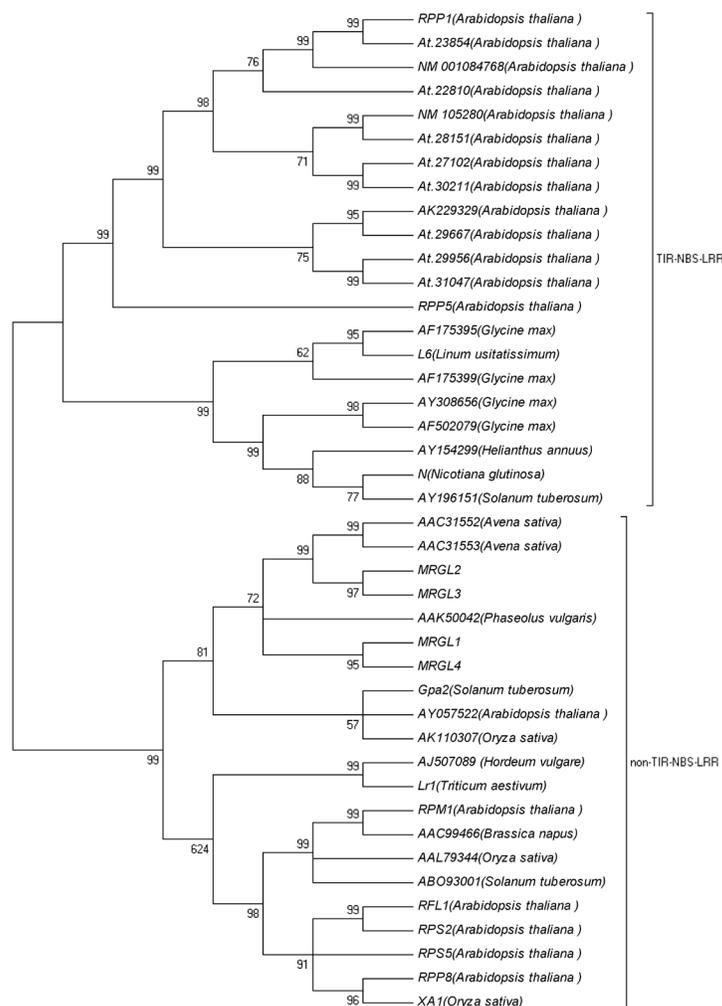

Figure 3. Phylogenetic tree for plant NBS-LRR-type resistance genes and their homologues based on amino acid sequences in the NBS region. This tree was constructed by the neighbor-joining method (Saitou and Nei, 1987) as implements in the MEGA 4.1 software. The length of each branch is proportional to the average substitutions per site as indicated by the scale. Numbers on the branches represent the percentage of bootstrap replications (1000), which support that particular branch point (only number above 50% are shown here). For abbreviations, see legend to Table 3.

DISCUSSION

The PCR approach designed for RGA discovery in monocotyledon species was effective in *Musa*. Four cloned RGAs belonged to the non-TIR-NBS-LRR subfamily, with considerable divergence observed at the amino acid level (Figure 3). RGAs could be obtained with two primer sets targeting universal non-TIR NBS motifs. A number of factors may have contributed to the success rate of the primers. Our design strategy for monocotyledons took into account the number of degeneracies, primer length, nucleotide composition, degeneracy position within each primer, and prevalence of putative targets in the sequences analyzed. Universal primer combinations designed for both TIR and non-TIR NBS motifs in dicot sequences were relatively inefficient, with a maximum of 29% of sequences homologous to RGAs when amplified with the primer combination (LF2/LR1B). Amplification was most efficient using non-TIR targeting primers, where both NBS and LRR domains have only been described in plant resistance genes so far. Given that the primer combination (LF2/LR1B) produced amplicons from the NBS kinase 2 to a conserved motif within the LRR domain, efficiency in amplification of targets involved in disease resistance is therefore potentially greater. We also found that numerous pseudogenes were also co-amplified. These likely arise through point mutations, insertions or nucleotide deletions, acting as reservoirs for variation and offering the potential for recombination or gene conversion between R gene alleles or paralogs (Michelmore and Meyers, 1998).

Multiple-sequence alignment of the *Musa* RGAs and known R genes showed significant homology to the kinase-1a and hydrophic GLPL motifs of the NBS domain of the known genes. All four *Musa* RGAs were also classified into the non-TIR-NBS-LRR subfamily of the known genes. The NBS-LRR genes are usually grouped into two different subfamilies (Meyers et al., 1999; Pan et al., 2000): subfamily I contains the TIR element and has been found only in dicots, while subfamily II lacks the TIR domain and has been found in both dicots and monocots. The partial sequence of the NBS portion is usually sufficient to assign a given gene to either subfamily I or II. The last residue of the kinase-2 domain can be used to predict with 95% accuracy whether an RGA belongs to the TIR-NBS or to the non-TIR-NBS family; conservation of tryptophan (W) at this location is tightly linked to non-TIR R genes (*RPS2*, *RPS5* and *RPP8* of *A. thaliana*), whereas conservation of aspartic acid (D) or its uncharged derivative aspartate (N) is characteristic of the TIR class of R genes (*N* and *L6*) (Meyers et al., 1999; Pan et al., 2000; Jeong et al., 2001; Peñuela et al., 2002). This clearly supports the hypothesis that *Musa* has resistance genes of the non-TIR-NBS-LRR class.

The cluster analysis of the NBS analogs identified their group based on similarity. The phylogenetic tree based on neighbor joining using % identity of the deduced amino acid sequence of *Musa* NBS analogs identified their relatedness with other non-TIR-NBS-LRR genes *RPS2*, *XA1*, while TIR-NBS-LRR genes *N* and *L6* were placed in the other group. Another distance tree was made relating the *Musa* NBS analogs and RGAs from other plants. *MRGL1*, *MRGL2*, *MRGL3*, and *MRGL4* were grouped to non-TIR-NBS-LRR, in agreement with the hypothesis that the TIR subfamily is restricted to dicotyledonous taxa (Pan et al., 2000). The existence of the TIR motif has now been reported in the rice genome, albeit in reduced numbers (Bai et al., 2002; Meyers et al., 2002). Lack of detection in the *Musa* monocotyledon genome may therefore reflect limitations in PCR amplification.

ACKNOWLEDGMENTS

REFERENCES

Aarts MG, te Lintel HB, Holub EB, Beynon JL, et al. (1998). Identification of R-gene homologous DNA fragments genetically linked to disease resistance loci in *Arabidopsis thaliana*. *Mol. Plant Microbe Interact.* 11: 251-258.

Agrios GN (1997). Plant Pathology. 4th edn. Academic Press, New York.

Altschul SF, Madden TL, Schaffer AA, Zhang J, et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389-3402.

Anderson PA, Lawrence GJ, Morrish BC, Ayliffe MA, et al. (1997). Inactivation of the flax rust resistance gene M associated with loss of a repeated unit within the leucine-rich repeat coding region. *Plant Cell* 9: 641-651.

Bai J, Pennill LA, Ning J, Lee SW, et al. (2002). Diversity in nucleotide binding site-leucine-rich repeat genes in cereals. *Genome Res.* 12: 1871-1884.

Bailey TL and Elkan C (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2: 28-36.

Bent AF, Kunkel BN, Dahlbeck D, Brown KL, et al. (1994). RPS2 of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. *Science* 265: 1856-1860.

Carlier J, Fouré E, Gauhl F and Jones DR (2000). Fungal Diseases of the Foliage. In: Diseases of Banana Abacá and Enset (Jones DR, ed.). CABI Publishing, Wallingford, 37-141.

Dixon MS, Jones DA, Keddie JS, Thomas CM, et al. (1996). The tomato Cf-2 disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. *Cell* 84: 451-459.

Eddy SR (2007). HMMER: Profile Hidden Markov Models for Biological Sequence Analysis. Available at [http://hmmer.janelia.org]. Accessed March 15, 2010.

FAO (2005). (Food and Agricultural Organization). Available at [http://www.fao.org/lim500/nphwrap.pi?productioncrops.primary&Domain=SUA&]. Accessed December 3, 2009.

Feuillet C, Schachermayr G and Keller B (1997). Molecular cloning of a new receptor-like kinase gene encoded at the Lr10 disease resistance locus of wheat. *Plant J.* 11: 45-52.

Grant MR, Godiard L, Straube E, Ashfield T, et al. (1995). Structure of the *Arabidopsis* RPM1 gene enabling dual specificity disease resistance. *Science* 269: 843-846.

Jeong SC, Hayes AJ, Biyashev RM and Saghai MMA (2001). Diversity and evolution of a non-TIR-NBS sequence family that clusters to a chromosomal "hotspot" for disease resistance genes in soybean. *Theor. Appl. Genet.* 103: 406-414.

Jones DA, Thomas CM, Hammond-Kosack KE, Balint-Kurti PJ, et al. (1994). Isolation of the tomato Cf-9 gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* 266: 789-793.

Jones DR (2000). Introduction of Banana. In: Diseases of Banana, Abacá and Enset (Jones DR, ed.). CABI Publishing, Wallingford, 1-36.

Joshi RK, Mohanty S, Subudhi E and Nayak S (2010). Isolation and characterization of NBS-LRR- resistance gene candidates in turmeric (*Curcuma longa* cv. surama). *Genet. Mol. Res.* 9: 1796-1806.

Lawrence GJ, Finnegan EJ, Ayliffe MA and Ellis JG (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.

Leister D, Kurth J, Laurie DA, Yano M, et al. (1999). RFLP- and physical mapping of resistance gene homologues in rice (*O. sative*) and Barley (*H. vulgare*). *Theor. Appl. Genet.* 98: 509-520.

Marín DH, Romero RA, Guzmán M and Sutton TB (2003). Black sigatoka: an increasing threat to banana cultivation. *Plant Disease* 87: 208-222.

Martín GB, Brommonschenkel SH, Chunwongse J, Frary A, et al. (1993). Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* 262: 1432-1436.

Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, et al. (1999). Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant J.* 20: 317-332.

Meyers BC, Morgante M and Michelmore RW (2002). TIR-X and TIR-NBS proteins: two new families related to disease resistance TIR-NBS-LRR proteins encoded in *Arabidopsis* and other plant genomes. *Plant J.* 32: 77-92.

Michelmore R (2000). Genomic approaches to plant disease resistance. *Curr. Opin. Plant Biol.* 3: 125-131.

Michelmore RW and Meyers BC (1998). Clusters of resistance genes in plants evolve by divergent selection and a birth-

and-death process. *Genome Res.* 8: 1113-1130.

Noel L, Moores TL, van Der Biezen EA, Parniske M, et al. (1999). Pronounced intraspecific haplotype divergence at the RPP5 complex disease resistance locus of *Arabidopsis*. *Plant Cell* 11: 2099-2112.

Ori N, Eshed Y, Paran I, Presting G, et al. (1997). The I2C family from the wilt disease resistance locus I2 belongs to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes. *Plant Cell* 9: 521-532.

Page RD (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12: 357-358.

Pan Q, Wendel J and Fluhr R (2000). Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. *J. Mol. Evol.* 50: 203-213.

Pei X, Li S, Jiang Y, Zhang Y, et al. (2007). Isolation, characterization and phylogenetic analysis of the resistance gene analogues (RGAs) in banana (*Musa* spp.). *Plant Sci.* 172: 1166-1174.

Penuela S, Danesh D and Young ND (2002). Targeted isolation, sequence analysis, and physical mapping of nonTIR NBS-LRR genes in soybean. *Theor. Appl. Genet.* 104: 261-272.

Peraza-Echeverria S, James-Kay A, Canto-Canche B and Castillo-Castro E (2007). Structural and phylogenetic analysis of Pto-type disease resistance gene candidates in banana. *Mol. Genet. Genom.* 278: 443-453.

Robert NGM, David JB, Franc CB, Candice MRS, et al. (2008). Analysis of non-TIR NBS-LRR resistance gene analogs in *Musa acaminata* Colla: isolation, RFLP marker development, and physical mapping. *BMC Plant Biol.* 8: 15.

Saitou N and Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406-425.

Salmeron JM, Oldroyd GE, Rommens CM, Scofield SR, et al. (1996). Tomato Prf is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the Pto kinase gene cluster. *Cell* 86: 123-133.

Sambrook J and Russell DW (2001). Molecular Cloning: A Laboratory Manual. 3rd edn. Cold Spring Harbor, USA, Cold Spring Harbor Laboratory, New York.

SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, et al. (1996). Nested retrotransposons in the intergenic regions of the maize genome. *Science* 274: 765-768.

Song WY, Wang GL, Chen LL, Kim HS, et al. (1995). A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. *Science* 270: 1804-1806.

Staden R (1996). The Staden sequence analysis package. *Mol. Biotechnol.* 5: 233-241.

Tai TH, Dahlbeck D, Clark ET, Gajiwala P, et al. (1999). Expression of the Bs2 pepper gene confers resistance to bacterial spot disease in tomato. *Proc. Natl. Acad. Sci. U. S. A.* 96: 14153-14158.

Thevissen K, Cammue BP, Lemaire K, Winderickx J, et al. (2000). A gene encoding a sphingolipid biosynthesis enzyme determines the sensitivity of *Saccharomyces cerevisiae* to an antifungal plant defensin from dahlia (*Dahlia merckii*). *Proc. Natl. Acad. Sci. U. S. A.* 97: 9531-9536.

van Der Biezen EA and Jones JD (1998). The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Curr. Biol.* 8: R226-R227.

Wang ZX, Yano M, Yamanouchi U, Iwamoto M, et al. (1999). The Pib gene for rice blast resistance belongs to the nucleotide binding and leucine-rich repeat class of plant disease resistance genes. *Plant J.* 19: 55-64.

Whitham S, Dinesh-Kumar SP, Choi D, Hehl R, et al. (1994). The product of the tobacco mosaic virus resistance gene N: similarity to toll and the interleukin-1 receptor. *Cell* 78: 1101-1115.

Wiame L, Swennen R and Sági L (2000). PCR-based cloning of candidate disease resistance genes from banana (*Musa acuminata*). *Acta Hortic.* 521: 51-57.

Xie YB, Tang DZ, Zhang YS and Li WM (1998). Isolation of homologous sequences of R gene from rice. *Chin. Sci. Bull.* 43: 277-281.

Yamanashi Y, Okada M, Semba T, Yamori T, et al. (1993). Identification of HS1 protein as a major substrate of protein-tyrosine kinase(s) upon B-cell antigen receptor-mediated signaling. *Proc. Natl. Acad. Sci. U. S. A.* 90: 3631-3635.

Yoshimura S, Yamanouchi U, Katayose Y, Toki S, et al. (1998). Expression of *Xa1*, a bacterial blight-resistance gene in rice, is induced by bacterial inoculation. *Proc. Natl. Acad. Sci.* 95: 1663-1668.