Host response profiling to fungal infection: molecular cloning, characterization and expression analysis of \textit{NPR1} gene from coffee (\textit{Coffea arabica})

C. F. Barsalobres-Cavallari$^{1,2,3}$, A. S. Petitot$^1$, F. E. Severino$^2$, L. G. Maia$^2$ and D. Fernandez$^1$

$^1$Résistance des Plantes aux Bioagresseurs, IRD, 911 Avenue d'Agropolis, BP64501, 34394, Montpellier, France
$^2$Laboratório de Biotecnologia e Genética Molecular, IB, UNESP, 18618-000, Botucatu, São Paulo, Brazil
$^3$Present address: Departamento de Química e Biologia, Centro de Educação, Ciências Exatas e Naturais, Universidade Estadual do Maranhão, Av. Lourenço Vieira da Silva, s/n., Cidade Universitária Paulo VI, 65054-970 - São Luís, Maranhão, Brazil.

*Author for correspondence: barsalobres@gmail.com

The \textit{Arabidopsis thaliana} NPR1 protein (for Nonexpressor of PR genes) is a key positive regulator of the systemic acquired resistance (SAR), a general defense response in plants that is characterized by the expression of pathogenesis-related (PR) genes. Herein, we present a combination of coffee genome sequence resources, gene cloning experiments, and transient expression assays to make possible the functional characterization of the orthologous \textit{NPR1} (\textit{CaNPR1}) gene from coffee (\textit{Coffea arabica}). Data obtained suggest that this gene likely play a role during coffee defense responses to pathogen infection. The \textit{CaNPR1} gene may represent a promising candidate for engineering resistance in coffee to achieving broad spectrum resistance to pathogens.

**Keywords** plant-pathogen interaction; orange rust fungus; disease resistance.

1. General remarks

Inducible defenses against a number of pathogens are quickly expressed in the primary infection sites and are often visible as necrotic local lesions on the leaves of resistant plant varieties. This local reaction is referred to as the hypersensitive response (HR). Subsequently, the HR triggers a general resistance mechanism that renders uninfected parts of the plant tolerant to further pathogen attack, a phenomenon called systemic acquired resistance (SAR) [1]. The elicitation of the HR and SAR reactions is accompanied by the coordinated induction of a heterogeneous group of proteins which are impaired in the absence of salicylic acid (SA) [2-5]. Thus, the rise in SA levels triggers the activation of a number of defense genes. Among these genes, the best characterized code for an essential regulator of SAR, namely the NPR1 gene (for Nonexpressor of PR genes; also known as NIM1 - non-immunity, or SAI1 - salicylic acid insensitivity) [6-8]. According to Cao et al. [8], the NPR1 protein is constitutively expressed and activated following pathogen recognition and SA treatment, suggesting that this protein acts as a key positive regulator of the SA-dependent resistance signaling pathway [9-10]. In healthy tissues, NPR1 is known to be an oligomeric, cytosolic protein.

Following SA treatment, \textit{Arabidopsis} NPR1 (AtNPR1) dimers become monomers and move into the nucleus to interact with TGA transcription factors [11]. It was demonstrated that NPR1 directly control the expression of \textit{PR} genes which encode a basic protein reported to be pathogen-responsive [12-14]. In Fig. 1, we summarize the data concerning the role of NPR1 proteins in plant defense responses to pathogens.

NPR1 proteins are highly conserved across many plant species and a conserved role in disease resistance has been suggested, as described by Cantu et al. [21]. By instance, overexpression of AtNPR1 in carrot tissues offered the ability to control a wide range of different pathogens [22]. Given the pivotal role of NPR1 gene in plant defense, we assessed the contribution and regulation of this protein during biotic stress responses in coffee (\textit{Coffea arabica}), a key export and cash crop in tropical and subtropical developing countries. One of the major limiting factors for coffee production is coffee leaf rust (or orange rust) disease caused by the Basidiomycete fungus \textit{Hemileia vastatrix} Berkeley and Broome [23]. This fungus infects the lower surface of the leaves where it produces large, orange colonies of uredospores, leading to premature leaf fall and yield losses in coffee (\textit{C. arabica}) production [23].

Combining \textit{in silico} databases searches, a genome walking strategy and phylogenetic analyses, \textit{NPR1} orthologous sequences (termed \textit{CaNPR1}) were isolated from the genomic DNA of \textit{C. arabica}. Moreover, the promoter region of \textit{CaNPR1} gene was characterized and transiently expressed by agroinfiltration in leaves of \textit{Nicotiana benthamiana} for functional analyses. According to our results, transcript levels of \textit{CaNPR1} were not affected following infection of coffee plants with \textit{H. vastatrix}. Interestingly, the activity of the \textit{CaNPR1} promoter was found to be SA-dependent. These data demonstrate that SA strongly stimulates the transcriptional activity of the \textit{CaNPR1} promoter, and suggest that different \textit{cis}-acting elements located in this region, such as the plant hormone- and/or infection-related motifs, might contribute to SA responsiveness.
2. Findings

2.1. Isolation and characterization of the complete coffee NPR1 gene sequence

To isolate the NPR1 cDNA from *C. arabica* (*CaNPR1*), ESTs encoding a NPR1 protein were retrieved in subtractive libraries, associated with expression of early resistance mechanisms of coffee plants to the leaf rust pathogen *H. vastatrix*, constructed by Fernandez et al. [24] and in the Brazilian Coffee EST database [25]. These ESTs contained a 687 bp coding region that covered the C-terminal part of the encoded protein, as determined by a BLASTX search. They were used to design a pair of nested gene-specific primers to conduct the genome walking procedure for amplification of the 5' flanking region of *CaNPR1*. Hence, the 5'-flanking region of the *CaNPR1* gene was amplified using a set of nine gene-specific primers (CaNPR1-1 to CaNPR1-9). The first two (CaNPR1-1 and CaNPR1-2) were designed based on the starting EST sequence selected in the database while the others were designed based on the nucleotide sequences of the fragments amplified by genome walking.

A 4375 bp-long genomic fragment encompassing the entire *CaNPR1* coding region was obtained (Fig. 2a). The *CaNPR1* open reading frame encoded a protein of 609 amino acids (Fig. 2b). The deduced amino acid sequence of *CaNPR1* displayed highest identity with the NPR1 proteins from different Solanaceae including *Capsicum annuum* (ABG38308, 71%), *Nicotiana tabacum* (ABH04326, 71%) and *Solanum lycopersicon* (AAT57637, 71%). Detailed analysis of the deduced amino acid sequence of *CaNPR1* revealed the typical highly conserved features of NPR1 proteins [7, 26-28] (Fig. 2b). Previous studies have demonstrated that residues Cys82, Cys150, Cys155, Cys160 and Cys216 present in the BTB/POZ domain are involved in oligomer-monomer transition [10, 29]. All these 5 cysteine residues are conserved in *CaNPR1* (data not shown). Curiously, an unexpected leucine zipper (LZ) motif was found in the *CaNPR1* protein (Fig. 2b). This domain has never been reported in NPR1 proteins. The LZ are usually found as part of a DNA-binding domain in various transcription factors [30], and are therefore involved in many gene regulatory proteins such as octamer-binding transcription factor 2 (Oct-2/OTF-2) [31], CCATT-box and enhancer binding protein (C/EBP) [32], nuclear oncogene fos and jun [33], cAMP response element (CRE) binding proteins (CREB, CRE-BP1, and ATFs) [34], and C-myc, L-myc and N-myc oncogenes [35]. According to Viedma *et al.* [36], the LZ motif modulates the equilibrium between monomeric and dimeric forms of a protein. Interestingly, this motif was also found in Solanaceae studied herein (data not shown).
2.2. Phylogenetic analysis of NPR1 homologs

Phylogenetic analysis was based on the entire amino acid sequences of NPR1 proteins retrieved from public databases: *S. lycopersicum* (AAT57637.1, Le), *Capsicum annum* (ABG38308.1, Ca), *Nicotiana tabacum* (AAM62410.1, Nt), *Ricinus communis* (EEF48081.1, Rc), *Musa acuminata* (ABL63913.1, Ma), *Oryza sativa* (ABE11613.1, Os), *Hordeum vulgare* (CAJ19095.1, Hv), *Brassica napus* (AAM88865.2, Bn), *Brassica juncea* (ABC94642.2, Bj), *Arabidopsis thaliana* (NP_176610.1), *Musa spp.* (ABB (AJC04030.1, Md), *Zea mays* (NP_001147587.1, Zm), *Glycine max*: GmNPR1-1 (ACJ45013.1) and GmNPR1-2 (ACJ45015.1), *Populus trichocarpa* (XP_002322351.1, Pt), *Malus x domestica* (ACC77697.1, Mp), *Pyrus pyrifolia* (ABK62792.1, Pp), as well the sequences from *A. thaliana* NPR2 (NP_194342.1, At), NPR3 (NP_199324.2, At), NPR4 (NP_193701.2, At), NPR5 or AtBOP1 (ABH04470.1, At), and NPR6 or AtBOP2 (AAU90063.1, At).

Analysis using bootstrap consensus for neighbor joining showed that CaNPR1 grouped with members of the Solanaceae family like tobacco (*N. tabacum*), tomato (*S. lycopersicum*) and pepper (*C. annuum*) (Fig. 3). This suggests a conserved pattern in the NPR1 protein among plant families that belong to the Asterid I clade of dicots [37]. The closer taxonomic affinities of coffee and Solanaceae are also paralleled by a number of striking botanic and genetic similarities, including the production of fleshy berries [38], similar genome content [39] and similar basic number and architecture of the chromosomes [40-42]. Accordingly, plants belonging to the Brassicaceae family (*A. thaliana*, *Brassica napus* and *B. juncea*) formed a separate sub-clade. Within the monocots and dicots, a relative close grouping was obtained between NPR1 proteins from the Asterid I clade of dicots and from monocots belonging to different families (Fig. 3), which is in agreement with reports by Henanff et al. [27] and Zhao et al. [43].

Fig. 3 Evolutionary relationship of NPR1 orthologues. The evolutionary history was inferred using the Neighbor-Joining method [44]. Bootstrap values based on 1000 replicates are indicated beside 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. The evolutionary distances were computed using the Poisson correction method [45] and are in the units of the number of amino acid substitutions per site. The analysis involved 23 amino acid sequences. All ambiguous positions were removed for each sequence pair. There were a total of 734 positions in the final dataset. Evolutionary analyses were conducted in MEGAS5 [46].
2.3. Expression pattern of CaNPR1 during compatible interaction with Hemileia vastatrix

Analysis of rust infection in susceptible coffee varieties showed that the fungus germinates at the leaf surface, penetrates through stomata and produces intracellular infection structure (haustoria) in leaf cells as soon as 24 hours after infection (hai) [47]. Drastic changes in defense gene expression may be observed at 6 to 12 hai [47-48]. To test whether the expression of CaNPR1 is affected by biotic stress, transcript accumulation in coffee plants exposed to H. vastatrix infection was monitored by quantitative real-time PCR during the first 24 hai. No significant changes in CaNPR1 transcript levels were observed (Fig. 4), suggesting that CaNPR1 expression is not regulated at the transcriptional level upon H. vastatrix infection. This is in line with previous studies showing constitutive expression of NPR1 genes under pathogen infection [11, 27, 49].

Fig. 4 Expression pattern of CaNPR1 following infection with H. vastatrix. Leaves of C. arabica cv Mundo Novo were inoculated with H. vastatrix race II (10 mg spores mL-1). Control leaves were sprayed with water. Leaves were collected at different time points as indicated. The log-transformed gene expression levels are presented and indicate normalized expression levels in inoculated leaves compared to normalized expression levels observed in water-treated leaves at the same time point. The GAPDH gene was used for normalization. The two thin dashed lines represent the minimal standard deviation, with a log expression levels difference between 0.5 and -0.5. hpi: hours post inoculation.

2.4. Sequence analysis of CaNPR1 promoter

The nucleotide sequence of the 5′ upstream region of CaNPR1 was identical among 11 clones analyzed, suggesting either that the upstream sequences of the two CaNPR1 copies of the tetraploid C. arabica genome are high similar or that our gene walking strategy impaired cloning of only one of the duplicated sequence.

To assess the set of putative cis-acting elements present in the CaNPR1 promoter, named pCaNPR1, an in silico search was performed in PLACE using the 1444-bp sequence located upstream of the start codon as query. In this analysis we looked for TATA-box consensus sequences and for consensus sequences described as binding sites for defense-related transcription factors. A TATA-box was identified 496 nucleotides upstream of the start codon. The presence of CAAT motifs that are generally associated with gene regulation was also observed. The CAAT-boxes are detected in the core promoter sequences important in binding RNA Polymerase II [50]. Additional important regulatory elements detected in the CaNPR1 promoter were those related to defense as well as ELRECOREPCRP1 [51], MYBPLANT [52], T/GBOXATPIN2 [53], GT1GMSCAM4 [54], PALBOXAPC [55] and W-box [56], suggesting a biotic regulation [57-62]. Moreover, several cis-elements reported to be associated with the responsiveness to plant hormones, such as salicylic acid (ASFI1MOTIFCAMV), auxin (NTBBF1ARROLB), jasmonic acid (JERE), citokin (CPBCSPOR), ethylene (ERE, WBOXNTERF3), abscisic acid (MYB1AT, MYCCONSENSUSAT, MYB2CONSENSUSAT, WRKY71OS, DPBF COREDCDC3, LTRECOREATCOR15, ABRELATERD1), as well as gibberellin (WRKY71OS, PYRIDINEBOXOSRAMY1A, TATCCACHVAL21, CAREOSREPI) were observed. The presence of such large number of cis-elements suggests that CaNPR1 expression might be regulated by hormonal action.

2.5. CaNPR1 promoter activity is SA-dependent

GUS expression levels driven by the CaNPR1 promoter was evaluated using a transient expression assay in leaves of N. benthamiana and compared to that obtained using the CaMV35S promoter. Transient GUS activity was assayed fluorometrically at 3 day post-infiltration (Fig. 5). It was observed that the level of GUS activity driven by the CaNPR1 promoter was very weak compared to that driven by the 35S promoter. As previously mentioned, the NPR1 gene is a key regulator of the SA-mediated SAR pathway [3, 13, 63-64] and its expression levels is slightly induced in A. thaliana upon foliar SA treatment [7, 65]. Given that the upstream promoter region of CaNPR1 harbors cis-acting elements implicated in plant hormone responses, we examined whether this promoter showed a SA-inducible expression pattern. As a result, a significant alteration in the level of GUS activity was detected in SA-treated leaves as compared to control leaves treated with water (Fig. 5). These data demonstrate that SA strongly stimulates the transcriptional activity of the CaNPR1 promoter.
3. Taken to acquire knowledge

*Coffea arabica* cv. Mundo Novo (IAC 388-17-1) and *C. arabica* cv. Caturra were obtained from IAC (Instituto Agrônômico de Campinas - Campinas, São Paulo, Brazil). Plants were grown in potting soil for 4 months (16h/8h light/dark; 28°C; 70% RH) before use for leaf DNA extraction (*C. arabica* cv. Caturra) or pathogen inoculation (*C. arabica* cv. Mundo Novo). Plants of *Nicotiana benthamiana* were grown in potting soil under a 16/8h photoperiod for 6 weeks prior to being used for *Agrobacterium* infiltration. *Hemileia vastatrix* race II was kindly provided by Dr. Maria do Céu Silva (CIFC, Centro de Investigação das Ferrugens do Cafeeiro - Oeiras, Portugal).

For biotic stress assays, 4-month old *Coffea arabica* cv. Mundo Novo plants were kept in a growth chamber (16h/8h light/dark; 23°C; 70% RH) for at least one week, before being inoculated with the coffee leaf rust fungus *H. vastatrix* race II, that elicits a compatible reaction in coffee. The urediniospores (100 mg) were diluted in 10 mL of sterile water under dark conditions. Leaves from the second pair of plagiotropic shoots from the apex of coffee plants were inoculated with urediniospores. To allow spore germination, the inoculated leaves were covered with a wet black plastic film for 24h. The inoculated leaves were not detached from the plants. Mock-inoculated controls as well as non-inoculated controls were performed. The biological samples were obtained from three independent experiments. Leaves were randomly sampled at different time-points after inoculation: 0, 8, 12, and 24 hours, and immediately deep-frozen. To confirm the infection by the leaf rust fungus, some inoculated leaves were maintained in the plants.

RNA extraction, reverse transcription, and quantitative PCR experiments were performed as described in Barsalobres-Cavallari et al. [66]. The specific oligonucleotides used for gene expression studies are shown in Table 1.

**Table 1** Sequence of primers used for quantitative PCR in coffee.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CaNPR1 F</td>
<td>TGGACGAAGGAGCGTGATGTCAGCTT</td>
</tr>
<tr>
<td>CaNPR1 R</td>
<td>ACATCCTTGGCGACGGCCTAACTCT</td>
</tr>
<tr>
<td>CaGAPDH F</td>
<td>TTGAAGGGCGGTGCAA</td>
</tr>
<tr>
<td>CaGAPDH R</td>
<td>AACATGGTGTCATCCTTGGCT</td>
</tr>
</tbody>
</table>

The genomic DNA from *C. arabica* cv. Caturra was isolated using the DNeasy Plant Minikit (Qiagen, France) following the manufacturer’s recommendations. Gene walking by PCR was performed as described in Petitot et al. [67]. The promoter region of CaNPR1 was amplified using nine GSPs (termed CaNPR1-1 to CaNPR1-9), where the first two (CaNPR1-1 and CaNPR1-2) were designed based on the EST sequence retrieved in the coffee database, while the others were designed based on the nucleotide sequences of the genome walking amplified fragments (Table 2). All the amplified genomic fragments were cloned into the pGEM-T Easy vector (Promega) and sequenced (Genome Express, France).
Table 2 Primers used for genome walking.

<table>
<thead>
<tr>
<th>PRIMER NAME</th>
<th>GENE-SPECIFIC PRIMERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>CaNPR1-1</td>
<td>5'CATACGCGAGATCATCGGCTGCCATAGC3'</td>
</tr>
<tr>
<td>CaNPR1-2</td>
<td>5'CCCCGCGAGTGTCGAGAGGAGTCCCCTGCG3'</td>
</tr>
<tr>
<td>CaNPR1-3</td>
<td>5'GGCCCTGCAAACCAATAGTGGGCGTCCCTGCG3'</td>
</tr>
<tr>
<td>CaNPR1-4</td>
<td>5'CATCTGCAGGACCTTTTGAGAAGACG3'</td>
</tr>
<tr>
<td>CaNPR1-5</td>
<td>5'GGAAAGATTAGTGGGCGTCCCTGCG3'</td>
</tr>
<tr>
<td>CaNPR1-6</td>
<td>5'CCAAAGCTCAGGAGAGCACGAGC3'</td>
</tr>
<tr>
<td>CaNPR1-7</td>
<td>5'GGGGAGGGGCTTGAGTGAGACGACG3'</td>
</tr>
<tr>
<td>CaNPR1-8</td>
<td>5'GACGACTGAACTGAGCCGGAGC3'</td>
</tr>
<tr>
<td>CaNPR1-9</td>
<td>5'CGGTCGATCGGCTAGGAAGACTGAGC3'</td>
</tr>
</tbody>
</table>

The DNA sequences were analyzed with the BLAST network services at the National Center for Biotechnology Information [68]. The software tool TSSP was used for the prediction of the transcription start site, RNA polymerase II binding site and TATA-box [69], and the PLACE database were used for the prediction of cis-acting elements [70]. For protein analysis, it was used the follows software: the Pfam database [71], the PSITE from Softberry [72] and the Bioinformatics Toolkit from Max-Planck Institute for Developmental Biology [73].

For the phylogenetic and molecular evolutionary analyses, the neighbor joining method with 1000 bootstraps was used. The NPR1 amino acid sequences used were: Solanum lycopersicum (AAT57637.1, Le), Capsicum annuum (ABG38308.1, Ca), Nicotiana tabacum (AAM62410.1, Nt), Ricinus communis (EEF48081.1, Rc), Musa acuminata (ABL63913.1, Ma), Oryza sativa (ABE11613.1, Os), Hordeum vulgare (CAJ19095.1, Hv), Brassica napus (AAM88652.2, Bn), Brassica juncea (ABC94642.2, Bj), Arabidopsis thaliana (NP_176610.1), Musa spp. ABB (ACJ04030.1, Md), Zea mays (NP_001147587.1, Zm), Glycine max: GmNPR1-1 (ACJ45013.1) and GmNPR1-2 (ACJ45015.1), Populus trichocarpa (XP_002322351.1, Pt), Malus x domestica (ACC77697.1, Mp), Pyrus pyrifolia (ABK62792.1, Pp), as well as the sequences from A. thaliana NPR2 (NP_194342.1, At), NPR3 (NP_199324.2, At), NPR4 (NP_193701.2, At), NPR5 or AtBOP1 (AH04470.1, At), and NPR6 or AtBOP2 (AAU90063.1, At).

The construction of β-glucuronidase (GUS) reporter cassettes, as well as the Agrobacterium-mediated transient expression assays, abiotic treatments and the GUS activity measurement were performed as described in Petitot et al. [67] and Barsalobres-Cavallari [74].

4. Conclusions

We have isolated the orthologous sequence of the Arabidopsis SAR controlling master switch NPR1 gene in coffee, that we called CaNPR1. Typical features of NPR1 proteins were found in CaNPR1, including the BTB/POZ domain, an ankyrin (ANK) repeat domain, and a nuclear localization signal (NLS). As expected from A. thaliana, soybean and grapevine results [8, 27, 49], transcript levels of CaNPR1 were not affected following pathogen infection. Interestingly, the activity of the CaNPR1 promoter was found to be strongly activated by SA-treatment. The different cis-acting elements identified in the pCaNPR1 sequence, such as the plant hormone- and/or infection-related motifs, might contribute to SA responsiveness. From these results, it is tempting to speculate that the mechanisms underlying C. arabica resistance to the rust pathogen are SA-dependent, as already suggested by other molecular studies [47, 75]. Our data also indicate that CaNPR1 may represent a promising candidate for engineering resistance in coffee to achieving broad spectrum resistance to pathogens. Further functional studies are required for assessing the role of this gene in coffee disease resistance responses. We anticipate that transgenic plants containing CaNPR1 promoter-reporter constructs will provide tools for the future analysis of defense signaling networks in coffee.

Acknowledgements The authors thank Dr. Mirian Perez Maluf (EMBRAPA/IAC) for providing the plant materials used in this study. We are greatly indebted to the research center on coffee rusts (CIFC, Oeiras, Portugal) for providing rust isolates. CFB-C was recipient of a fellowship from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil, and also of a scholarship from the Colégio Doutoral Franco-Brasileiro from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil. FES was recipient of a fellowship from CAPES, Brazil. This project was supported by Consórcio Brasileiro de Pesquisa e Desenvolvimento do Café (CBP&D-Café), Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), CAPES, Comité Français d’Évaluation de la Coopération Universitaire et Scientifique avec le Brésil (COFECUB-project Sv 555/07) and the Institut de Recherche pour le Développement (IRD).
References


© FORMATEX 2013


