NEP1 orthologs encoding necrosis and ethylene inducing proteins exist as a multigene family in Phytophthora megakarya, causal agent of black pod disease on cacao

Hanhong BAE1, John H. BOWERS1, Paul W. TOOLEY2 and Bryan A. BAILEY1

1 USDA/ARS, Plant Sciences Institute, Sustainable Perennial Crops Laboratory, Beltsville Agricultural Research Center-West, Beltsville, MD 20705, USA.
2 USDA/ARS, Foreign Disease-Weed Science Research Unit, Ft. Detrick, MD 21702-5023, USA.
E-mail: rbae@asrr.arsusda.gov

Received 16 December 2004; accepted 20 July 2005.

Phytophthora megakarya is a devastating oomycete pathogen that causes black pod disease in cacao. Phytophthora species produce a protein that has a similar sequence to the necrosis and ethylene inducing protein (Nep1) of Fusarium oxysporum. Multiple copies of NEP1 orthologs (PmegNEP) have been identified in P. megakarya and four other Phytophthora species (P. citrophthora, P. capsici, P. palmivora, and P. sojae). Genome database searches confirmed the existence of multiple copies of NEP1 orthologs in P. sojae and P. ramorum. In this study, nine different PmegNEP orthologs from P. megakarya strain Mk-1 were identified and analyzed. Of these nine orthologs, six were expressed in mycelium and in P. megakarya zoospore-infected cacao leaf tissue. The remaining two clones are either regulated differently, or are nonfunctional genes. Sequence analysis revealed that six PmegNEP orthologs were organized in two clusters of three orthologs each in the P. megakarya genome. Evidence is presented for the instability in the P. megakarya genome resulting from duplications, inversions, and fused genes resulting in multiple NEP1 orthologs. Traits characteristic of the Phytophthora genome, such as the clustering of NEP1 orthologs, the lack of CATT and TATA boxes, the lack of introns, and the short distance between ORFs were also observed.

INTRODUCTION

Black pod of cacao is caused by Phytophthora species and can result in complete yield loss under some conditions (Hoopen et al., 2001, Bennett 2003). This loss of yield has a devastating effect on local agricultural economies. Several Phytophthora species, such as P. megakarya, P. palmivora, P. citrophthora, and P. capsici, attack the tropical tree Theobroma cacao (cacao) causing seedling blights, stem cankers, and pod rots. P. megakarya is the most aggressive among the four species on cacao and attacks all plant parts.

The necrosis and ethylene inducing protein 1 (Nep1), is a 24 kDa extracellular protein that is produced by Fusarium oxysporum in liquid cultures (Bailey 1995, Bailey, Jennings & Anderson 1997). Nep1 causes necrosis in dicotyledonous plants, but is inactive in monocotyledonous plants (Bailey 1995, Jennings et al. 2000, Veit et al. 2001). Nep1-like proteins (NLPs), such as necrosis-inducing protein PsojNIP (GenBank accession no. AAK01636, isolated from P. sojae), necrosis-inducing Phytophthora protein 1 NPP1 (AAK19753, isolated from P. parasitica), and necrosis-inducing protein NIP1 (AAK25828, isolated from P. infestans), range from 24 to 26 kDa and have been identified in many different plant pathogens including fungi and bacteria (Veit et al. 2001, Fellbrich et al. 2002, Qutob, Kamoun & Gijzen 2002, Pemberton & Salmond 2004). Nep1 and NLPs cause induction of pathogen-related (PR) genes, K+ and H+ channel fluxes, callose apposition, accumulation of reactive oxygen species (ROS) and ethylene, altering cell respiration, creating a hypersensitive response (HR), and localized cell death (Jennings et al. 2001, Veit et al. 2001, Fellbrich et al. 2002, Keates et al. 2003). Nep1 also induced the breakdown of the cuticle, chloroplast envelope and internal membrane structures after 1–4 h of treatments in Arabidopsis thaliana, spotted knapweed (Centaurea maculosa), and dandelion (Taraxacum officinale) (Keates et al. 2003). In addition to disease resistance genes, Nep1 treatment induced genes putatively involved in plant stress responses, including wounding, drought and senescence (Keates et al. 2003). When Nep1 protein was applied in combination with the pathogen Pleospora papaveracea, it enhanced disease development and death of opium poppy (Papaver somniferum; Bailey et al. 2000). Qutob et al. (2002) demonstrated that PsojNIP was produced by P. sojae, a
pathogen of soybean (Glycine max), during the necrotrophic phase of the disease. Therefore, NLPs might function as necrosis inducing toxins rather than inducers of functional plant resistance responses (Jennings et al. 2000).

In P. sojae, two additional predicted open reading frames (PsojORF II and PsojORF III) related to PsojNIP were found clustered with PsojNIP within a 3515-bp region (AF511649; Qutob et al. 2002). PsojNIP and PsojORF II did not have introns, while the PsojORFIII had a 25-bp intron. The three ORFs had shared identities between 67% and 69%. The size of two intergenic sequences between the three ORFs was 674 bp and 713 bp, respectively. According to RT-PCR, only PsojNIP (AK01636) was transcriptionally active. Alignment of NLPS of Phytophthora species revealed that the encoded proteins were highly similar to each other (69–89% identity). However, the sequence identity to Fusarium oxysporum Nep1 was relatively low (34–37%).

This study was initiated to characterize the gene family encoding NLPS in Phytophthora species causing black pod, a disease that severely reduce cacao yields around the world. NEP1 orthologs were cloned and characterized in the P. megakarya strain Mk-1. Multiple copies of NEP1 orthologs in P. megakarya were identified, as well as in four other Phytophthora species. Of nine NEP1 orthologs cloned, seven of them were transcriptionally active in P. megakarya, both in mycelia and during infection of the cacao leaf.

**MATERIALS AND METHODS**

**Plant growth and inoculation of leaf disks**

Seeds of Theobroma cacao cv. ‘Comun’ (Lower Amazon Amelonado type) were planted in pots filled with a soilless mix (2:2:1, sand:perlite:promix). Seedlings were grown in a greenhouse up to 9 months with ambient light, 75% humidity, and temperatures of 20-29 °C. Leaf development was separated into four stages: (1) unexpanded leaves (UE), less than 1 cm long with limited pigmentation; (2) young red leaves (YR), 5-10 cm long and pliable; (3) immature green leaves (IG), 10–20 cm long and pliable; and (4) mature green leaves (MG), 10–20 cm and rigid. For the zoospore infection study, leaves that were between IG and MG stage were used.

The cacao leaf disks were inoculated with zoospores of P. megakarya strain Mk-1 (Pierre Tondje, IRAD, Youande, Cameroon) that were synchronously produced in axenic culture (Bowers & Mitchell 1991). Identification of the P. megakarya strain Mk-1 was confirmed by morphology and phylogenetic analysis of the ribosomal DNA ITS region (data not shown). Leaf disks (85 mm) were cut from detached leaves and placed abaxial side up on a moist No. 2 Whatman paper in Petri dishes (100 × 15 mm). The disks were inoculated with multiple drops of 20 μl of water containing zoospores of P. megakarya (3 × 10⁶ zoospores ml⁻¹), and control leaf disks were treated with multiple 20 μl drops of distilled water. The Petri dishes were sealed with parafilm and incubated at room temperature for up to 3 d under ambient light conditions. Three replications of each Petri dish were harvested 3 d after inoculation.

**PCR and RACE (Rapid Amplification of cDNA Ends)**

DNA and RNA were extracted from Phytophthora megakarya mycelia grown in clarified V8 (Campbell Soup Company, Camden, NJ) liquid culture (Martin & Tooley 2003) using DNaseasy Plant Maxi Kit and RNeasy Mini Kit, respectively (Qiagen, Valencia, CA). Total RNA from P. megakarya zoospore-infected and control leaf disks was extracted by the adaptation of several established methods including those of Chang, Puryear & Cairney (1993), McKenzie et al. (1997), and the RNeasy Mini Kit (Qiagen) (Verica et al. 2004). The extracted RNA was treated with DNase I (Roche, Indianapolis, IN) according to manufacturer’s recommendation and was quantified using a spectrophotometer and the integrity was verified with agarose gels.

To amplify the transcriptionally active NEP1 ortholog, 2 μg of DNase I treated total RNA, extracted from zoospore-infected cacao leaf disks, were used to generate cDNA using the QIAGEN OneStep RT-PCR Kit (Qiagen) with 11F/16R primer set. Primer sequences were designed based on the consensus sequence of the NEP1 orthologs of P. sojae (AK01636), P. infestans (AK25828), and P. parasitica (AAK19753) (11F 5’- AACGGCTGCCACCCGTACCCTGC-3’, 16R 5’- CGTGAG(T/C)TGGTCCCACATGATCAG-3’). Transcripts were amplified using the following conditions: reverse transcription at 50 °C for 30 min, initial activation of HotStar Taq DNA polymerase and inactivation of the reverse transcriptase at 95 °C for 15 min, denaturation for 30 s at 95 °C, annealing for 30 s at 57 °C, extension for 1 min at 72 °C, and an additional 10 min extension at 72 °C, 40 cycles. The PCR product was electrophoresed in 0.8% agarose gel, purified by using QIAEX II Gel Extraction System (Qiagen), cloned into pCR2.1-TOPO vector for small inserts or pCR-XL-TOPO vector for large inserts (Invitrogen, Calsbad, CA), and prepared for sequencing. DNA sequences were determined by the University of Maryland, DNA Sequencing Facility (http://www.umbi.umd.edu/~ecbr/dna.html). Half μg of mycelial DNA was also used as a template for PCR amplification. This was performed using the following conditions: initial activation of Taq DNA Polymerase at 94 °C for 5 min, denaturation for 30 s at 94 °C, annealing for 30 s at 55 °C, extension for 1 min at 72 °C, and an additional 10 min extension at 72 °C, with 30 cycles. In order to clone the full length of NEP1 orthologs, both 5’ and 3’ RACE were performed using the RACE system (Invitrogen).

Semiquantitative RT-PCR was used to detect the transcripts of NEP1 orthologs in P. megakarya mycelia and in P. megakarya zoospore-infected cacao leaf
tissues. Eight sets of *PmegNEP* ortholog-specific primers are listed in Table 1. Four μg of DNase I treated total RNA was used to generate the first strand complementary DNA (cDNA) using the oligo (dT)20 primer. The first strand cDNA was then synthesized from the total RNA, using SuperScript II RNase H+ reverse transcriptase, according to the manufacturer’s instructions (Invitrogen). Then the first strand cDNA was used as a template for semiquantitative RT-PCR. This was performed using the following conditions: initial activation of Taq DNA Polymerase at 94° C for 5 min, denaturation for 30 s at 94°, annealing for 30 s at 55°, extension for 1 min at 72°, and an additional 10 min extension at 72°, with 30 cycles for *PmegNEP* orthologs and 22 cycles for *P. megakarya* ACTIN. Mycelial DNA was used as a PCR template to verify the existence of nine *PmegNEP* ortholgs in the genome of *P. megakarya* using the same set of ortholog-specific primers mentioned above (Table 1). The primer sequences for 459-bp ACTIN fragment were designed based on *P. infestans* (M59715) and was used as a PCR control. RT-PCR was repeated at least three times with different RNA extractions from separate plants or sources of *P. megakarya* mycelia.

DNA and RNA blot analyses

For Southern blot analysis, 2.5 μg of the mycelial DNA was digested with restriction enzymes *Bgl*I and *Pst*I at 37° overnight, separated in a 0.8% agarose gel, blotted onto a membrane, and hybridized with the 32P-labeled *PmegNEP*. Northern blot analyses were repeated two times on separate samples to confirm the reproducibility of the results. Blots were washed and imaged at 200 μm resolution on a Typhoon 8600 Variable Mode Imager (Molecular Dynamics/Amersham-Pharmacia Biotech, Sunnyvale, CA). Probe hybridization and washing blots were performed under medium stringency condition according to Keates et al. (2003). Three other *Phytophthora* species pathogenic on cacao (Edna Luz, CEPLAC, Itabuna, Brazil) were used for Southern blot analyses: *P. citrophthora* 62E, *P. capsici* 197E and *P. palmivora* 252E.

Data analyses

DNA and protein sequence analyses were performed with Biology WorkBench (http://workbench.sdsc.edu/). Genome database searches of *Phytophthora sojae* and *P. ramorum* were performed with *Phytophthora* Molecular Genetics Research Collaboration Network (http://pmgn.vbi.vt.edu/). The sequences were examined for signal peptides using SignalP v. 3.0 (http://www.cbs.dtu.dk/services/SignalP/). The degree of glycosylation of PmegNep clones was assessed using motif search programs NetOGlyc 3.1 Server (http://www.cbs.dtu.dk/services/NetOGlyc/). A phylogenetic tree was constructed from amino acid sequences, using neighbor-joining method of ClustalW with the number of differences model as implemented in Mega2 (http://www.megasoftware.net/).

RESULTS

Multiple copies of NEP1 orthologs detected

The universal primer set (11F and 16R) was used for PCR and the PCR products were cloned into

| Table 1. Ortholog specific primer sequences for nine *PmegNEP* clones and the ACTIN gene. F, forward primer: R, reverse primer. |
|--------------------------|-----------------------------------------------|------------------------|
| Clone name (GenBank accession no.) | Sequences (5’ to 3’) | Expected size (bp) |
| *PmegNEP1* (AY741082) | F: TTCGACCTGGTATAACGGA | 300 |
| R: CGTCTCACCACTATCCGAT | |
| *PmegNEP2* (AY741083) | F: CACGTACAAAGGGGTCTAC | 300 |
| R: GGTCTCTCCCACATCTGTG | |
| *PmegNEP3* (AY741084) | F: GAAAGGAGCCCTAGTGCTA | 250 |
| R: CAGAATCTTGGGGTTTTCA | |
| *PmegNEP4* (AY741085) | F: AGGCCCAACCGGTTCTGGAG | 250 |
| R: CAGATTTTGGCCGCTGCAAC | |
| *PmegNEP4-2* (AY741086) | F: GACTCGGGTCCTAGTTCA | 250 |
| R: GATAATCTTCGGCTCCCGGA | |
| *PmegNEP5* (AY741087) | F: AGTGATCCCTGAGCCCTT | 320 |
| R: ATGATATCTTCGGCTCCCGGA | |
| *PmegNEP6* (AY741088) | F: CAACATGCACATGAAAGCCTA | 320 |
| R: ATTACATCTTCGAGCCTC | |
| *PmegNEP6-2* (AY741089) | F: TACCAATGACATGGAGGT | 245 |
| R: CCAGACGTCAATGGA | |
| *PmegNEP7* (AY741090) | F: CAAGGAGAAGCTGGACGTACA | 457 |
| R: TGTACGTCAAGCTTTTCTTCTTGG | |
Table 2. Percent identity between amino acid sequences of PmegNep clones. Partial amino acid sequences of 11F-16R PCR products were used to compare PmegNep clones. For PmegNep6-2 and PmegNep7, partial sequences were used to compare with other PmegNep clones due to lack of sequence information.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Nep1c</td>
<td>71</td>
<td>62</td>
<td>65</td>
<td>64</td>
<td>50</td>
<td>66</td>
<td>68</td>
<td>57</td>
</tr>
<tr>
<td>Nep2</td>
<td>59</td>
<td>64</td>
<td>64</td>
<td>42</td>
<td>64</td>
<td>64</td>
<td>67</td>
<td>46</td>
</tr>
<tr>
<td>Nep3</td>
<td>52</td>
<td>53</td>
<td>64</td>
<td>64</td>
<td>52</td>
<td>56</td>
<td>63</td>
<td>63</td>
</tr>
<tr>
<td>Nep4</td>
<td>96</td>
<td>42</td>
<td>81</td>
<td>83</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>Nep4-2</td>
<td>43</td>
<td>79</td>
<td>82</td>
<td>82</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>Nep5</td>
<td>43</td>
<td>51</td>
<td>57</td>
<td>57</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>Nep6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>44</td>
</tr>
<tr>
<td>Nep6-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>44</td>
</tr>
</tbody>
</table>

Fig. 1. Schematic diagrams of the two clusters of PmegNep clones amplified from DNA using PCR with the 11F primer (A), and PmegNep1 Forward and PmegNep6 Reverse primers (B). Each cluster contained two partial NEP1 orthologs and one full length NEP1 ortholog. The size of the PCR product does not include a 23-bp primer sequence on both ends.

pCR2.1-TOPO cloning vector for sequencing. According to sequence results, three different PCR products were identified from two different PCR templates (DNA and RNA from mycelia). The size of PCR fragments were 400, 400 and 397 bp without primer sequences. The three partial NEP1 orthologs were named as P. megakarya necrosis and ethylene inducing proteins, PmegNep1 (AY741082), PmegNep2 (AY741083), and PmegNep3 (AY741084), respectively (Table 2). The deduced amino acid sequences of PmegNep1, 2 and 3 shared identities between 53 and 84% with database sequences of three Nep1-like proteins (NLPs) in Phytophthora species (P. sojae, AAK01636; P. infestans, AAK25828; P. parasitica, AAK19753).

The same primer set (11F and 16R) was used to amplify additional NEP1 orthologs in the genome of P. megakarya with longer extension time (10 min instead of 1 min at 72° C). More than one DNA PCR products were detected on the gel. A 2390-bp (2344 bp without primer sequences) PCR fragment was cut out of the gel and was cloned into pCR-XL-TOPO vector for sequencing. Sequencing results revealed that the 2390-bp PCR fragment was only amplified by the 11F primer, and contained an additional three NEP1 orthologs (Fig. 1A). The fragment contained two partial copies and one full length copy of NEP1 orthologs, and the three orthologs were named as PmegNep4 (AY741085), PmegNep6 (AY741088) and PmegNep5 (AY741087), respectively. Because one NEP1 ortholog (PmegNep5) was inverted, one primer could anneal to 5’ and 3’ regions resulting in the amplification of 2390-bp PCR fragment. To verify the existence of the 2390-bp PCR product in the genome of P. megakarya, ortholog specific primer set for PmegNep1 Forward and PmegNep6 Reverse were designed (Table 1) and
PCR was performed with mycelial DNA. The PCR products around 2.4 kb were cloned into pCR-XL-TOPO and more than one colonies were picked for sequencing. The existence of the 2390-bp PCR product was confirmed by sequence result. In addition, a 2353-bp PCR product (without primer sequences) that contained new NEP1 orthologs was detected. The 2353-bp PCR fragment contained three NEP1 orthologs (Fig. 1B). The fragment contained two partial and one full length copies of NEP1 orthologs. The first partial copy was the same as the previously detected PmegNEP1. However, the full length and the other partial length copies were not the same as previously detected copies. The two copies had 96% identity with the amino acid sequences of PmegNep4 and PmegNep6, and were named as PmegNep4-2 (AY741086) and PmegNep6-2 (AY741089), respectively.

For 5’ RACE, primers were designed based on the three PmegNEP consensuses sequences (1, 2 and 3) in the middle of the PmegNEP sequences. During the process of 5’ RACE, a ninth clone of a NEP1 ortholog (PmegNEP7, Accession No. AY741090) was identified. The 525-bp 5’ RACE product predicted to encode 175 amino acids and consisted of two partial genes that were fused together. The N-terminal region of the 5’ RACE product matched with the C-terminal region of the unknown functional protein of SAV7207 of Streptomyces avermitilis MA-4680 (NP828383). The sequence of the 5’ RACE product ranging from amino acids 4 to 74 corresponded to the Jacalin-like lectin domain of SAV7207, and the sequence between amino acids 94 to 175 matched the partial sequences of the necrosis-inducing-like protein of P. sojae (AAM48171).

An alignment of the deduced amino acid sequences of the partial NEP1 orthologs (region corresponding to the primer set of 11F and 16R) revealed that the nine PmegNEP clones were different from each other (Table 2). The deduced amino acid sequences of nine clones shared identities between 42 and 96% with each other. The existence of the nine clones of NEP1 orthologs in the P. megakarya genome was confirmed by conducting PCR using P. megakarya mycelial DNA as a template, and the ACTIN gene as a control (Fig. 3). PCR with P. megakarya DNA amplified the expected size band for each PmegNEP clone, confirming the existence of nine clones in the P. megakarya genome including the fused gene PmegNEP7. PCR using DNA from other Phytophthora species showed different band...
patterns or was unable to amplify any fragment (data not shown). To test whether each PmegNEP clone was transcriptionally active, RT-PCR was performed (Fig. 3). Total RNA from P. megakarya zoospore-infected cacao leaf tissue and P. megakarya mycelia were used as templates for RT-PCR. The common primer set for PmegNEP4 and PmegNEP4-2 was used, while clone specific primer sets were used to amplify PmegNEP6, PmegNEP6-2, and the remaining orthologs (Table 1). Six PmegNEP clones (1, 2, 3, 5, 6, 6-2) were transcriptionally active in leaf during infection (3 days after inoculation) and in mycelial tissue. RT-PCR products of PmegNEP1, 2, 6 and 6-2 had a greater intensity than the intensity of PmegNEP3 and 5 in mycelial tissue. On the other hand, in P. megakarya zoospore-infected tissue, the intensity of PmegNEP1, 2, 6 and 6-2 was higher than any other PmegNEP clones. For PmegNEP7, we tried three different primer sets that covered: (1) the two fused genes (forward primer 5’-AAGTACCAAGAGTACATTAC-3’; reverse primer 5’-CCAGACGTCATTTGAACCAAG-3’); (2) the Jacalin-like lectin domain of SAV7207 (forward primer 5’-GTACATTACGTCGATGGAAG-3’; reverse primer 5’-CAAATCGCACCGATGAGGTC-3’); and (3) the PsojNIP domain (PmegNep7, primers sequences in Table 1). When the primer set that covered the entire region of two fused genes (SAV7207 + PsojNIP) was used, transcript was not detected by RT-PCR. However, the transcript for PmegNEP7 was detected when the primer set that covered only the NLP domain was used. The transcript of the Jacalin-like domain of SAV7207 was not detected by RT-PCR. The PCR fragments from DNA and RT-PCR were the same size indicating there were no introns in PmegNEP orthologs.

In order to determine the genomic characteristics of the NEP1 orthologs, Southern blot analysis was conducted. The genomic DNA of five Phytophthora species was digested with two restriction enzymes, BglII and PstI, and was probed with PmegNEP2. Multiple bands with different patterns were detected in all species tested, confirming that there were multiple copies of NEP1 orthologs (Fig. 4). Band intensities varied which may be due to the length of the hybridized region, the degree of homology to the probe, or to the presence of more than one band.

**Phytophthora megakarya Nep1 clones share high homology with NLPs of other organisms**

Recently, a number of NLPs have appeared in public databases, including representatives from various oomycete and bacterial species. A phylogenetic tree and the deduced amino acid sequence alignment of seven PmegNEP clones containing the region that corresponded with the primer set 11F and 16R, and 14
other NLPs were generated (Figs 5–6). The alignment of the partial deduced amino acid sequences of PmegNEP clones revealed that seven PmegNEP clones and 14 NLPs of the Phytophthora species clustered together (Fig. 5). The percentage identities between the deduced amino acid sequence of PmegNep1 and other NLPs ranged from 24–93%. The highest matches existed with NLPs of Phytophthora, 93% between PmegNep1 and P. sojae (AAM48171, necrosis-inducing-like protein), 87% between PmegNep2 and P. sojae (AAM48172), 84% between PmegNep2 and P. parasitica (AAK19753), and 82% between PmegNep2 and P. infestans (AAK25828, necrosis-inducing protein). The identities between Fusarium oxysporum (AAC97382) and the deduced amino acid sequences of PmegNEP clones were relatively low (34 to 39%).

The deduced sequence encoding the full length of PmegNep4-2 and PmegNep6 precursor proteins consisted of 235 and 233 amino acids, respectively. Algorithms for the prediction of secondary structures did not reveal any characteristic domains within the two deduced amino acid sequences of PmegNEP clones (data not shown). Two cysteine residues were present in all of the deduced amino acid sequences of PmegNEP orthologs (C55 and C81, C56 and C82, C54 and C80, and C54 and C80 of the pre-protein in PmegNep2, PmegNep4-2, PmegNep6, and PmegNep6-2, respectively); although PmegNep4, PmegNep4-2, PmegNep6, and PmegNep6-2 were distinguished by the existence of the third and/or fourth cysteine residues (e.g. cysteine 126 of PmegNep6) of the pre-protein (Fig. 6). The third cysteine was also detected in PsojNIP at position 128 (AAK01636; Qutob et al. 2002). The GC contents of PmegNep4-2 and PmegNep6 were 54.7% and 50.9%, respectively, which were lower than other Phytophthora NLPs, such as 56.9% for P. infestans, 59.3% for P. sojae, and 56.6% for P. parasitica.

Using SignalP v.3.0, signal peptide sequences were analyzed for PmegNep2, 4-2, 6, 6-2 and other NLPs of Phytophthora species (Bendtsen et al. 2004) (Table 3). Nineteen to 17 amino acid sequences were predicted as signal peptides for the PmegNep clones, while nineteen signal peptides were predicted for other Phytophthora NLPs.

Transposable-like sequences are abundant in the Phytophthora genome. A copia-like retrotransposable element was amplified during the PCR process with PmegNEP3 Forward and PmegNEP5 Reverse primers,

Fig. 5. Phylogenetic analysis of the partial sequences of Nep1 orthologs from oomycetes and other organisms. The 11F and 16R region of Nep1 orthologs was used to create a phylogenetic tree. The tree was constructed from amino acid sequences by the neighbor-joining method of ClustalW and Mega2 (http://www.megasoftware.net/). Numbers at the nodes are bootrap values from 100 replications and represent weighted amino acid substitutions, with the scale bar represents 10% weighted sequence divergence. GenBank accession numbers are shown to the right of the organism.
yielding a product size of 1,283 bp (Accession No. AY741092). However, no evidence of recombination of \textit{PmegNEP} sequence with the transposable element was found. In other words, the sequence of the \textit{NEP1} ortholog was not found in the PCR fragment.

\textbf{Multiple \textit{Nep1} orthologs in \textit{Phytophthora sojae} and \textit{P. ramorum}}

Recently, full genome sequence information has been made available for \textit{Phytophthora sojae} and \textit{P. ramorum}. Searches of the genome databases of \textit{P. sojae} and \textit{P. ramorum} for "necrosis-inducing protein" detected many putative NLPs (\textit{Phytophthora} Molecular Genetics Research Collaboration Network, http://pmg.vbi.vt.edu/). Identical amino acids are highlighted with black. Dashes indicate gaps that were introduced to optimize the alignment. The arrow head indicates the conserved cysteine residue (C80) and a third cysteine residue (C128) of \textit{PmegNep4-2}. GenBank accession numbers are shown in parenthesis to the right of the organism.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{alignment.png}
\caption{Alignment of the partial amino acid sequences of \textit{Nep1} orthologs from oomycetes and other organisms. Alignment of 11F and 16R region of \textit{Nep1} orthologs was performed using ClustalW and BOXSHADE sequence alignment program (http://seqtool.sdsc.edu/CGI/BW.cgi#!). Identical amino acids are highlighted with black. Dashes indicate gaps that were introduced to optimize the alignment. The arrow head indicates the conserved cysteine residue (C80) and a third cysteine residue (C128) of \textit{PmegNep4-2}. GenBank accession numbers are shown in parenthesis to the right of the organism.}
\end{figure}
known PsojNIPs in the phylogenetic tree, and were selected for alignment with the known PsojNIPs and PmegNep6 (Fig. 7). In a search of the P. ramorum genome, 61 putative NLPs matched with a necrosis-inducing protein. 30 of the 61 putative necrosis-inducing proteins shared over a 50% identity with 3 previously reported PsojNIP (AAK01636, AAM48171 and AAM48172) or PmegNep6 amino acid sequences. Nine putative necrosis-inducing proteins clustered together with the known PsojNIPs in the phylogenetic tree, and were selected for further alignment (Fig. 7). The highest identities were 85% (P. sojae 18) and 88% (P. ramorum 38) between the PmegNep6 and the putative NLPs from P. sojae or P. ramorum, respectively.

**DISCUSSION**

*NEP1* orthologs constitute a multigene family  

Analysis of the *Phytophthora megakarya* genome using Southern blot and PCR analyses showed that there are multiple *NEP1* orthologs. PmegNEP2 probe was cross-hybridized with multiple DNA fragments under medium stringency condition on Southern blots. The resulting DNA hybridization patterns showed that polymorphisms are present in five species of *Phytophthora*. That is, the copy number and gross structural features of the *NEP1* multigene family are not conserved among *Phytophthora* species. In previous studies, one major band was detected in Southern blot analyses in *P. parasitica* and *P. sojae* under high stringency condition (Fellbrich et al. 2002, Qutob et al. 2002). However, multiple weakly hybridizing bands were detected when genomic DNA was probed under medium stringency (data not shown, Qutob et al. 2002). Fellbrich et al. (2002) were able to detect up to four bands in five *Phytophthora* species (*P. parasitica, P. infestans, P. palmivora, P. cactorum,* and *P. nicotianae*) as well as *Pythium vexans* under medium stringency condition. In this study, a complex gene family of *NEP1* orthologs was detected in *P. megakarya* and *P. sojae* as well as three other species studied using Southern blot analysis. In addition to Southern blot analysis, genome database searches of the *P. sojae* and the *P. ramorum* also confirmed the existence of a *NEP1* multigene family. While the identities between PmegNep6 and the three PsojNIPs were around 70%, two of the putative PsojNIPs (P. sojae 14 and 15 in Fig. 7) showed identities over 80% with PmegNep6. In addition, there are three putative NLPs that have identities over 80% with PmegNep6 in the genome *P. ramorum*. These results strongly suggest that there are additional PsojNIPs and that there are multiple copies of NLPs in the genome of other *Phytophthora* species.

**Multiple PmegNEPs expressed during the infection of cacao leaves and mycelia growth**

Nine PmegNEP orthologs were verified in the genome of *Phytophthora megakarya* when mycelial DNA was used as a PCR template. The accumulation of PmegNEP transcripts during the infection of cacao leaf disks and during mycelial growth was investigated by Northern blot analysis, however, no transcript was detected (data not shown). The greater sensitivity of RT-PCR analysis allowed for the detection of PmegNEP transcripts both in inoculated cacao leaf tissue and in mycelia. Expression of only six of nine orthologs were verified in the genome *P. megakarya* product of PmegNEP6 and the three PsojNIPs were around 70%, two of the putative PsojNIPs (P. sojae 14 and 15 in Fig. 7) showed identities over 80% with PmegNep6. In addition, there are three putative NLPs that have identities over 80% with PmegNep6 in the genome *P. ramorum*. These results strongly suggest that there are additional PsojNIPs and that there are multiple copies of NLPs in the genome of other *Phytophthora* species.
active. The same possibility can also apply to the other PCR products in Fig. 3. That is, even the expected sizes of the PCR products were detected, the products could be new orthologs that have homologous sequences to the primers used. Qutob et al. (2002) detected the PsojNIP transcripts during the transition from biotrophy to necrotrophy, that is, the transcript was not detected during the early stage of infection. These results suggest differential induction of NEP1 orthologs in Phytophthora species.
**PmegNEP orthologs show characteristics of the genome structure of Phytophthora spp.**

Some characteristics of the *Phytophthora* genome organization were revealed under further analysis of the *P. megakarya* genome for *PmegNEP* orthologs. Two cassettes were found that contained at least three *Phytophthora* genome is a short distance between ORFs (Kamoun 2003). For example, the distance between *PmegNEP4* and *PmegNEP6* (also between *PmegNEP4-2* and *PmegNEP6-2*) was 532 bp, and the distance between *PmegNEP1* and *PmegNEP4-2* was 268 bp. The distance between the two stop codons of *PmegNEP6* and *PmegNEP5* was only 56 bp. Similar observations were made in *P. sojae*. In *P. sojae* the average distance between ORFs was 497 bp (range 102 to 714 bp) in the region of seven genes including three genes encoding PssoNIP and PssoNIP-like proteins clustered in a 10835-bp BamH1 fragment (Qutob et al. 2002). Analysis of three promoter regions showed that there were no motifs corresponding to CAAT and TATA boxes in *P. megakarya*, which is a typical phenomenon in the *Phytophthora* genome (Judelson & Michelmore 1991, Judelson, Tyler & Michelmore 1992, Kamoun 2003). The clustering of genes and the short distance between ORFs also matches the characteristics of the highly compact transcripts with short untranslated regions found in *Phytophthora* spp. (Pieterse et al. 1994, Pesole et al. 2000, Kamoun 2003). Introns were not found in any of the *PmegNEP* orthologs, which is also a common characteristic of the genome of other *Phytophthora* spp. (Kamoun 2003).

The existence of genes in small clusters has been reported previously in other *Phytophthora* species. For examples, two clusters of *in planta-induced* (ipi) genes were detected in *P. infestans* (Pieterse et al. 1994). Three similar *ipiB* genes were arranged start-to-stop codons with short intergenic sequences (820 bp) and two *ipiO* genes were arranged start-to-start codons with a 2200-bp intergenic sequence. In *P. cinamomi*, there are at least six polygalacturonase genes in several clusters; in addition, at least three pyruvate phosphate dikinase genes were arranged in a small cluster (Marshall et al. 2001).

**PmegNep clones share sequence characteristics with other Nep1-like proteins (NLPs)**

The structural features of the *PmegNep* clones include a hydrophobic signal peptide at the N-terminus. The signal peptides for *PmegNeps* had 17–19 amino acid sequences, while 19 amino acid signal peptides were predicted for other *Phytophthora* NLPs. A conserved amino acid sequence could not be detected for the signal peptides. Unlike typical elicitors, *PmegNep* proteins are predicted to have a glycosylated region, suggesting localization to the cell wall or membranes. According to the result obtained from transmission electron microscopy, Nep1 treatment induces rapid breakdown of thylakoid membrane structures. This result might indicate that the Nep1 protein is localized to membrane structures. However, no glycosylated residue was found in the active PssoNIP of *P. sojae* (AAK01636, Qutob et al. 2002). This result indicates that glycosylation by *Phytophthora* spp. is not important in the ability of NLPs to induce necrosis.

Two full length *PmegNEP* orthologs were cloned: *PmegNEP4-2* and *PmegNEP6* predicted to yield proteins with 235 and 233 amino acids, respectively. Cysteine residues have important roles in folding, structure, and function of proteins (Boissy et al. 1999, Qutob et al. 2002). When cysteine residues are exposed to air, the sulphhydril (SH) groups oxidize and are joined by a disulphide bond. Deletion mutation analysis in *P. parasitica* NPP1 indicated the elictor activity was dependant upon the tertiary structural characteristics of the protein and the pattern of spacing of the conserved cysteine residue (Fellbrich et al. 2002). The importance of cysteine residues has also been reported for other secreted elicitors (Templeton, Rikkerink & Beever 1994). In this study, it was determined that the second cysteine residue (C82 and C80 in *PmegNep4-2* and *PmegNep6*, respectively) was conserved in the predicted amino acid sequences of all nine *PmegNEP* orthologs. Only partial sequences (corresponding to 11F and 16R) were used for alignment due to the lack of sequence information regarding the sequences in some of *PmegNEP* orthologs. Additional information about the N-terminal region for *PmegNep2* and the C-terminal region for *PmegNep1* was gained through 5’ and 3’ RACE. The first cysteine residue (C55, 56, 54, and 54 in *PmegNep2*, *PmegNep4-2*, *PmegNep6* and *PmegNep6-2*, respectively) was also conserved in all the predicted amino acid sequences of *PmegNEP* orthologs when sequence data was available (Fig. 7). Most notably, there are 25 amino acids between the first and second cysteines in most NLPs except in *Neurospora crassa* (26 amino acids, EAA30528) and *Magnaporthe grisea* (63 amino acids, EAA54347). This indicates that spacing between the two cysteines may be important for activity. The importance of the spacing between cysteines was also observed in the deletion mutation of *P. parasitica* NPP1 (Fellbrich et al. 2002) and in other secreted elicitors (Templeton et al. 1994). Interestingly, additional cysteine residues were predicted in some of the *PmegNEP* orthologs, for example, the third or/and fourth cysteine residues in *PmegNep4*, C98 and C128 in *PmegNep4-2*, C107 and C126 in *PmegNep6*, and C126 in *PmegNep6-2*. Further analysis is required to examine the importance of these additional cysteine residues.

**Genetic diversity and genome stability**

Of the multiple copies of *PmegNEP*, nine clones that were analyzed in this study had unique ORFs, encoding...
different forms of PmegNep. Comparisons of the two cassettes (Fig. 1) that contain PmegNEP orthologs suggest that duplication and mutation has occurred in the P. megakarya genome. Based on comparisons within the cassettes (Fig. 1), polymorphisms in the PmegNEP ORFs indicate that the amplification of some PmegNEP orthologs might be a remote event so that there has been sufficient time for genetic variation to accumulate. Alternatively, there is mechanism for rapid genetic variation in Phytophthora species, such as transposable elements, gene conversion, mitotic recombination, and/or dispensable chromosomes (Martin 1995, Goodwin 1997, Chamnanpunt et al. 2001, Judelson 2002). The overall identities among PmegNEP orthologs ranged from 42–83%. In contrast, the second cassette (Fig. 1B) contained three putative PmegNEP orthologs and two out of three were only slightly different from previously found PmegNEP orthologs. That is, while PmegNEP1 was identical with a previously detected sequence, PmegNEP4-2 and PmegNEP6-2, although different, were closely related (96% identity) to the previously detected clones PmegNEP4 and PmegNEP6, respectively. In addition, while the length of the intergenic sequences between PmegNEP4 (PmegNEP4-2) and PmegNEP6 (PmegNEP6-2) were the same (532 bp) they had slightly different sequences (14 bp out of 532 bp were different). Based on the sequence comparison between the two cassettes, the duplication of PmegNEP from one cassette to the other cassette is a relatively recent event, resulting in the generation of PmegNEP orthologs with high homology. In addition, the existence of the DNA ortholog PmegNEP7 that contains two genes fused together suggests instability in the genome of P. megakarya.

The survival of an organism depends on the ability to adapt to various environments and new or modified gene functions are a means of generating such adaptability. Duplication and divergence is the classical model of the generation of these novel genes (Hooper & Berg 2003). The dosage effect might be advantageous primarily for weakly expressed genes (Hooper & Berg 2003). Rapid generation of genetic variation has been reported in Phytophthora species grown in culture (Fry 1982) and under field conditions (Chamnanpunt, Shan & Tyler 2001, Judelson 2002). The mechanisms of generating such variation are not clear, but possible sources may be transposable elements, gene conversion, mitotic recombination, and/or dispensable chromosomes (Martin 1995, Goodwin 1997, Chamnanpunt et al. 2001, Judelson 2002). Gao et al. (2002) postulated that the copy number diversity of the Telomere-Linked Helicase (TLH) gene family in Magnaporthe grisea isolates could be explained by unequal crossing over between repetitive elements associated with the TLH gene family. Repetitive sequences are abundant in the genomes of Phytophthora species. For example, five families of tandemly repeated sequences were found in P. sojae (Mao & Tyler 1996) and, 33 distinct families of repeated DNA have been detected in P. infestans (Judelson & Randall 1998). The repeated sequences cover at least 50% of the genome and copy numbers ranged from 70–8400 per haploid genome. A role for transposable elements in gene amplification has been reported in Pyrenophora tritici-repentis, the causal agent of tan spot in wheat, which produces a host-selective toxin, ToxB (Gao et al. 2002, Martinez, Oesch & Ciuffetti 2004). It was found that ToxB loci were associated with truncated retrotransposon-like sequences. The association of a retrotransposon-like sequence with ToxB provided for the opportunity of unequal crossing over with the same or similar sequences. Serendipitously, we amplified a copia-like retrotransposable element from P. megakarya DNA using the specific primer set of PmegNEP6 Forward/PmegNEP3 Reverse primers. The DNA PCR fragment was 1283-bp long and had a moderate homology with the copia-like retrotransposable element of Arabidopsis thaliana (41% identity with E-value 3e-46, BAB01972). However, no trace of PmegNEP sequence could be found in the PCR fragment.

Based on two characteristics, Qutob et al. (2002) suggested that PsojNIP is a virulence factor rather than an avirulence factor: (1) PsojNIP was active in a host cultivar that was susceptible to most P. sojae isolates; and (2) PsojNIP-like proteins are found in taxonomically different organisms with different hosts. A similar mechanism to PsojNIP has been proposed for Nep1 from Fusarium oxysporum (Jennings et al. 2001), and for various non-proteinaceous toxins from Alternaria alternata f. sp. lycopersici (Wang, Bostock & Gilchrist 1996), Cochliobolus victoriae (Navarre & Wolpert 1999), and Fusarium moniliforme (Stone et al. 2000). In addition, the existence of a complex multigene family encoding NLPs may also support the hypothesis that NLPs are virulence factors.

The data presented indicates that the NLP gene family structure and its regulation are much more complex than previously thought. At least nine copies of PmegNEP orthologs were detected in P. megakarya. Multiple copies of PmegNEP were expressed during the infection of cacao leaves and mycelial growth, and were regulated differently. The P. megakarya genome is unstable, resulting in the generation of duplications, inversions, and fused genes. The multiple and ubiquitous existence of the NLP gene family in the Phytophthora species, as well as many other organisms suggests a significant role of the gene in pathogenicity of microbes. The existence of a gene family containing NLPs may be advantageous for a wide range of plant pathogens. Certainly, efforts aimed at understanding the function(s) of NPLs in microbial biology and plant microbe interactions, whether as avirulence or virulence factors, will be impacted by the complex nature of the NPL gene family and the potential for generation of genes with new functions due to the instability of the Phytophthora genome.
REFERENCES


