CHAPTER THREE

COMPARATIVE ANALYSIS OF TWO BANANA NPR1-LIKE CODING SEQUENCES

3.1 Abstract

The *Non-expressor of pathogenesis-related*1 (*NPR1*) gene is known to exist in more that one copy in most plants like *Arabidopsis*, rice, soybean, and banana. In Banana, three *NPR1*-like genes have already been isolated. The *Musa* (*M*)*NPR1A* and *MNPR1B* genes were isolated from a Cavendish banana with an AAA genome (Grand Naine) while *MdNPR1* was isolated from a banana cultivar with a ABB genome (Dongguan dijiao). Characterization of the expression pattern of the two *MNPR1* genes in various banana cultivars following treatment with *Fusarium* and elicitors like salicylic acid depicts a differential mode of activation of these genes. However, what controls this differential response is largely unknown and the role expression of *MNPR1* genes contributie to pathogen defense, if at all, has also not been reported. To complicate this picture further, characterization of expression of the genes following treatment of banana with the hemibiotroph *Xanthomonas campestris* further unveiled a sequential expression pattern for these genes which correlated with *pathogenesis-related* (*PR*)*-1* and *PR-3*.transcription. Hence, using various sequence analysis tools, it is reported here that these banana cultivars share different levels of identities with 39 previously described NPR1-like sequences, some of which are well-known to positively mediate pathogen defense response. These MNPR1 coding sequences further harbor functional conserved motifs, in some cases with some amino acid dissimilarities within the motifs relative to known sequences. Phylogenetic analysis revealed that the banana NPR1 sequences are closely related to those of other monocotyledons like rice and *Hordeum vulgare*. *Cis*-analysis of the two AAA banana NPR1 sequences showed that they both contain various regulatory elements reported to be sensitive to pathogens and elicitors.

3.2 Introduction

Since its isolation in 1994, the gene encoding the Non expressor of pathogenesis related1 (NPR1) protein has been isolated in many monocotyledons and dicotyledons to date. In most of these plants like *Arabidopsis*, rice, soybean, grapevine, *Brassica* and banana, more that one copy of the gene has been reported (Hepworth *et al*., 2005; Yuan *et al*., 2007; Endah *et al*., 2008; Zhao *et al*., 2009; Sandhu *et al*., 2009). In banana, two copies of the *NPR1*-like genes were isolated from a Cavendish banana cultivar with an AAA genome (*MNPR1A* and *MNPR1B*; Endah *et al*., 2008) and one from a banana cultivar with an ABB genome (*MdNPR1*; Zhao *et al*., 2009).

Studies on the characterization of expression of the two *MNPR1* genes in a *Fusarium*-tolerant Cavendish banana cultivar (GCTCV-218) and a susceptible cultivar (Grand Naine) indicates that these genes respond differentially to salicylic acid (SA) treatment and to a necrotrophic pathogen *Fusarium oxysporum forma specialis* (f. Sp.) *cubense* Race 4 (*Foc*; Endah *et al*., 2008). *Foc* elicits predominantly a jasmonic acid- (JA) and or ethylene- (ET) dependent defense response pathway in plants (Oliver and Ipcho, 2004; Glazebrook, 2005; Spoel *et al*., 2007). Further, these *MNPR1* genes were sequentially expressed in response to the banana hemi-biotrophic pathogen *Xanthomonas campestris* (Endah *et al*., 2010), which, leads to increased concentrations of both SA and JA in infested plant (Martinez *et al*., 2000; Delannoy *et al*., 2005). However, the control of this differential response and that of expression of *MNPR1* genes in pathogen defense is still largely unknown. It was observed, however, that expression of these genes correlated with expression of two *pathogenesis-related* (*PR-1* and *PR-3*) genes used during these studies (Endah *et al*., 2008; Endah *et al*., 2010) hence implicating them as possible defense genes.

Preliminary sequence analysis of previously reported NPR1-like sequences has been carried out in many studies using selected sequences from the National Center for Biotechnology information (NCBI) database. In 2008 analysis done by Endah and colleagues revealed that the MNPR1 amino acid sequences share a 78% sequence identity to each other when compared with eight other NPR1 sequences. They further grouped closed with the monocotyledon NPR1 homologue (NH1 or *Os*NPR1) from rice known to mediate plant defense responses (Yuan *et al*., 2007). In 2010, a more comprehensive sequence alignment study was done by Begeault *et al*. (2010) using 31 NPR1-like sequences pooled from the database From their study, the two banana MNPR1 sequences still grouped in the same clade as the rice *Os*NPR1 sequence. What is missing from both studies, however, is how the two MNPR1 sequences compare with the third banana ABB NPR1 sequence. Moreover, more than 31 NPR1-like sequences currently exist in the NCBI database and a more detailed comparison is essential for an in-depth understanding of the *MNPR1* genes and how their structure relates to their function. Additionally, information is also lacking on the type and representation of *cis*-regulatory elements present within the *MNPR1* coding region.

The objective of this part of the study was to therefore understand at the sequence level whether these two *MNPR1* genes belong to the same group as other known NPR1-like sequences when more 39 NPR1-like sequences are compared. A further objective was to identify recently published conserved functional motifs and defense *cis*-regulatory elements present in the MNPR1 coding sequences.

3.3 Materials and methods

3.3.1 Alignment of NPR1-like sequences and identification of conserved regions

Using the *Arabidopsis* NPR1 coding sequence as a reference (At1g64280), 41 NPR1-like coding sequences from different plant species were obtained from the NCBI database (Table 3.1). These sequences were edited in BioEdit v7.0.0 to obtain the correct reading frame and aligned using the ClustalW software program (Thompson *et al*., 1994). The percentage identity among sequences was provided directly by ClustalW while conserved regions were identified using published data from well-characterized NPR1 like sequences.

3.3.2 Phylogenetic analysis

The full length amino acid coding region of the *Musa* NPR1-like sequences (Table 3.1) were used to establish a phylogenetic relationship with the other 39 *NPR1* sequences. All sequences were aligned using the multiple sequence alignment software (MAFFT version 5) program (Katoh *et al*., 2005). Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 5 for Neighbour Joining (NJ) following the software instructions (Tamura *et al*., 2007) and the bootstrap was set to 1000 replicates.

Table 3.1 NPR1-like amino acid sequences used for alignment.

Organism	Sequence name	Genbank reference
Musa acuminate (AAA)	*M*NPR1A	ABI93182.1
M. acuminate (AAA)	*M*NPR1B	ABL63913.1
M. spp. ABB group	*Md*NPR1	ACE86413.1
Arabidopsis. Thaliana	*At*NPR1	ABR46023.1
A. thaliana	*At*NPR2	NP_194342.1
A. thaliana	*At*NPR3	NP_199324.2
A. thaliana	AtNPR4	NP_193701.2
A. thaliana	AtNPR5	ABH04470.1
A. thaliana	AtNPR6	AAU90063.1
A. lyrata	*Al*NPR1	ABR46032.1
Oryza sativa	*Os*NPR1	ABE11614.1
O. sativa	*Os*NPR2	ABE11616.1
O. sativa	*Os*NPR3	ABE11618.1
O. sativa	*Os*NPR4	ABE11620.1
O. sativa	*Os*NPR5	ABE11622.1
Brassica napus	*Bn*NPR1	AAM88865.2
B. juncea	*Bj*NPR1	ABC94642.2
Beta vulgaris	*Bv*NIM1	AAT57640.1
Capsicum annuum	*Ca*NPR1	ABG38308.1
C. chinense	*Cc*NPR1	CAP12787.1
Carica papaya	*Cp*NPR1	AAS55117.1
Glycine max	*Gm*NPR1-1	ACJ45013.1
G. max	*Gm*NPR1-2	ACJ45015.1
Gossypium hirsutum	*Gh*NPR1	ABC54558.1
Helianthus annuus	*Ha*NIM1	AAT57642.1
Hordeum vulgare	*Hv*NPR1	CAJ19095.1
Ipomoea batatas	*Ip*NPR1	ABM64782.1
Malus x *domestica*	*M*xd NPR1	ACC77697.1
Nicotiana tabacum	*Nt*NIM1	AAT57641.1
N. tabacum	*Nt*NPR1	ABH04326.1
N. glutinosa	*Ng*NPR3	ABN45747.1
Populus trichocarpa	*Pt*NIM1	XP_002300863.1
P. trichocarpa	*Pt*NPR1	XP_002322351.1
P. trichocarpa	*Pt*NIM1b	ABF48718.1
Ricinus communis	*Rc*NPR1	XP_002514127.1
R. communis	*Rc*NPR1B	XP_002520549.1
Solanum lycopersicum	*Sl*NIM1	AAT57638.1
S. lycopersicum	*Sl*NIM2	AAT57639.1
Theobroma cacao	*Tc*NPR1	ADI24348.1
Vitis vinifera.	*Vv*NPR1.1	CAO65332.1
V. vinifera	*Vv*NPR1.2	CAN67078.1

3.3.3 Identification of defense-response *cis*-elements within the *MNPR1* coding region

To identify defense response *cis*-elements within the *MNPR1* coding sequences, the coding region of *MNPR1A* and *MNPR1B* were individually analyzed using the PLACE/Signal scan web-based program as instructed by the software manufacturer (http://www.dna.affrc.go.jp/PLACE/signalup.html; Higo *et al*., 1999). Each identified *cis*-element was then individually analyzed to identify elements known to mediate elicitor and/or pathogen responses in plants.

3.4 Results

3.4.1 NPR1 sequence analysis and identification of conserved regions

Of the 41 sequences aligned, the two banana MNPR1 (MNPR1A and MNPR1B) coding sequences were 78% identical to each other (Table 3.2) but only 42% and 44%, respectively identical to the ABB banana NPR1-like sequence (MdNPR1). The rice (*Os*)NPR1 sequence was 63% and 65% identical to MNPR1A and MNPR1B, respectively while *Arabidopsis* NPR1 was only 47% and 48% identical to the two MNPR1 sequences, respectively. *Capsicum annuum* had the highest identity to both MNPR1A (64%) and MNPR1B (68%) and the least identity was with the rice NH4 sequence which was only 4% identical to MNPR1A and 10% identical to MNPR1B.

Table 3.2 Percentage identity of 41 NPR1-like amino acid coding sequences

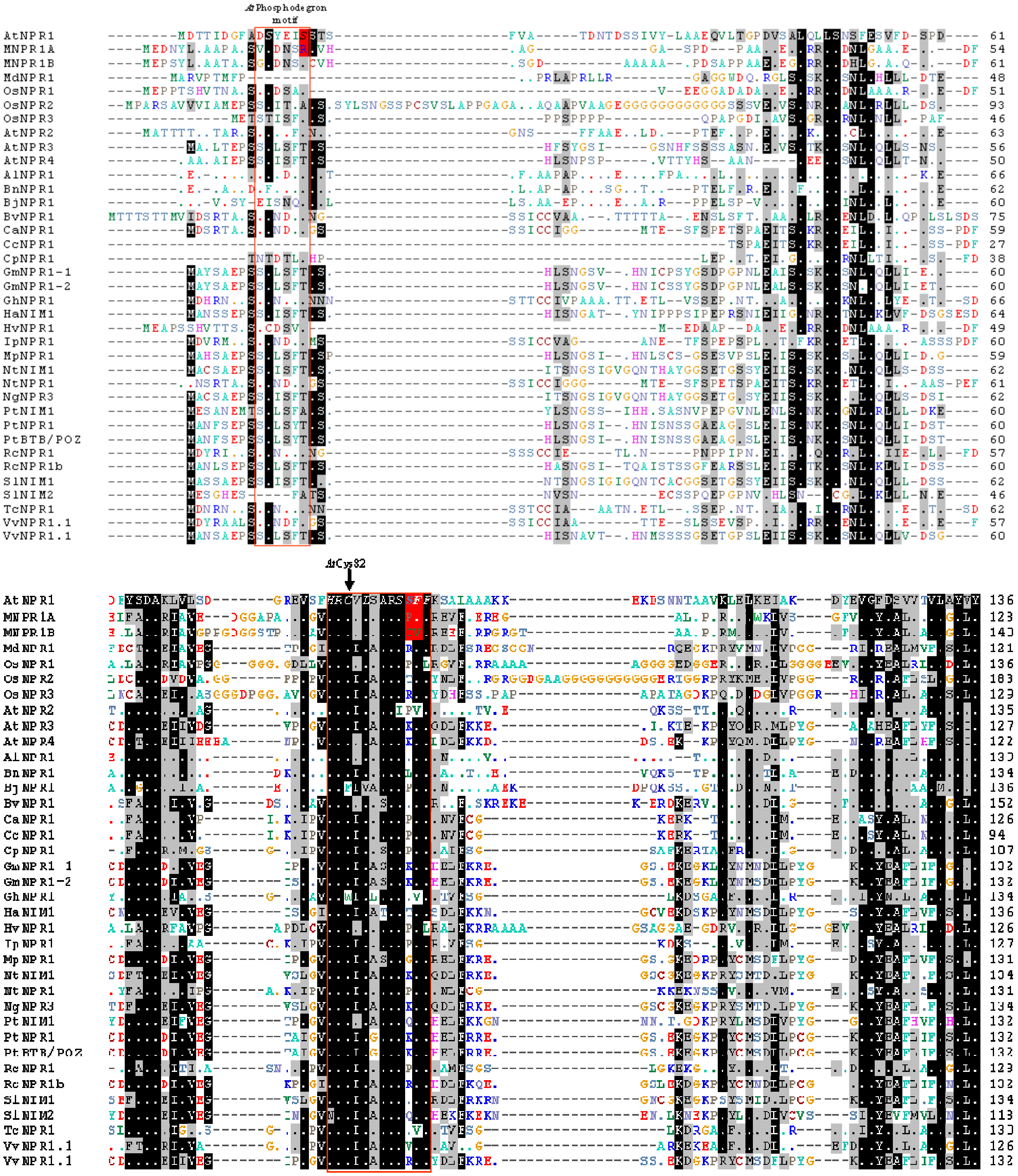
Name	MNPR1 A	MNPR1 B	MdNPR 1	AtNPR1	AtNPR 2	ATNPR 3	ATNPR 4	ATNPR 5	ATNPR 6	AlNPR1	NH1	NH2	NH3	NH4	NH5	BNNPR1	BjNPR1	BvNPR 1	CaNPR 1	CcNPR 1	CpNPR 1
MNPR1A																					
MNPR1B	78.0																				
MdNPR1	42.0	44.0																			
AtNPR1	47.0	48.0	34.0																		
AtNPR2	46.0	45.0	36.0	61.0																	
AtNPR3	39.0	40.0	52.0	35.0	36.0																
AtNPR4	40.0	40.0	52.0	36.0	37.0	71.0															
AtNPR5	22.0	24.0	25.0	21.0	20.0	25.0	25.0														
AtNPR6	20.0	21.0	26.0	20.0	18.0	24.0	24.0	83.0													
ALNPR1	49.0	49.0	36.0	89.0	60.0	36.0	37.0	22.0	20.0												
NH1	63.0	65.0	41.0	45.0	42.0	38.0	37.0	25.0	23.0	45.0											
NH2	43.0	43.0	65.0	36.0	34.0	49.0	49.0	26.0	23.0	37.0	42.0										
NH3	40.0	41.0	52.0	37.0	37.0	45.0	46.0	23.0	23.0	37.0	41.0	52.0									
NH4	4.0	10.0	8.0	9.0	7.0	7.0	8.0	9.0	9.0	9.0	13.0	10.0	11.0								
NH5	22.0	22.0	25.0	21.0	18.0	22.0	21.0	68.0	70.0	21.0	24.0	25.0	21.0	9.0							
BNNPR1	46.0	46.0	37.0	69.0	56.0	33.0	33.0	20.0	18.0	69.0	43.0	35.0	37.0	8.0	19.0						

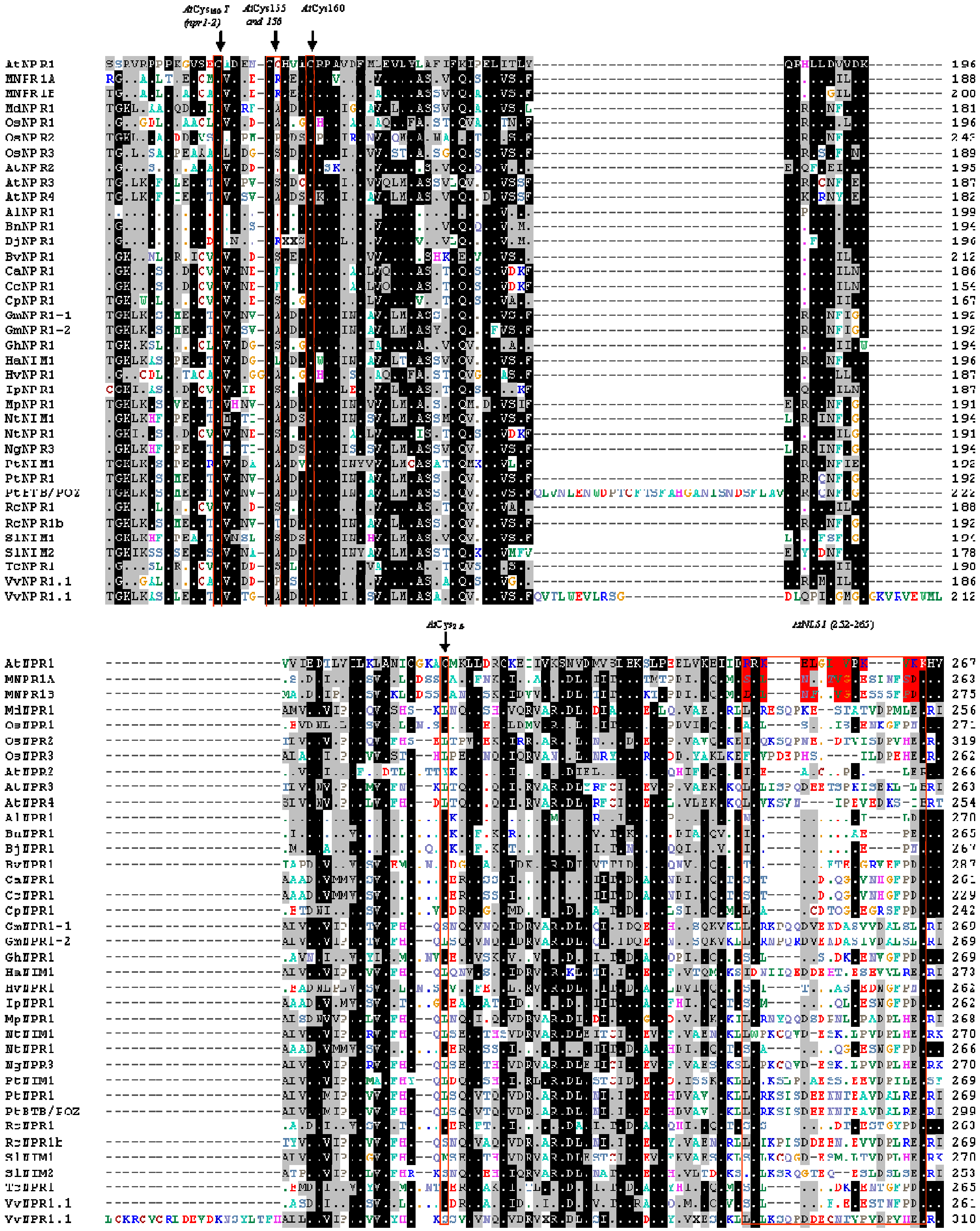
Name	MNPR1 A	MNPR1 B	MdNPR 1	AtNPR1	AtNPR 2	ATNPR 3	ATNPR 4	ATNPR 5	ATNPR 6	AlNPR1	NH1	NH2	NH3	NH4	NH5	BNNPR1	BjNPR1	BvNPR 1	CaNPR 1	CcNPR 1	CpNPR 1
BjNPR1	45.0	43.0	34.0	62.0	54.0	34.0	34.0	21.0	19.0	64.0	42.0	33.0	33.0	6.0	19.0	67.0					
BvNPR1	59.0	60.0	43.0	50.0	49.0	38.0	40.0	23.0	20.0	52.0	56.0	41.0	38.0	13.0	23.0	51.0	46.0				
CaNPR1	60.0	63.0	43.0	49.0	47.0	39.0	37.0	23.0	21.0	52.0	58.0	42.0	40.0	10.0	21.0	51.0	47.0	67.0			
CcNPR1	64.0	68.0	45.0	54.0	52.0	42.0	40.0	23.0	21.0	55.0	62.0	45.0	43.0	8.0	22.0	54.0	49.0	69.0	99.0		
CpNPR1	60.0	63.0	43.0	52.0	51.0	39.0	39.0	23.0	21.0	54.0	56.0	41.0	39.0	9.0	20.0	52.0	50.0	67.0	71.0	73.0	
GmNPR1-1	43.0	44.0	63.0	37.0	37.0	60.0	62.0	27.0	25.0	38.0	41.0	57.0	48.0	7.0	23.0	36.0	35.0	44.0	43.0	47.0	45.0
GmNPR1-2	43.0	43.0	63.0	36.0	36.0	59.0	62.0	26.0	26.0	38.0	41.0	58.0	47.0	6.0	24.0	35.0	34.0	44.0	42.0	46.0	45.0
GhNPR1	58.0	59.0	42.0	53.0	50.0	40.0	40.0	22.0	20.0	53.0	57.0	41.0	38.0	9.0	21.0	51.0	50.0	64.0	67.0	70.0	74.0
HaNIM1	40.0	38.0	58.0	36.0	36.0	54.0	53.0	23.0	22.0	37.0	39.0	53.0	47.0	7.0	22.0	34.0	34.0	42.0	41.0	45.0	40.0
HvNPR1	60.0	62.0	39.0	42.0	42.0	38.0	38.0	22.0	21.0	43.0	79.0	40.0	39.0	11.0	21.0	41.0	39.0	53.0	56.0	60.0	54.0
IpNPR1	57.0	59.0	43.0	48.0	48.0	38.0	39.0	24.0	21.0	51.0	56.0	41.0	39.0	9.0	22.0	50.0	46.0	66.0	80.0	81.0	66.0
MxdNPR1	41.0	42.0	62.0	36.0	36.0	59.0	57.0	25.0	26.0	38.0	39.0	55.0	46.0	9.0	25.0	36.0	34.0	40.0	41.0	45.0	43.0
NtNIM1	41.0	42.0	62.0	36.0	36.0	55.0	55.0	25.0	23.0	37.0	39.0	55.0	45.0	8.0	24.0	36.0	34.0	43.0	41.0	45.0	41.0
NtNPR1	58.0	62.0	43.0	50.0	49.0	39.0	39.0	23.0	21.0	52.0	55.0	42.0	40.0	8.0	22.0	48.0	45.0	65.0	92.0	91.0	69.0
NgNPR3	43.0	43.0	62.0	36.0	36.0	56.0	56.0	25.0	24.0	38.0	40.0	55.0	46.0	11.0	25.0	37.0	35.0	42.0	41.0	44.0	40.0
PtNIM1	44.0	43.0	66.0	36.0	37.0	60.0	59.0	27.0	25.0	38.0	42.0	60.0	48.0	9.0	25.0	34.0	33.0	43.0	43.0	47.0	44.0
PtNPR1	62.0	61.0	43.0	51.0	50.0	40.0	41.0	26.0	23.0	51.0	57.0	43.0	38.0	9.0	23.0	50.0	48.0	65.0	70.0	73.0	73.0
PtNPR1	62.0	61.0	43.0	51.0	50.0	40.0	41.0	26.0	23.0	51.0	57.0	43.0	38.0	9.0	23.0	50.0	48.0	65.0	70.0	73.0	73.0
PtNIM1b	40.0	39.0	59.0	32.0	32.0	50.0	50.0	27.0	27.0	34.0	35.0	53.0	45.0	8.0	24.0	32.0	34.0	40.0	38.0	41.0	39.0
RcNPR1	62.0	62.0	44.0	53.0	52.0	39.0	42.0	25.0	24.0	55.0	56.0	43.0	38.0	8.0	23.0	53.0	50.0	67.0	72.0	75.0	76.0
RcNPR1b	44.0	43.0	65.0	37.0	38.0	61.0	59.0	26.0	24.0	38.0	41.0	58.0	49.0	8.0	26.0	37.0	35.0	44.0	45.0	48.0	45.0
SlNIM1	41.0	41.0	61.0	37.0	35.0	55.0	53.0	25.0	25.0	36.0	39.0	55.0	47.0	8.0	23.0	35.0	34.0	44.0	42.0	45.0	41.0
SlNIM2	39.0	39.0	55.0	37.0	36.0	49.0	47.0	26.0	23.0	37.0	36.0	49.0	47.0	9.0	25.0	35.0	34.0	41.0	40.0	41.0	38.0
TcNPR1	60.0	61.0	42.0	54.0	53.0	39.0	40.0	24.0	24.0	55.0	58.0	41.0	39.0	10.0	21.0	53.0	50.0	67.0	71.0	74.0	76.0
VvNPR1.1	62.0	64.0	44.0	52.0	52.0	41.0	41.0	21.0	20.0	55.0	58.0	43.0	40.0	7.0	21.0	52.0	49.0	69.0	72.0	74.0	71.0
VvNPR1.2	43.0	44.0	67.0	37.0	37.0	60.0	59.0	27.0	25.0	39.0	43.0	59.0	50.0	7.0	24.0	36.0	35.0	43.0	43.0	47.0	43.0

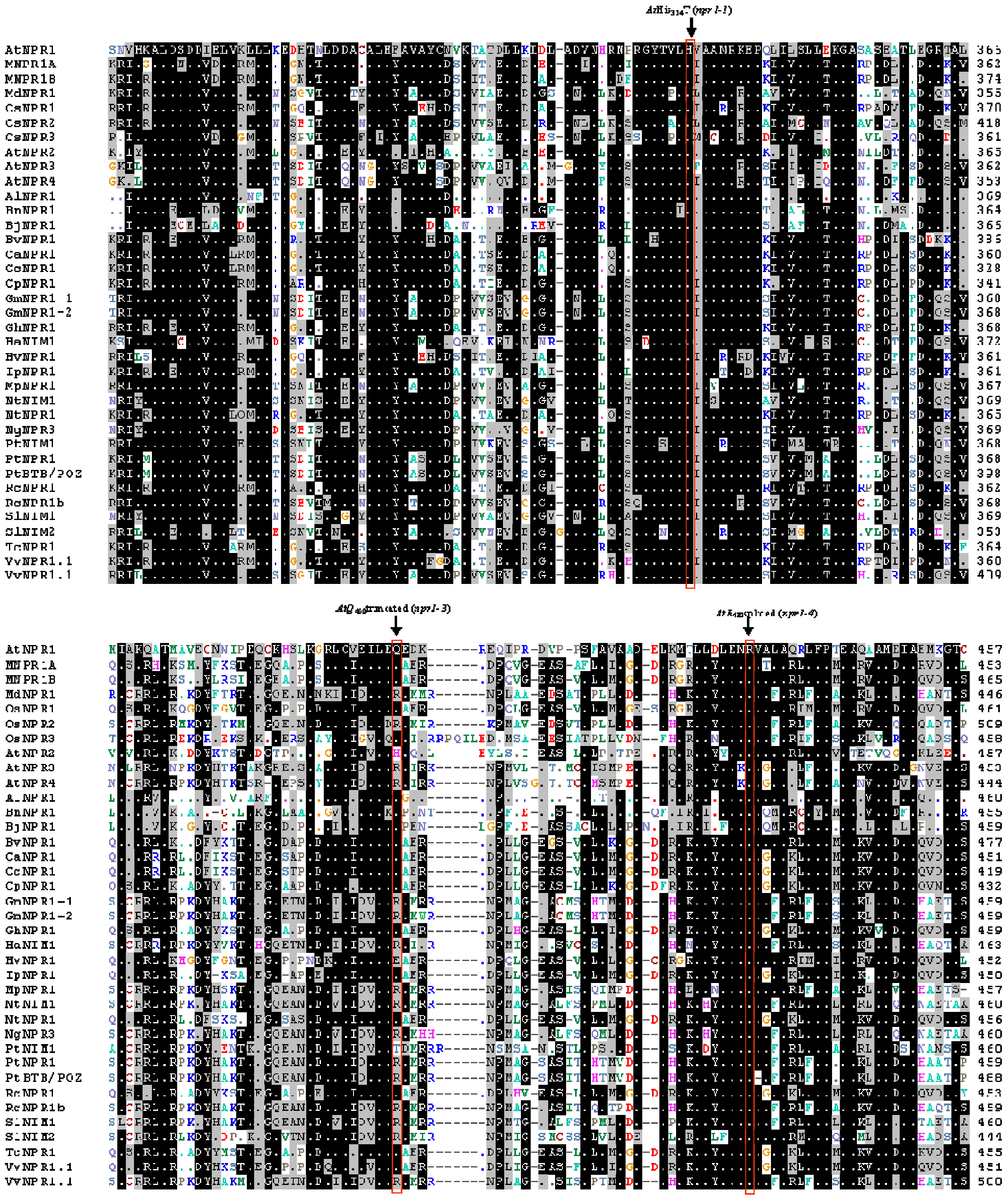
Name	GmNPR 1-1	GmNPR 1-2	GhNPR 1	HaNIM1	HvNP R1	IpNPR1	MxdN PR1	NtNIM 1	NtNPR 1	PtNIM 1	PtNPR 1	PtNI M1b	RcNPR 1	RcNPR 1b	SlNIM 1	SlNIM 2	TcNPR 1	VvNPR 1.1	VvNPR 1.2
GmNPR1-1																			
GmNPR1-2	96.0																		
GhNPR1	44.0	44.0																	
HaNIM1	68.0	67.0	41.0																
HvNPR1	41.0	38.0	54.0	36.0															
IpNPR1	41.0	40.0	64.0	41.0	56.0														
MxdNPR1	75.0	74.0	42.0	62.0	40.0	39.0													
NtNIM1	70.0	69.0	41.0	64.0	37.0	39.0	67.0												
NtNPR1	44.0	43.0	65.0	40.0	56.0	78.0	42.0	41.0											
NgNPR3	70.0	70.0	40.0	65.0	38.0	40.0	67.0	93.0	41.0										
PtNIM1	80.0	79.0	42.0	68.0	39.0	42.0	74.0	70.0	43.0	71.0									
PtNPR1	45.0	44.0	71.0	42.0	53.0	67.0	43.0	42.0	69.0	42.0	44.0								
PtNIM1b	62.0	61.0	38.0	57.0	35.0	37.0	58.0	58.0	38.0	58.0	61.0	41.0							
RcNPR1	45.0	44.0	73.0	40.0	54.0	70.0	43.0	40.0	71.0	40.0	44.0	80.0	40.0						
RcNPR1b	78.0	76.0	43.0	67.0	39.0	42.0	73.0	69.0	44.0	70.0	83.0	45.0	60.0	45.0					
SlNIM1	69.0	69.0	41.0	64.0	38.0	41.0	66.0	85.0	42.0	85.0	69.0	43.0	57.0	41.0	69.0				
SlNIM2	58.0	56.0	40.0	52.0	36.0	40.0	56.0	51.0	42.0	52.0	56.0	39.0	54.0	40.0	56.0	51.0			
TcNPR1	43.0	44.0	85.0	41.0	56.0	68.0	41.0	41.0	69.0	41.0	42.0	76.0	39.0	77.0	44.0	42.0	40.0		
VvNPR1.1	46.0	45.0	68.0	40.0	55.0	69.0	42.0	42.0	70.0	42.0	43.0	71.0	41.0	72.0	45.0	43.0	42.0	73.0	
VvNPR1.2	79.0	78.0	44.0	69.0	41.0	42.0	74.0	73.0	43.0	74.0	79.0	43.0	61.0	43.0	78.0	72.0	57.0	44.0	43.0

Gaps on table represent areas were sequence comparison results have already been presented

The two MNPR1 sequences also shared conserved regions that have been described in other NPR1-like genes (Fig. 3.1). These regions consist of a phosphodendron motif, the BTB/POZ domain, ankyrin repeats, and nuclear localization regions. However, within these domains, there exist some amino acid dissimilarities within the two MNPR1 sequences. Moreover, these sequence dissimilarities also exist among homologous NPR1 sequences from the same plants like rice, *Arabidopsis*, grape vine, soybean and *Brassica*. For instance, in the phosphodendron motif, the characterized serine amino acid at position 13 was present in the two MNPR1 sequences, *Os*NPR1, *Os*NPR2, *At*NPR1, *At*NPR2, *At*NPR3, but completely absent from MdNPR1 and replaced by a tyrosine in *Os*NPR3. Serine at position 15 on the other hand was present in MNPR1B, *At*NPR1, *At*NPR2, *Os*NPR1 and *Os*NPR3 but replaced by an arginine in MNPR1A. This serine was also completely absent from the MdNPR1 banana sequence. Another pronounced dissimilarity was found for the already characterized cysteine at position 529 in the *At*NPR1 sequence. This was replaced by a tyrosine in MNPR1A, MdNPR1, and *At*NPR4; by a glycine in MNPR1B; by a serine in *Os*NPR1, *At*NPR2, by aspartic acid in *Os*NPR2, by a leucine in *Os*NPR3, and by a histidine in *At*NPR3.


At Phosphodegron motif
At Cys82
AtNPR1
MNPR1A
MNPR1B
MdNPR1
OsNPR1
OsNPR2
OsNPR3
AtNPR2
AtNPR3
AtNPR4
AlNPR1
BnNPR1
BjNPR1
BvNPR1
CaNPR1
CcNPR1
CpNPR1
GmNPR1-1
GmNPR1-2
GhNPR1
HaNIM1
HvNPR1
IpNPR1
MpNPR1
NtNIM1
NtNPR1
NgNPR3
PtNIM1
PtNPR1
PtBTB/POZ
RcNPR1
RcNPR1b
SlNIM1
SlNIM2
TcNPR1
VvNPR1.1
VvNPR1.1


AtCys150Y (npr1-2)
AtCys155 and 156
AtCys160
AtCys216
AtNLS1 (252-265)
AtNPR1
MNPR1A
MNPR1B
MdNPR1
OsNPR1
OsNPR2
OsNPR3
AtNPR2
AtNPR3
AtNPR4
AlNPR1
BnNPR1
BjNPR1
BvNPR1
CaNPR1
CcNPR1
CpNPR1
GmNPR1-1
GmNPR1-2
GhNPR1
HaNIM1
HvNPR1
IpNPR1
MpNPR1
NtNIM1
NtNPR1
NgNPR3
PtNIM1
PtNPR1
PtBTB/POZ
RcNPR1
RcNPR1b
SlNIM1
SlNIM2
TcNPR1
VvNPR1.1
VvNPR1.1


AtHis334Y (npr1-1)
AtQ400truncated (npr1-3)
AtR432replaced (npr1-4)
AtNPR1
MNPR1A
MNPR1B
MdNPR1
OsNPR1
OsNPR2
OsNPR3
AtNPR2
AtNPR3
AtNPR4
AlNPR1
BnNPR1
BjNPR1
BvNPR1
CaNPR1
CcNPR1
CpNPR1
GmNPR1-1
GmNPR1-2
GhNPR1
HaNIM1
HvNPR1
IpNPR1
MpNPR1
NtNIM1
NtNPR1
NgNPR3
PtNIM1
PtNPR1
PtBTB/POZ
RcNPR1
RcNPR1b
SlNIM1
SlNIM2
TcNPR1
VvNPR1.1
VvNPR1.1

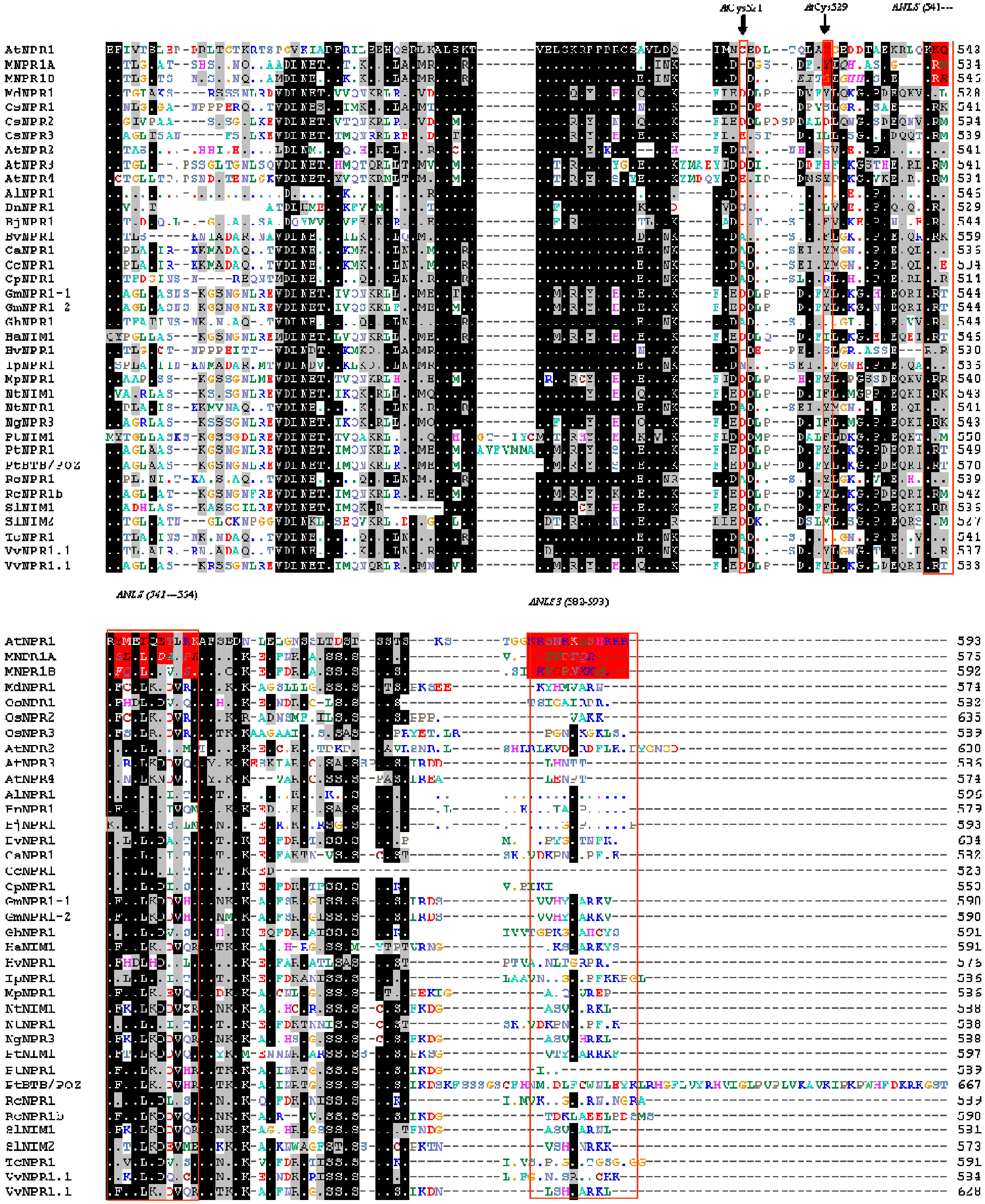

AtCys521
AtCys529
ANLS (541---
ANLS (541—554)
ANLS3 (580-593)
AtNPR1
MNPR1A
MNPR1B
MdNPR1
OsNPR1
OsNPR2
OsNPR3
AtNPR2
AtNPR3
AtNPR4
AlNPR1
DnNPR1
BjNPR1
BvNPR1
CaNPR1
CcNPR1
CpNPR1
GmNPR1-1
GmNPR1-2
GhNPR1
HaNIM1
HvNPR1
IpNPR1
MpNPR1
NtNIM1
NtNPR1
NgNPR3
PtNIM1
PtNPR1
PtBTB/POZ
PcNPR1
RcNPR1b
SlNIM1
SlNIM2
TcNPR1
VvNPR1.1
VvNPR1.1

Figure 3.1 Multiple alignment of MNPR1A and MNPR1B with other plant NPR1-like amino acid coding sequences. Amino acid sequences were aligned using the ClustalW multiple alignment program (Thompson *et al*., 1994). Identical amino acids are represented with dots, vertical red rectangles represent conserved domains and specific arrows show other single conserved amino acids with reference to the *Arabidopsis* NPR1 sequence. Differences in amino acids in the conserved domains between the *Arabidopsis* NPR1 and banana MNPR1 sequences are shaded in red while differences between the two banana MNPR1 sequences are shaded in red and italicized.

3.4.2 Phylogenetic grouping of MNPR1 coding sequences

Bootstrap consensus for neighbour joining (NJ) produced a total of three distinct groups in the phylogenetic tree. The two MNPR1 coding sequences (Fig 3.2; red) belonged to the first group and were closely related with each other (Fig. 3.2; dark red). They further grouped closely with NPR1 sequences of two other monocotyledons (*Hordeum vulgare* and the NHI sequence from rice). Although the banana ABB NPR1-like sequence belonged to a different group (Fig. 3.2 green) from the two MNPR1 banana sequences, it however grouped closely with the rice NH2 and NH3 sequences as well as many sequences known to mediate defense responses. NH5 *At*NPR5 and *At*NPR6 were the most distant NPR1-like sequences, while NH4 was used to root the tree.

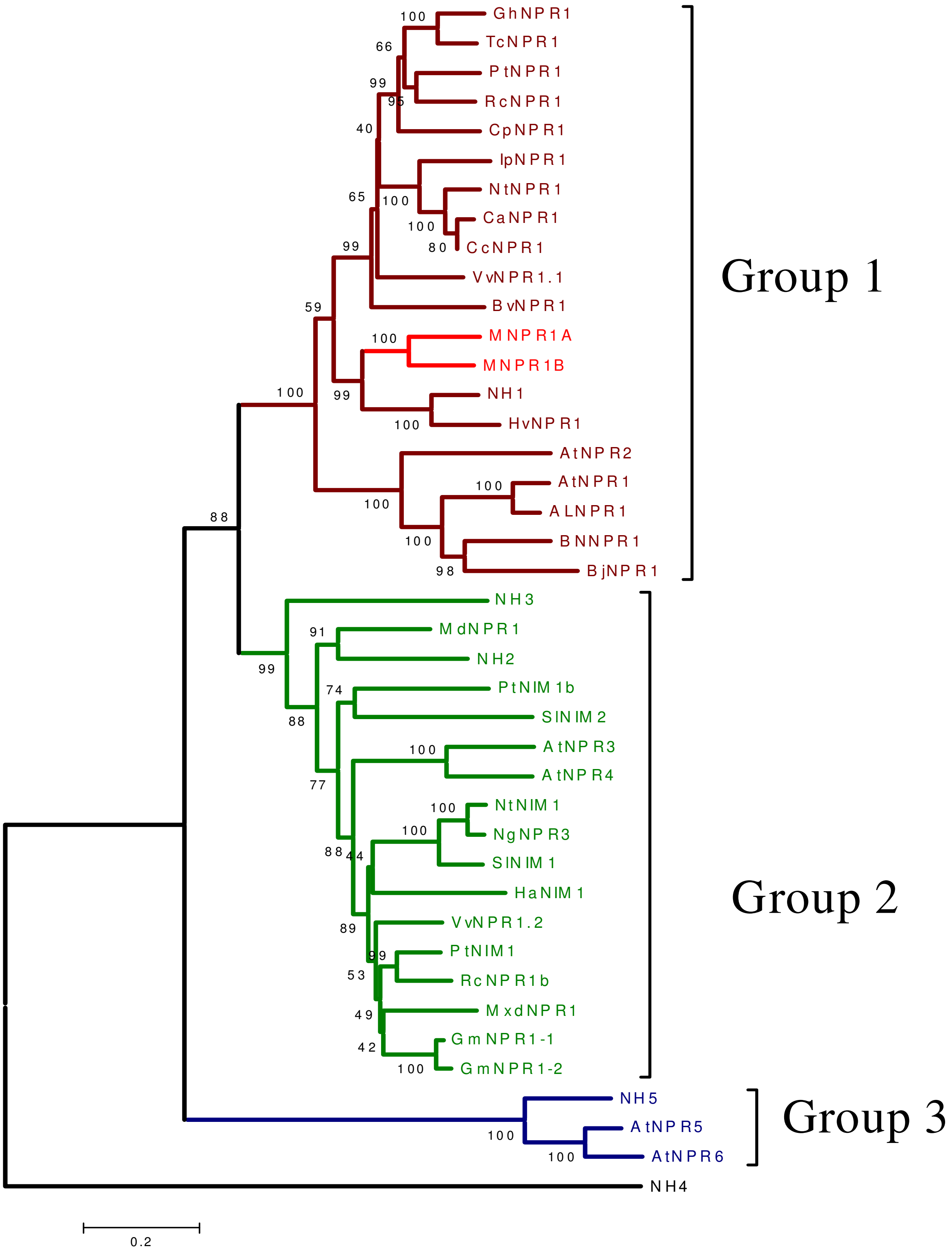


Figure 3.2 Multiple alignment and evolutionary relationship of the various plant NPR1-like coding sequences. Amino acid coding sequences from 41 plants were retrieved from Genbank

and aligned using the Mafft software program. The alignment was then edited in Bioedit to obtain the correct reading frame. The evolutionary history was inferred using the Neighbor-Joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree.

3.4.3 Comparison of defense-related *cis*-elements within the *MNPR1* coding sequences

From PLACE/Signal scan results it was apparent that the two banana *NPR1*-like coding regions harbored known *cis*-elements involved in defense responses mediated by pathogens, JA, SA, ET and ABA (Table 3.3). The *MNPR1A* gene was more enriched in the number of occurrence of each defense *cis*-elements (Fig. 3.3). For instance the transcription factor WRKY71*Os* appeared eight times in the positive *MNPR1A* strand but only six times in the positive *MNPR1B* strand. The ET-responsive factor3 (ERF3) was also over-represented in *MNPR1A* in comparison to *MNPR1B*. However, *MNPR1B* had three additional *cis*-elements (TCA, AGCBOXPGLB, ELRECOREPCRP1 and ethylene responsive element) known to be responsive to SA which were absent in *MNPR1B* (Table 3.3).

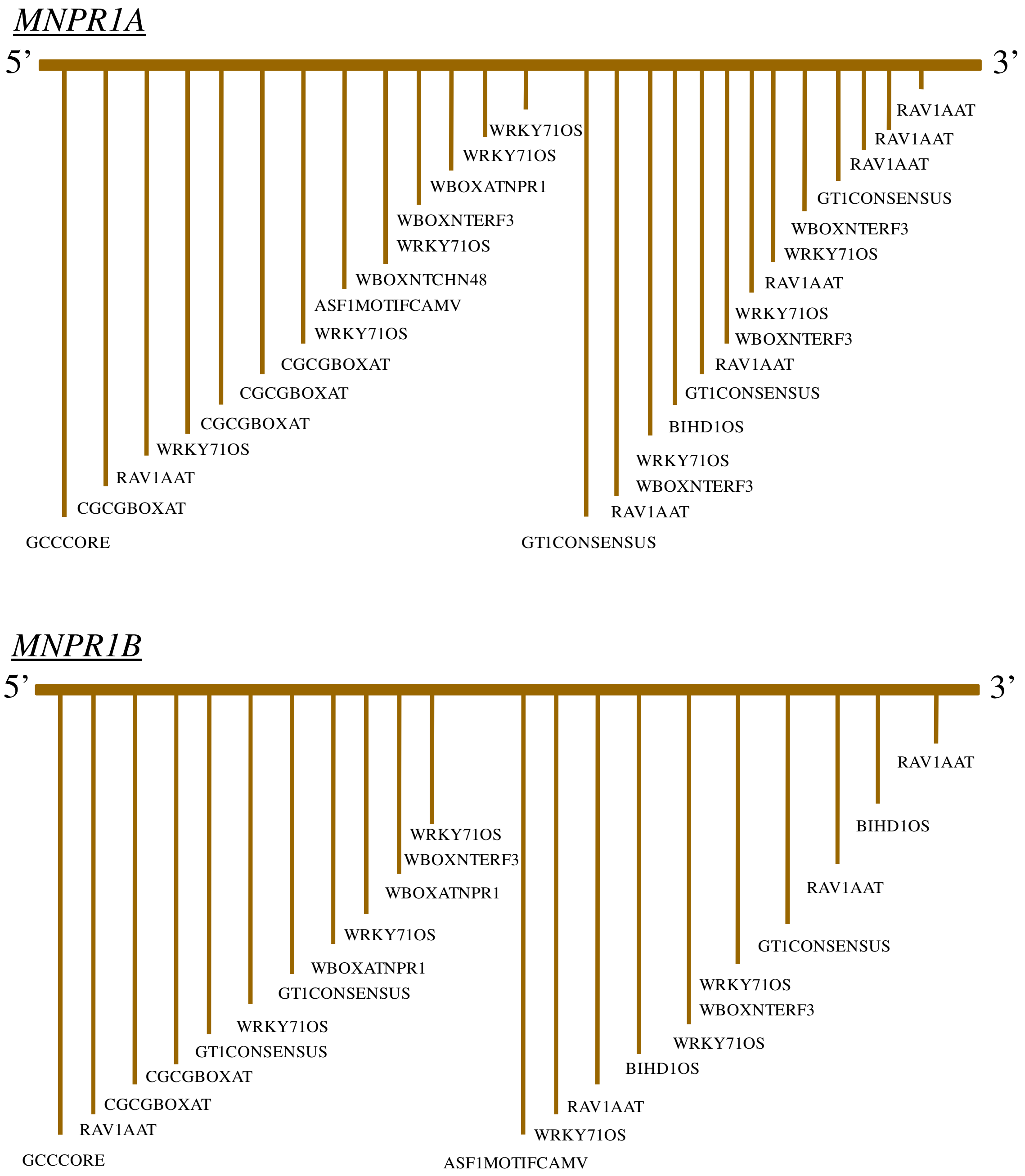


Figure 3.3 Defense-related *cis*-regulatory elements identified within the *MNPR1* coding sequences. *MNPR1* coding sequences were retrieved from the NCBI database. The defense *cis*-regulatory elements in the positive strand of *MNPR1A* and *MNPR1B* were identified on

PLACE/Signal scan database (http://www.dna.affrc.go.jp/PLACE/signalup.html; Higo *et al*., 1999) and represented linearly.

Table 3.3 Elicitor/light-responsive defense *cis*-elements in the coding region of the *MNPR1A*, and *MNPR1B* sequences.

Sequence-*Cis* element	**Strand/ position**		**Function**	**Organism**	**Reference**
	"+"	"-"			
GCCGCC **–** **GCCCORE**	22^{a} 22^{b}	73^{a} 81^{b}	Ethylene responsive element present in PDF1.2 Thi2.1, *PR-4*. Plays a role in JA and ET responsive gene expression	*A. thaliana*, *L. esculentum*	Brown *et al.*, (2003); Chakravarthy *et al.*, 2003); Berrocal-Lobo and Molina, (2004), Oñate-Sánchez *et al.*, (2007)
GRWAAW – GT1 CONSENSUS	585^{a}, 778^{a}, 811^{a}, 1396^{a} 363^{b}, 814^{b}, 1432^{b}	771^{a}, 1363^{a}, 1505^{a}, 1506^{a} 1540, 1541^{b}	Light-regulated genes, stabilizes TATA box complex; GT1 interact with the promoter of *PR-1a* to regulate SA and SAR gene products.	*L. esculentum*	Le Gourrierec *et al.*, (1999); Buchel *et al.*, (1999); Buchel *et al.*, (1996) Terzaghi *et al.*, 1995; Villain *et al.*, 1996. Zhou, (1999).
GAAAAA **–** **GT1 GMSCAM4**		771^{a}, 1505^{a} 1540^{b}	Mediates pathogen- and NaCl-induced gene expression. Found in GT-1 promoters like the soybean calmodulin (CaM) isoform	*L. esculentum*	Park *et al.*, (2004), Hung *et al.*, 2004).
CCGAC – LTRECOREATCOR15	44^{a}, 89^{a}, 173^{a} 44^{b}, 194^{b}	220^{a}, 343^{a}, 1702^{a} 142^{b}, 211^{b}, 379^{b}	Core of low temperature responsive element (LTRE) of cor15a gene, responsive to ABA and induces BN115 gene during cold stress.	*A. thaliana*; *Brassica napus*	Baker *et al.*, (1994); Jiang *et al.*, (1996), Busk *et al.*, (1998).

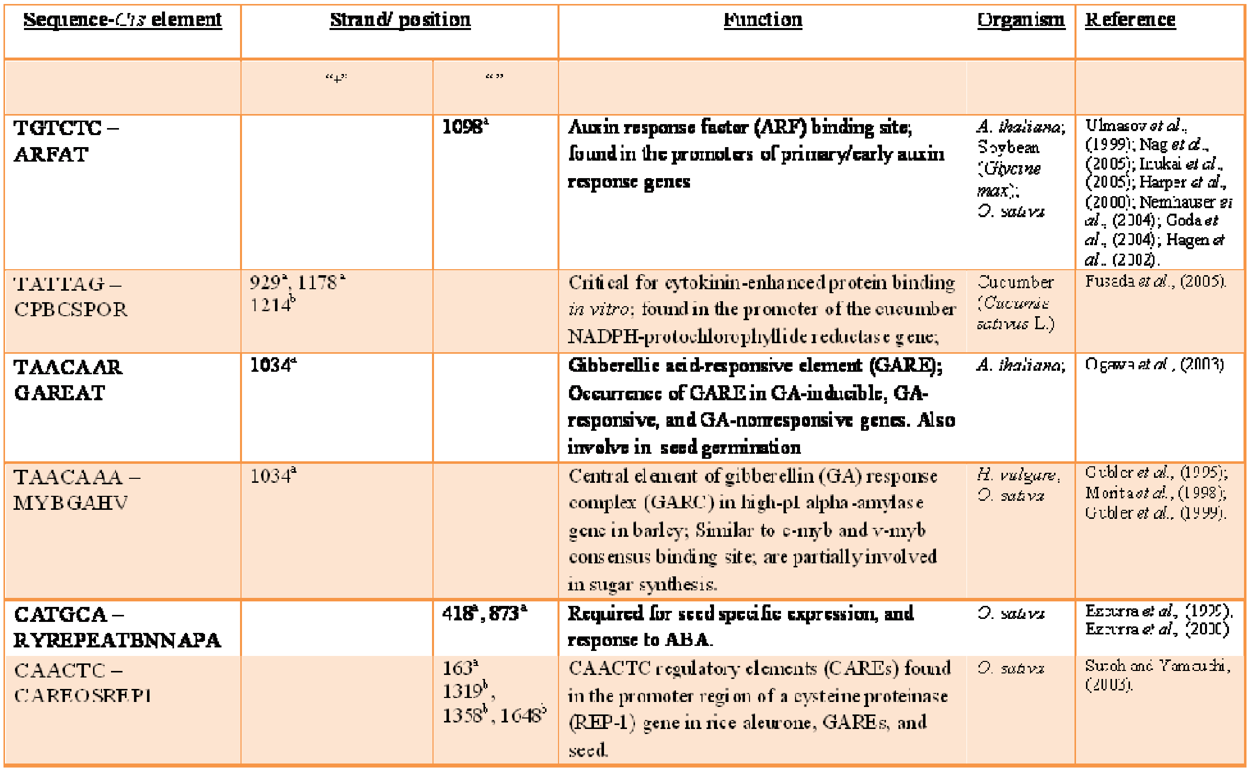
Sequence-*Cis* element	Strand/ position		Function	Organism	Reference
	"-"	"+"			
MACGYGB – MACGYGB		29^{a}	An ABRE-related sequence found upstream of the Ca^{2+}-responsive unregulated genes	*Arabidopsis thaliana*	Kaplan *et al.*, (2006)
AGCAGC - ANAERO2CONSENSUS; TCATCAC - ANAERO3CONSENSUS, GTTTHGCAA- ANAERO4CONSENSUS	52^{b}, 799^{b}, 802^{b}	85^{c} 1244^{a} 851^{b}, 1280^{b} 635^{a}	Anaerobic stress/ respiration	Maize (*Zea mays*); *A. thaliana*; Pea (*Pisum sativum*); Sunflower (*Helianthus vulgare*); Rice (*Oryza sativa*); Petunia (*Petunia hybrida*); Tomato (*Lycopersicon esculentum*)	Mohanty *et al.*, (2005).
RCCGAC DRECRTCOREAT	172^{c} 193^{b}	343^{a}, 1702^{a}	Core motif of dehydration-responsive element/C-repeat (DRE/CRT) cis-acting element. DRE/CRT are involve in drought; high-light; cold stress	*O. sativa*; *Z. mays*, (*H. annuus*	Dubouzet *et al.*, (2003); Qin *et al.*, (2004); Diaz-Martin *et al.*, (2005); Suzuki *et al.*, (2005) Skinner *et al.*, (2005);
TGTCA BIHD1OS	711^{c}, 786^{c} 1061^{b}, 1664^{b}	126_{a}, 558^{a}, 576^{a} 612^{b}, 846^{b}, 1070^{b}	Rice BELL homeodomain TF in disease response. Activated by BTH and *Magnaporthe grisea*. Leads to *PR-1* gene activation	*O. sativa*	Luo *et al.*, (2005)
VCGCGB – CGCGBCXAT	29^{a}, 211^{a}, 245^{a}, 247^{a} 113^{b}, 329^{b}	29^{a}, 211^{a}, 245^{a}, 247^{a} 113^{b}, 329^{b}	Calmodulin-involved in multiple signaling Induced by MeJA, SA, H_2O_2, wounding, ABA, salt, UVB stress	*A. thaliana*	Liu *et al.*, (2009), Wang *et al.*, (2009), Yang and Poovaiah (2002), Park *et al.*, (2004)

Sequence-*Cis* element	**Strand/ position**		**Function**	**Organism**	**Reference**
	"+"	"-"			
TAACTG – MYB2AT		656^{a}	Regulates water stress-related genes. Binds to *AtMYB2*.	*A. thaliana*	Urao *et al.*, (1993).
YAACKG – MYB2CONSENSUSAT	1130^{a}	656^{a}	MYB recognition site present in the promoter of the dehydration-responsive gene *rd22*	*A. thaliana*	Abe *et al.*, (2003).
CNGTTR – MYBCORE	656^{a}, 941^{a}, 1523^{a} 532^{b}, 954^{b}	48^{a}, 920^{a}, 1130^{a} 48^{b}, 956, 1419^{b}	Binding site for all animal MYB and at least two *Arabidopsis* MYB proteins AtMYB2 regulates water stress responsive genes. MYB protein (MYB.Ph3) contributes in regulating the flavonoid biosynthetic pathway.	*A. thaliana*; animal; *P. hybrida*;	Luscher and Eiseman, (1990), Urao *et al.*, (1993), Solano *et al.*, (1995)
CATGTG – MYCATERD1	420^{a}, 875^{a} 915^{b}		MYC recognition sequence necessary for expression of early responsive to dehydration (erd1) gene during dehydration	*A. thaliana*;	Chinnusamy *et al.*, (2003), Chinnusamy *et al.*, (2004); Abe *et al.*, (2003), Oh *et al.*, 2005); Lee *et al.*, (2005), Hartmann *et al.*, (2005), Agarwal *et al.*, (2006).
CANNTG – MYCCONSENSUSAT	420^{a}, 539^{a}, 875^{a}, 1130^{a} 54^{b}, 675^{b}, 915^{b}, 983^{b}, 1088^{b}, 1265^{b}, 1489^{b}	420^{a}, 539^{a}, 875^{a}, 1130^{a} 54^{b}, 675^{b}, 915^{b}, 983^{b}, 1088^{b}, 1265^{b}, 1489^{b}	MYC recognition site found in the promoters of the *dehydration-responsive* gene *rd22*. Regulates transcription of the cold stress related gene CBF/DREB1).	*A. thaliana*;	
CACATG – MYCATRD22		420^{a}, 875^{a} 915^{b}	Drought-, ABA- regulated gene expression by MYB and MYC	*A. thaliana*	Anderson *et al.*, (2004); Abe *et al.*, (1997); Busk *et al.*, (1998).

Sequence-*Cis* element	Strand/ position		Function	Organism	Reference
	"+"	"-"			
TGACY – **WBOXNTERF3**	**404^{a}, 698^{a}, 906^{a} 861^{b}, 1284^{b}**	**522^{a}, 662^{a}, 1127^{a} 233^{b}, 558^{b}, 693^{b}, 698^{b}, 1560^{b}**	**Activators of ERF3 gene by wounding in tobacco leaves.**	*N. tabacum*	Nishiuchi *et al.*, (2004).
TGAC – WRKY71OS	126^{a}, 362^{a}, 404^{a}, 558^{a}, 576^{a}, 698^{a}, 906^{a}, 1248^{a} 612^{b}, 846^{b}, 861^{b}, 906^{b}, 1070^{b}, 1284^{b}	523, 663, 712, 787, 1128^{a}, 1285^{a} 234^{b}, 559^{b}, 694^{b}, 699^{b}, 823^{b}, 1062^{b}, 1561^{b}, 1665^{b}, 1731^{b}	Repressor of GA signaling binds to PR-10 promoters; induced by ABA. Upregulated by SA, MeJA, ACC, wounding and pathogen infection.	*O. sativa*; Parsley (*Petroselinum crispum*)	Liu *et al.*, (2007), Zhang *et al.*, (2004); Xie *et al.*, (2005); Eulgem *et al.*, (1999); Eulgem *et al.*, (2000).
TTGAC – **WBOXATNPR1**	**557^{a} 845^{b}, 860^{b}**	**712^{a}, 787^{a} 234^{b}, 694^{b}, 823^{b}, 1665^{b}**	**WRKY- (W) –box are SA inducible. Present in *NPR1* promoters.**	*A. thaliana*	Yu *et al.*, (2001); Chen *et al.*, (2002); Eulgem *et al.*, (2000); Chen *et al.*, (2002); Maleck *et al.*, (2000); Xu *et al.*, (2006).
CTGACY – WBOXNTCHN48	403^{a}	522^{a}, 662^{a} 558^{b}, 698^{b}	WRKY1, 2, 4 in tobacco related *cis*-elements of basic chitinase gene. Possibly involve in elicitor-responsive transcription of defense genes in tobacco.	*Nicotiana tabacum*	Yamamoto *et al.*, (2004).

Sequence-*Cis* element	Strand/ position		Function	Organism	Reference
	“+”	“-”			
TCATCTTCTT – **TCA1MOTIF**		**1765^{b}**	**SA inducible expression of genes. Found in barley beta-1,3-glucanase and over 30 different plant genes known to be induced by one or more forms of stress**	*H. vulgare*, *N. tabacum*	Goldsbrough *et al*., (1993); Mhiri *et al*., 1997).
CAACA – RAV1AAT	48^{a}, 636^{a}, 860^{a}, 920^{a}, 1662^{a}, 1694^{a}, 1715^{a} 48^{b}, 956^{b}, 1566^{b}, 1736^{b}	822^{a}, 842^{a} 633^{b}, 858^{b}, 1301^{b}	RAV1 box, positively regulates early responses to pathogens, SA, wounding and ET.	*A. thaliana*	Sohn *et al*., (2006); Kagaya *et al*., (1999).
TGACG – **ASF1MOTIFCAMV,**	**362^{a}** **906^{b}**	**822^{b}, 1730^{b}**	**Activation sequence factor1 (ASF-1) binding site in CaMV 35S promoter. Found in genes induced by auxins, SA genes involved in biotic and abiotic stress. TGA binding site in SA responsive *PR-1* promoters. Differentially stimulate the activity of the "as-1 element" during xenobiotic stress, SAR, Disease resistance.**	Cauliflower mosaic virus (CaMV), *N. tabacum*, *A. thaliana*	Després *et al*., (2003); Terzaghi *et al*., (1995); Benfey *et al*., (1990); Katagiri *et al*., (1989); Xiang *et al*., (1997); Klinedinst *et al*., (2000); Redman *et al*., (2002).

Sequence-*Cis* element	Strand/ position		Function	Organism	Reference
	“+”	“-”			
AGCCGCC – **AGCBOXNPGLB**		**81^{b}**	**Ethylene responsive factors (ERFs); ERF1,2,5; activates GCC BOX dependent transcription of stress signal responsive factors; binding sites are conserved in *PR* genes and MAPK**	*N. plumbaginifolia*; *N. sylvestris*; *A.thaliana*; *O. sativa*	Hart *et al*., (1993); Fujimoto *et al*., (2000); Sato *et al*., (1996); Ohme-Takagi *et al*., (2000); Rushton *et al*., (2002); Zhang *et al*., (2004).
TTGACC – ELRECOREPCRP1		233^{b}	Elicitor responsive elements (WRKY1, 2); PR-1, SA inducible.	*P. crispum*; *N. tabacum*	Rushton *et al*., (1996); Chen and Chen (2000); Eulgem *et al*., (1999); Eulgem *et al*., (2000); Rushton *et al*., (2002); Laloi *et al*., (2004)
CCGTCC – **PALBOXAPC**		**195^{a}** **225^{b}**	**PAL gene family activation. Involved in the induction of transcriptional activity though not sufficient to confer UV-light or elicitor responsiveness.**	*P. crispum*	Lois *et al*., (1989); Logemann *et al*., (1995);



Sequence-*Cis* element	Strand/ position		Function	Organism	Reference
	“+”	“-”			
TGTCTC – ARFAT		**1098^{a}**	**Auxin response factor (ARF) binding site; found in the promoters of primary/early auxin response genes**	*A. thaliana*; Soybean (*Glycine max*); *O. sativa*	Ulmasov *et al*., (1999); Nag *et al*., (2005); Inukai *et al*., (2005); Harper *et al*., (2000); Nemhauser *et al*., (2004); Goda *et al*., (2004); Hagen *et al*., (2002).
TATTAG – CPBCSPOR	929^{a}, 1178^{a} 1214^{b}		Critical for cytokinin-enhanced protein binding *in vitro*; found in the promoter of the cucumber NADPH-protochlorophyllide reductase gene;	Cucumber (*Cucumis sativus* L.)	Fusada *et al*., (2005).
TAACAAR GAREAT	**1034^{a}**		**Gibberellic acid-responsive element (GARE); Occurrence of GARE in GA-inducible, GA-responsive, and GA-nonresponsive genes. Also involve in seed germination**	*A. thaliana*;	Ogawa *et al*., (2003)
TAACAAA – MYBGAHV	1034^{a}		Central element of gibberellin (GA) response complex (GARC) in high-pI alpha-amylase gene in barley; Similar to c-myb and v-myb consensus binding site; are partially involved in sugar synthesis.	*H. vulgare*, *O. sativa*	Gubler *et al*., (1995); Morita *et al*., (1998); Gubler *et al*., (1999).
CATGCA – RYREPEATBNNAPA		**418^{a}, 873^{a}**	**Required for seed specific expression, and response to ABA.**	*O. sativa*	Ezcurra *et al*, (1995). Ezcurra *et al*, (2000)
CAACTC – CAREOSREP1		163^{a} 1319^{b}, 1358^{b}, 1648^{b}	CAACTC regulatory elements (CAREs) found in the promoter region of a cysteine proteinase (REP-1) gene in rice aleurone, GAREs, and seed.	*O. sativa*	Sutoh and Yamauchi, (2003).

Sequence-*Cis* element	Strand/ position		Function	Organism	Reference
	“+”	“-”			
RYCGAC - CBFHV	172^{a} 193^{b}, 463^{b}, 1744^{b}	343^{a}, 466^{a}, 1702^{a} 463^{b}, 502^{b}, 1164^{b}	C-repeat (CRT) binding factors; CBFs are also known as dehydration-responsive element (DRE) binding proteins (DREBs)	Barley (*Hordeum vulgare*)	Xue, (2002); Svensson *et al*., (2006).
WAACCA – MYB1AT	954^{a} 990^{b}		MYB recognition site found in the promoters of the dehydration-responsive gene *rd22*. Involve in ABA signaling.	*A. thaliana*.	Abe *et al*., (2003).
CAANNNNATC- CIACADIANLELHC	714, 1530^{a} 1566^{b}	550, 817^{a} 853^{b}	Required for circadian expression of tomato light harvesting complex (Lhc) gene;	*L. esculentum*	Piechulla *et al*., (1998).
GATAAG - IBOX		120^{a}, 1565^{a}	Conserved sequence upstream of light-regulated genes; Prsence in the promoter region of rbcS of plants.	L. *esculentum*; *A. thaliana*	Giuliano *et al*., (1988); Donald and Cashmore, (1990); Rose *et al*., (1999).
GATAA – IBOXCORE	778^{a}, 811^{a} 814^{b}	121^{a}, 1526^{a}, 1566^{a}	Conserved sequence upstream of light-regulated genes of both monocots and dicots;	Monocots; dicots	Terzaghi *et al*., (1995).
AAMAATCT – CCA1ATLHCB1		1297^{b}	CCA1 binding site; CCA1 protein (myb-related transcription factor) interact with the light related gene Lhcb1*3. Related to regulation by phytochrome in shoots leaves	*A. thaliana*	Wang *et al*., (1997).
TATTCT – -10PEHVPSBD	552^{a} 588^{b}	793^{a} 829^{b}	Involve in chloroplast gene expression, circadian rhythms and light regulation	*H. vulgare*	Thum *et al*., (2001)

Sequence-*Cis* element	Strand/ position		Function	Organism	Reference
	"+"	"-"			
GATA – GATABOX	550^{a}, 613^{a}, 706^{a}, 778^{a}, 811^{a}, 949^{a}, 1176^{a} 649^{b}, 709^{b}, 742^{b}, 814^{b}, 1212^{b}	13^{a}, 122^{a}, 593^{a}, 1527^{a}, 1567^{a} 272^{b}, 1262^{b}, 1306^{b}	GATA motif in CaMV 35S promoter; Binding with ASF-2; found in the promoter of Petunia chlorophyll a/b binding protein, *Cab22* gene; Required for light regulated, and tissue specific expression; Conserved in the promoter of all LHCII type I Cab genes;	CaMV; *P. hybrid,* *A. thaliana;* *O. sativa*	Lam and Chua (1989); Gilmartin *et al.*, (1990), Benfey and Chua, (1990) Giuliano *et al.*, (1989), Teakle *et al.*,(2002), Reyes *et al.*, (2004), Rubio-Somoza *et al.*, (2006)
SCGAYNRNNNNNNNNNNNNNNNNHD –PRECONS CRHSP70A	**44^{a} 44^{b}, 1447^{b}**	**579^{a}, 1983^{a} 192^{b}**	**Consensus sequence of (plastid response element PRE) in the promoters of HSP70A in *Chlamydomonas*; Involved in induction of HSP70A gene by light.**	*C. reinhardtii*	von Gromoff *et al.*, (2006).
GCCAC – SORLIP1AT	25^{b}, 214^{b}	394^{a}, 895^{a} 215^{b}, 333^{b}, 397^{b}, 430^{b}	One of "Sequences Over-Represented in Light-Induced Promoters (SORLIPs) in *Arabidopsis*; over-represented in light-induced cotyledon and root common genes and root-specific genes.	*A. thaliana;*	Hudson and Quail. (2003); Jiao *et al.*, (2005).
GGGCC – SORLIP2AT		**108^{a}**	**one of "Sequences Over-Represented in Light-Induced Promoters (SORLIPs) in *Arabidopsis*.**	*A. thaliana;*	Hudson and Quail. (2003)

The *cis*-element in the coding regions was analyzed in place scan database. The position of the element in each gene is represented by "a" for M*NPR1A*, and "B" for *MNPR1B* sequences. Full citations in table are found in the reference section.

3.5 Discussion

New information contributed by this study is the type and relative frequency of *cis*-regulatory elements indentified within the *MNPR1A* and *MNPR1B* coding regions. The *cis*-elements were more abundant in the *MNPR1A* sequence relative to *MNPR1B*. Over-representation of such *cis*-elements such as the ERF might mediate the *MNPR1A's* activity in necrotrophic pathogens that mediate JA responses in plants (Oliver and Ipcho, 2004; Glazebrook, 2005; Spoel *et al.*, 2007; Endah *et al.*, 2008) Of further interest from this study was also the fact that *MNPR1B* had additional motifs like the TCA and ELRECOREPCRP1 which have been identified in tobacco and sorghum SA-inducible *WRKY1*,2 and *PR-1* genes (Goldsbrough *et al.*, 1993; Mhiri *et al.*, 1997; Eulgem *et al.*, 1999). This might account to some extend for the increased transcription of the *MNPR1B* gene following SA treatment in our previous study (Endah *et al.*, 2008). The additional ethylene responsive element found in *MNPR1B* sequence buy not in *MNPR1A* might also have a functional role in mediating JA-defense response has also observed in the gene's response to MeJA (Endah *et al.*, 2008). All the *cis* elements identified during this study have also been reported to play many diverse functions in pathogen (*Pseudomonas*) and elicitor (SA, JA, ET) mediated responses as outlined in Table 3.2. How these *cis*-elements actually interact to mediate overall defense remains unknown. However, results following characterization of the expression of both genes in banana show that they are sequentially transcribed by the hemibiotrophic pathogen *X. campestris* which can elicit both a SA and JA response pathway in infested plants, differentially regulated by SA or the necrotrophic *Foc* pathogen which elicits a JA/ET dependent response, and also expressed following JA treatment (Glazebrook, 2005; Spoel *et al.*, 2007; Endah *et al.*, 2008; Endah *et al.*, 2010).

In correlation with previous studies, the two banana MNPR1 coding sequences group closely together but, had some sequence dissimilarities even within highly conserved regions. It is well-known that most plants harbour more that a single copy of the NPR1-like gene with *Arabidopsis* having 6 copies (Hepworth et al., 2005). Food crops like rice, Soybean, and grape vine are known to harbour 5, and 2 copies each of the gene, respectively (Yuan *et al*., 2007; Endah *et al*., 2008; Zhao *et al*., 2009; Sandhu *et al*., 2009). Sequence analysis from previous studies have already revealed that these homologs share some sequence dissimilarity to each other even within such highly conserved domains as the BTB/POZ and ankyrin domains (Sandhu *et al*., 2009) and this has again been observed with the banana MNPR1 sequences in this study. However, these differences might not be very important for adifferential function as seen with soybean *NPR1* and *NPR2* genes (Sandhu *et al*., 2009) or, it might influence the genes' behaviour to pathogens and elicitors and in different plant systems as seen for the rice *NPR1* homologues (Yuan *et al*., 2007).

Phylogenetic results from this study also correlated with results using 31 NPR1-like sequences or just the ankyrin domains of 10 NPR1-like sequences. The banana MNPR1 sequences grouped closely to each other and with the rice NH1 and the NPR1-like sequence from *H. vulgare*. In previous phylogenetic studies, such close groupings of the banana MNPR1 sequences with monocotyledons have been reported (Endah *et al*., 2008; Bergeault *et al*., 2010). Both Bergeault *et al*., (2010) working with 31 NPR1-like sequences and Hepworth *et al*. (2005) working with the 6 *At*NPR1-like sequences, also found 3 main groups during their phylogenetic study also found in this current study using 41 sequences. Both the MNPR1A and MNPR1B genes grouped with the *At*NPR1 sequence which is well-know to positively mediate defense response (Speol *et* *al*., 2009). MdNPR1 isolated from another banana cultivar with an ABB genome (Zhao *et al*., 2009) however grouped in another clade from the two banana AAA NPR1 sequences. Coupled to its low level of identity of less than 50% to the two MNPR1 sequences, it can be possible that this new copy of the banana NPR1 gene might have been contributed by the B genome as reported for the two *NPR1* gene copies in *Brassica juncea*, believed to have originated from two individual parental genomes (*B. rapa* and *B. nigra*; Meur *et al*., 2006). However, how the individual banana A and B genomes have evolved over-time is still unknown hence the origin of the three banana *NPR1*-like gene from the ABB genome is still highly speculative. This might however play important roles in mediating their responses during stress.

Data from this study clearly shows that two banana NPR1 like coding sequences (MNPR1A and MNPR1B) group closely with other defense response NPR1-like sequences but share sequence dissimilarities even in the highly conserved NPR1 functional regions. Furthermore, *MNPR1A* is over-represented in the occurrence of *cis* elements identified although *MNPR1B* has additional SA-responsive elements and an eythlene-responsive factor which might be important for their expression. The next part of the study focuses on a complementation study carried with *Arabidopsis npr1-2* mutants to demonstrate banana *NPR1* activity in protection against pathogens and to investigate if sequence dissimilarity in their coding sequence is responsible for differential activity of the genes.