



## Putative pathogenesis-related genes within *Solanum nigrum* L. var. *americanum* genome: isolation of two genes coding for PR5-like proteins, phylogenetic and sequence analysis

M. A. CAMPOS\*, S. G. RIBEIRO, D. J. RIGDEN, D. C. MONTE and M. F. GROSSI DE SA

<sup>1</sup>EMBRAPA Recursos Genéticos e Biotecnologia, P.O. Box 02372, 70770-900, Brasília, DF, Brazil and <sup>2</sup>Universidade de Brasília, Departamento de Biologia Celular, 70910-900, Brasília, DF, Brazil

(Accepted for publication 22 July 2002)

Pathogenesis-related (PR) proteins are involved in plant responses to biotic and abiotic stress and are grouped into 14 families (PR-1–PR-14). Two PR5-like genes were isolated from black nightshade (*Solanum nigrum* L. var. *americanum*) genome, a solanaceous weed. A complete open reading frame consisting of 744 bp, coding for a precursor of a neutral PR5-like protein, and a DNA fragment of 621 bp long, coding for another basic PR5-like protein, were obtained by using PCR amplification. Based on sequence comparisons with tobacco osmotin and osmotin-like proteins (OLPs), members of the antifungal PR-5 family, these proteins were named SnOLP (neutral) and SnOSML (basic). The two predicted mature proteins are 207 amino acids in length and contain the 16 cysteine residues involved in the eight disulfide bonds common to most PR-5 proteins. Southern blot analysis revealed that osmotin-like proteins are encoded by at least eight members of a multigene family in *S. nigrum*. This is the first report of the isolation of PR5-like genes from the *S. nigrum* genome. Computer modeling of the new sequences produces structures that, according to current hypotheses, are indicative of antifungal activity. Phylogenetic analysis of solanaceous PR-5 proteins was also carried out revealing three major groups with different characteristics.

© 2002 Elsevier Science Ltd. All rights reserved.

**Keywords:** pathogenesis-related proteins; osmotin-like proteins; PR-5 proteins; black nightshade; *Solanum nigrum* L. var. *americanum*; Solanaceae; PCR; phylogeny; modelling.

### INTRODUCTION

Pathogenesis-related proteins (PR proteins) comprise a large group of diverse plant defence proteins which are induced upon microbial attack through many signaling pathway elements, including different receptor components or chemical elicitors such as salicylic acid, jasmonic acid, systemin, and ethylene [17, 36, 59]. Together with other antimicrobial polypeptides and defence-related proteins, PR proteins are end products of the signal transduction cascades that lead to complex plant defence responses, known as the hypersensitive response (HR), occurring at the infection site of an invading pathogen [10, 35]. However, the set of PR proteins not only accumulate locally in the infected tissue, but are also induced systemically, associated with the development of systemic acquired resistance (SAR)

against further infection by fungi, bacteria and viruses [36, 41, 57]. The apoplastic and vacuolar localization of many of these proteins suggests that they may be positioned to come into contact with the pathogen during the infection process, although, PR proteins appear to be only part of the induced resistance phenotype [52, 53]. Originally, these proteins were classified into five main groups or families, PR-1–PR-5 [51]. This list has now been extended to contain 14 families, including fungal cell wall-degrading enzymes, such as  $\beta$ -glucanases (PR-2), and chitinases (PR-3, PR-4, PR-8, PR-11), antimicrobial polypeptides and components of signal transduction cascades [10, 53].

The family of PR-5 proteins comprises proteins with diverse functions, as well as roles in development [30, 43], protection against osmotic stress [20, 46, 63] and freezing tolerance [32, 62] in plants, PR-5 proteins also permeabilize fungal and oomycete plasma membranes [1, 2, 37]. Due to the sequence similarity between PR-5 proteins and thaumatin, a sweet-tasting protein from fruits of West African shrub *Thaumatococcus daniellii* [6, 21, 50], members of this family were first referred to as thaumatin-like

\* To whom all correspondence should be addressed. E-mail: camposma@cenargen.embrapa.br

Abbreviations used in text: HR, hypersensitive response; OLP, osmotin-like protein; PCR, polymerase chain reactions; PR, pathogenesis-related; SAR, systemic acquired resistance; TLP, thaumatin-like protein.

proteins (TLPs). Greater sequence similarity with a PR-5 family member, tobacco osmotin, a protein induced by osmotic stress [46], means that they are now more commonly referred to as osmotin-like proteins (OLPs). Osmotin and several others PR-5 proteins have antifungal activity, whereas the thaumatin protein does not [26]. The *in vitro* antifungal activity of PR-5 proteins is through inhibition of hyphal growth and spore germination, spore lysis and a reduction in viability of germinated spores. This activity has been observed against several pathogenic and non-pathogenic fungi [1, 54, 57, 65]. However, the antimicrobial molecular mechanisms of PR-5 proteins are still not completely understood. A fungal/oomycete membrane-permeabilization mechanism was proposed for several PR-5 proteins, including zeamatin of maize [37], osmotin of tobacco [1] and linusitin of flax [2]. All these proteins cause rapid leakage of cytoplasmic contents from fungal cells. The existence of a target specificity between PR-5 proteins and the target microorganisms was suggested, e.g. for *Saccharomyces cerevisiae* strains and tobacco osmotin interactions, where cell wall carbohydrates and/or a probable membrane receptor of the fungi might be involved [13, 61]. Nevertheless, the mechanism of membrane permeabilization or interaction with fungal receptor is still not clear. Recently, it was demonstrated that osmotin induces apoptosis in *S. cerevisiae* in a manner that was correlated with intracellular accumulation of reactive oxygen species and mediated by RAS2/cAMP signaling pathway, which is involved in cell division and growth in yeast [29].

Apparently, osmotin and OLPs from several plants have antimicrobial activity effectively directed towards the oomycete *Phytophthora infestans*, the causal agent of the destructive late blight disease of potato and tomato plants [58]. The *in vitro* lysis of sporangia and hyphal growth inhibition of *P. infestans* caused by tobacco and tomato osmotin have been demonstrated [1, 26, 58]. Moreover, the overexpression of tobacco osmotin and an OLP (OSML13) of potato (*Solanum commersonii*) conferred enhanced resistance to *P. infestans* in transgenic potato plants [22, 24, 25]. In addition to this evidence, results of actin-binding studies suggested that an OLP and a basic chitinase might be involved in cytoplasmic aggregation, an important event in potato cellular defence against *P. infestans* [48].

It has also been demonstrated that the constitutive expression of PR genes, including PR-5 group genes, may contribute to non-specific resistance to *P. infestans* in *Solanum* [55]. Osmotin and OLPs are encoded by a multi-gene family that is highly conserved in the Solanaceae [64]. The genus *Solanum* constitutes an extensive gene pool, in which a broad spectrum of pathogen resistance has accumulated throughout evolution [40, 56]. Black nightshade (*Solanum nigrum* L.), a solanaceous weed, is a remarkable example of a non-host plant possessing resistance to *P. infestans*, which survives in *P. infestans*-infected potato

fields. It was reported that even though *S. nigrum* has been continuously exposed to *P. infestans* since the introduction of this oomycete to Europe 150 years ago, it has remained resistant [4]. When *S. nigrum* was inoculated with *P. infestans* in laboratory conditions, it was observed that the penetration of the leaf epidermis by the oomycete was accompanied by rapid cell death (HR) of the penetrated plant cell [14, 56]. Although the PR-5 genes from *S. nigrum* have a potential for control of fungi or oomycetes, they have not yet been studied. In this report, we describe the first cloning and characterization of genes coding for neutral and basic PR5-like proteins from *S. nigrum* L. var. *americanum*. Their high degree of sequence identity with other PR-5 proteins with known antifungal activity suggests that the *S. nigrum* PR5-like genes encode for proteins with similar functions. Computer modeling of the novel sequences is in accord with current hypotheses regarding function. Phylogenetic analysis of solanaceous PR-5 proteins was also carried out; revealing three major groups with different characteristics.

## MATERIALS AND METHODS

### *Plant material*

Black nightshade (*S. nigrum* L. var. *americanum*) plants were grown in greenhouse conditions from seeds collected in open fields in the south of Minas Gerais, Brazil.

### *DNA isolation and southern hybridization analysis*

The genomic DNA of black nightshade was isolated from the younger leaves of 5-week old plants by the CTAB method [38]. Southern hybridization analysis [47] was performed following the protocol described by Romano [39]. Briefly, 20 µg of total DNA was digested with *Eco*R I and *Xba* I and fractionated on a 0.8% (w/v) agarose gel by electrophoresis. DNA was transferred onto nylon Hybond-N (Amersham) membrane by capillary transfer and blots were hybridized using *SnOLP* as probe (isolated by PCR amplification in this work). The DNA probe was prepared by using a random primed-DNA-labelling kit (Pharmacia) and labelled with <sup>32</sup>P. After hybridization at 65°C in hybridization solution containing 5× SSC (Sodium Chloride/Sodium Citrate) buffer, 5× Denhardt's solution, 0.5% SDS and 20 µg ml<sup>-1</sup> salmon sperm DNA, the membrane was washed three times with solutions containing 2×, 1× and 0.1× SSC plus 0.1% SDS at 65°C for 15 min each and subjected to autoradiography.

### *Primer design, PCR amplification and DNA sequencing*

The primers used in this study are presented in Table 1 and Fig. 1. Oligonucleotides were designed based on an

TABLE 1. Sequence of primers designed to amplify PR5-like sequences from *Solanum*. The primers PPS1, PPM1, NP1 and P5 were designed based on multiple alignment of sequences of PR-5 protein genes and homologous from the GenBank database, while PPM2 and PCTP were based on the sequence of a *S. gilo* gene (Campos et al., unpublished work). Primers NP1 and P5 were designed from highly conserved regions within the coding sequence of PR-5 genes. Primers PPS1 and PCTP were synthesized to amplify a putative ORF coding region for PR5-like genes. Primers PPM1 and PPM2 were constructed to amplify fragments of PR-5 genes coding for only putative mature proteins. PPS1 and PPM1 containing a suitable Bam HI cloning site whereas PCTP and PPM2 containing a Hind III site. Numbering under primer sequences correspond to the nucleotide position on SnOLP gene sequence (in this paper)

Primer	Nucleotide Sequence (5'–3') and location on <i>SnOLP</i> gene	Based on region
PPS1	CGCGGATCCATGGGCTACTTGAGATCT Bam HI 1 18	Non-conserved
PPM1	CGC GGATCCG CTGCGACTATCGAGGTACGC Bam HI 61 81	Semi-conserved
NP1	ACA ACTGTCCGTACACCG 83 100	Conserved
P5	GGGTTGTTACATCCTCCG 626 643	Conserved
PPM2	CCC AAGCTTA CCCTTAGGACAAAAGACAACCC Hind III 662 684	<i>SgOLP</i> gene
PCTP	CCC AAGCTT T TACTTGGCCACTTCATC Hind III 727 744	<i>SgOLP</i> gene

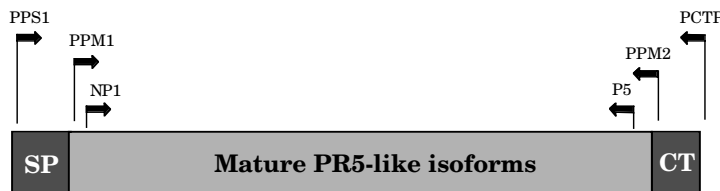


FIG. 1. Schematic representation of PR-5 proteins and annealing position of the primers described in Table 1 on coding sequences for *Solanum* PR-5. The orientation of the primers is indicated by arrows and the putative regions of PR-5 precursors are shown. SP, Signal Peptide; CT, Carboxy-Terminal propeptide.

alignment of sequences into PR-5 group genes and homologues from the GenBank database. For the five primer combinations (PPS1/PCTP, PPM1/PPM2, PPM1/PCTP, PPS1/PPM2 and NP1/P5), PCR were performed using 0.2  $\mu$ M of each primer, 0.06  $\mu$ g of template DNA from *S. nigrum* leaf, 200  $\mu$ M of each dNTP, and 1.0 U of Platinum High-Fidelity Taq Polymerase (Gibco, Taq Polymerase buffer and MgSO<sub>4</sub> concentrations according to manufacturer's procedures), in a final volume of 50  $\mu$ l. The conditions used were 94°C for 3 min, followed by 25 cycles of 1 min at 94°C, 1 min at 50°C, 2 min at 68°C, with a final extension of 5 min at 68°C.

PCR-amplified products were analysed by 0.8% agarose gel electrophoresis and their specificity was tested by blot hybridization using a partial PR-5 homologous gene isolated from *Solanum gilo* (Campos et al., unpublished work) as probe. The PPS1/PCTP and PPM1/PPM2 fragments were eluted from agarose gel electrophoresis and cloned into pGEM-T Easy vector (Promega). Recombinants were selected using the X-Gal/IPTG system. Three clones harbouring the expected size inserts

analysed by both digestions with *EcoR* I and PCR were selected for sequencing. Cloned DNA was sequenced entirely on both strands using an automated ABI sequencer with BigDye terminator cycle sequencing kit (Perkin-Elmer). Sequences were analysed using the UWGCG software Package (Version 9.1 Genetics Computer Group, Wisconsin, Madison, WI, U.S.A.), and DNA MAN (Version 4.0, Lynnon Biosoft).

#### Phylogenetic analyses

A set of PR-5 homologue protein sequences from solanaceous plants were selected from the results of a search of the nr database (<http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=Protein>) using Hidden Markov Model methods [15]. Sequence characteristics (signal peptide, pI for mature region and C-terminus propeptide) were evaluated for each protein using the resources of the protein analysis tools site ([www.expasy.ch](http://www.expasy.ch)). Multiple alignments were performed with the mature peptide sequences using PileUp program from UWGCG. Phylogenetic analysis of those alignments was done using either

the Protpars program or the combination of the Protdist and Neighbor programs of Phylip package [7]. Bootstrap analysis employing 100 replicates was carried out.

*Structural modelling*

Computer modelling of structures for SnOLP and SnOSML proteins was carried out using MODELLER [42] based on known three-dimensional structure of the related PR-5d protein, a neutral PR-5 protein from *Nicotiana sylvestris* [Protein Data Bank identifier 1AUN, 18].

*Sequence accession number*

Gene sequences corresponding to SnOLP and SnOSML proteins are deposited in GenBank as accession numbers AF450276 and AF450277, respectively.

**RESULTS**

*Cloning and molecular characterization of PR5-like sequences from Solanum nigrum var. americanum*

A complete open reading frame (ORF) of a gene coding for a neutral PR5-like protein and a partial sequence of a gene for a basic PR5-like protein were obtained from *Solanum nigrum* L. var. *americanum* DNA. All primer pairs, designed (Table 1, Fig. 1) to amplify DNA of sequences coding for PR-5 genes, generated single fragments by PCR with the expected sizes from the genome of *S. nigrum* L. var. *americanum*. The Southern blot presented in Fig. 2, shows that no non-specific PCR products were obtained, since comparison with the original gel of PCR products (data not shown) contains the same number of bands in identical positions. Restriction enzyme mapping and sequence analysis of three independent clones harbouring

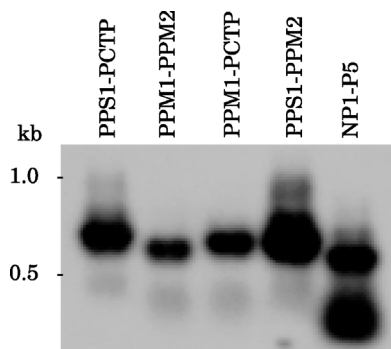


FIG. 2. Southern hybridization of *S. nigrum* PCR products amplified with the primers indicated in Table 1, probed with a homologous gene from *S. gilo* (Campos et al., unpublished work). In the last lane, the lower spot might result from hybridized primers.

the fragments produced by the two PPS1/PCTP and PPM1/PPM2 primer pairs revealed the presence of two different insert sequence groups, that are highly similar between themselves and with genes coding for PR-5 proteins of plants belonging to the Solanaceae family. Within the sequenced clones, fragments amplified with the same pair of primers showed no sequence variation.

The use of the synthetic oligonucleotides PPS1 and PCTP resulted in amplification of the complete ORF, whose nucleic acid sequence is 744 bp in length. The complete nucleotide sequence of this ORF of a putative PR5-like gene, denominated *SnOLP*, is shown in Fig. 3, along with the deduced primary structure of the protein. The coding sequence of the *SnOLP* gene lacked intron sequences and encoded a precursor protein of 247 amino acids residues in length. The first 21 amino acids of the predicted protein have typical features of a signal peptide [33], including length, a hydrophobic core, and a potential cleavage site situated between the amino acids 21 and 22 (TYA-TS), which is involved in the transport

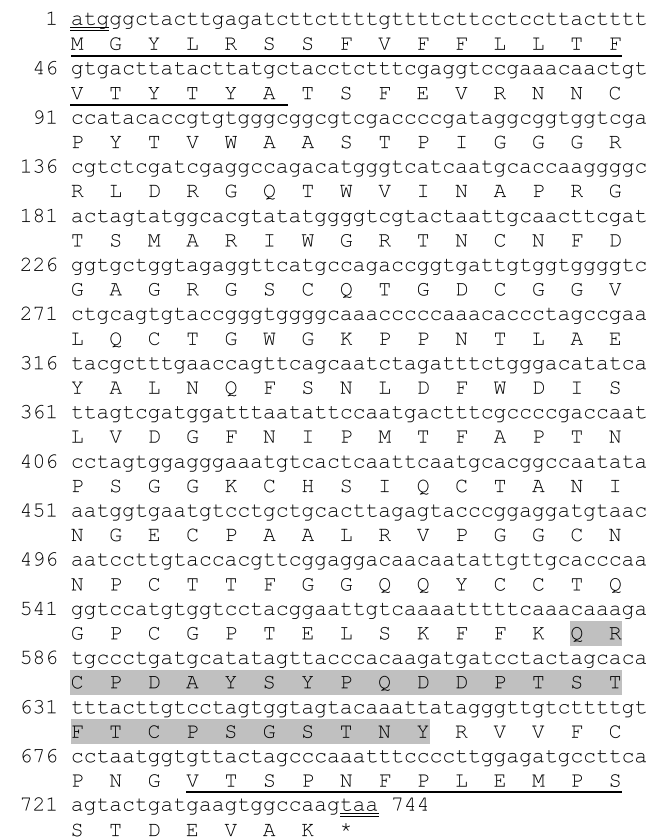


FIG. 3. Nucleotide and deduced amino acid sequences of the *SnOLP* gene. Putative translation initiation and termination codons are double-underlined. Putative amino-terminal signal and carboxy-terminal sequences are underscored. The potential PEST rich sequence is shown gray-shaded. The GenBank accession number is AF450276.

across the membrane of the endoplasmic reticulum. A carboxy-terminal propeptide of 19 amino acids is also present, which may be removed in a second processing step and may be involved in vacuolar targeting. These data suggest that the SnOLP polypeptide is synthesized as a preproprotein that may be processed into a mature slightly neutral isoform of 207 amino acids with a calculated molecular weight (MW) of 22, 274 Da and a theoretical pI of 7.28.

Conversely, the sequence analysis of the three clones from PCR products obtained using the PPM1/PPM2 primers revealed a DNA fragment of 621 bp long that seems to be part of a different gene coding for a second PR5-like protein in *S. nigrum*, *SnOSML* (Fig. 4). The predicted amino acid sequence data indicate the presence of a complete mature polypeptide with 207 amino acids in length, whose calculated MW is 22 207 Da and which has a theoretical pI of 8.89. These data suggest that the polypeptide, denominated SnOSML, coded by this partial gene sequence may be a basic mature form PR5-like protein, different from SnOLP but sharing 90 % identity with it (Fig. 5). Computer analysis suggest that a potential PEST (proline-glutamate-serine-threonine) rich sequence, as already reported for other PR-5 proteins, is present only in SnOLP, whereas the potential N-linked glycosylation site (Asn-X-Ser/Thr) observed in the neutral OLP of the *N. sylvestris* is present in neither protein [45]. Sixteen cysteine residues, which participate

in eight disulfide bonds common to most PR-5 proteins, are conserved in SnOLP and SnOSML (Fig. 5).

*Homology of PR5-like sequences from S. nigrum among those from the Solanaceae family*

A BLAST search for proteins with amino acid sequences similar to deduced SnOLP and SnOSML proteins indicated significant homology (45–98 % identity of amino acids) with around 100 proteins sequences available in the database (data not shown). Similar results were obtained when nucleotide sequences of the PR-5 genes were compared. Among the amino acid sequences compared, the highest degrees of identity for SnOLP and SnOSML were found with PR-5 homologues from other solanaceous plants. SnOLP shared 94, 93 and 92 % identity with precursors of the P23 and TPM-1 proteins of *Lycopersicon esculentum* and OSML13 of *S. commersonii*, respectively, whereas, SnOSML shared 98, 97 and 96 % of identity with NP24 of *L. esculentum*, OSML81 of *S. commersonii* and a PR-5 protein of *Capsicum annum*, respectively. The same identity percentages (79, 61 and 60 %) were found between both amino acid sequences of SnOLP and SnOSML with a neutral PR-5 protein of *Nicotiana tabacum*, zeamatin and MAI (alpha-amylase/trypsin inhibitor) of *Zea mays*, respectively.

Multiple alignment of the mature deduced amino acid sequences of SnOLP and SnOSML with seven PR-5 proteins from solanaceous plants revealed a high identity of amino acid sequences and several conserved regions (Fig. 6). The mature proteins were mostly about 207 amino acids long, contained the 16 conserved cysteine residues forming disulfide bridges and the amino acid residues that share the acidic cleft (E84, D97 and D102) in tobacco PR-5d protein were also present in all sequences. The amino and carboxy-terminus were the most variable regions among the compared sequences (not shown).

*Number of PR5-like genes in the Solanum nigrum L. var. americanum genome*

In order to investigate the presence of PR5-like sequences in *S. nigrum*, a Southern hybridization analysis was done using genomic DNA digested by a restriction enzyme that potentially do not cleave within the coding sequence of PR-5 (*EcoR* I). This revealed that at least eight large discrete restriction fragments hybridize to <sup>32</sup>P labelled *SnOLP* ORF fragment (Fig. 7). These results suggest that the *SnOLP* and *SnOSML* genes belong to a multigene family. Variation in the signal intensity among of the hybridizing bands suggested sequence diversity or presence of multiple copies (per restriction enzyme fragment).

```

1  gctgcgactatcgaggtagccaacaactgtccctacaccggtgtgg
   A A T I E V R N N C P Y T V W
46  gcggcgctcgaccccgatagcggtgctcgagctctcaatcgaggc
   A A S T P I G G G R R L N R G
91  caaacgtgggtcatcaatgcgccgaggggactaagatggctcgt
   Q T W V I N A P R G T K M A R
136 atatggggtcgtactggttgaacttcaatgctgcaggtagaggt
   I W G R T G C N F N A A G R G
181 tcatgtcagactggtgactgtggtggagcttgcggtgtacaggg
   S C Q T G D C G G V L R C T G
226 tgggggcaaaccaccaacacgcttggccgaatatgccttgaccag
   W G K P P N T L A E Y A L D Q
271 ttcggcaacctagatttctgggatatatcattagtcgatggattc
   F G N L D F W D I S L V D G F
316 aatattccgatgactttcgccccacaaacctagtggtgggaaa
   N I P M T F A P T K P S G G K
361 tgccatgcaattcattgcacggccaatataatggtgaatgtcct
   C H A I H C T A N I N G E C P
406 cgcgcccttaaggtagctggaggatgtaacaaccttgtaccacg
   R A L K V P G G C N N P C T T
451 ttcggaggacaagaatattgttgcacccaggggtccatgtcct
   F G G Q Q Y C C T Q G P C G P
496 acggagttgtcgaaattttcaagaaaagatgtcctgatgcgtat
   T E L S K F F K K R C P D A Y
541 agctaccacaagaatgatcctactagcacatttactgtcctgga
   S Y P Q D D P T S T F T C P G
586 ggtagtacaaactataggggttgtcctttgtcctaaggggt 624
   G S T N Y R V V F C P K G

```

FIG. 4. Nucleotide and deduced amino acid partial sequences of the *SnOSML* gene. PPM1 and PPM2 primers are underlined. The GenBank accession number is AF450277.

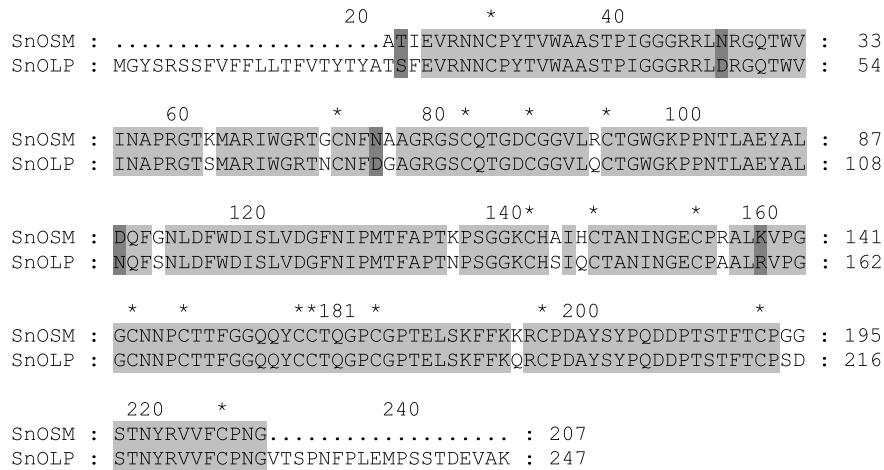


FIG. 5. Comparison of the deduced amino acid sequence of SnOLP with SnOSML. Identical and similar amino acid residues are shaded in light or dark gray, respectively. Asterisks indicate the 16 cysteine residues. The SnOLP sequence contains signal peptide and C-terminal cleaved regions not present in SnOSML.

#### Phylogenetic analyses of PR5-like genes within the Solanaceae plant family

For phylogenetic analysis a data set containing 18 solanaceous PR-5 mature protein sequences found by database searches, supplemented with the new *S. nigrum* deduced amino acid sequences, was used. The branching orders of neighbour-joining and maximum parsimony-based trees were identical. Bootstrap analysis (100 replicates) of the data showed that the tree branches are well supported. The neighbour-joining tree (Fig. 8) shows that the various PR-5 sequences are grouped into three major clusters. Four *N. tabacum* (NtPR-R1 major, NtPR-R2 minor, NtTLPSE39b and NtTLP) comprise the first cluster while three from different organisms-ScOSML15, LePR-5 and NsPR5d (from *S. commersonii*, *L. esculentum* and *N. sylvestris*, respectively) make up the second cluster. The remaining 13 sequences belonged to a third cluster, in which were placed the two sequences from *S. nigrum* (SnOLP and SnOSML), three osmotin derived from *N. tabacum* (1–3 OSM), four sequences derived from *L. esculentum* (LeNP24, LeSaltInd, Le23 and LeTMP-1), SdOLP from *Solanum dulcamara*, CaTLP from *C. annuum* and PHOSM from *Petunia x hybrida*. The phylogenetic tree derived from nucleotide sequences was similar to the one derived from the amino acid sequences (data not shown).

#### Computer-based analysis of the three-dimensional structure of the protein

The 84 or 80 % sequence identity, respectively, between SnOLP and SnOSML and the template used to construct their models ensures that the target-template alignment is correct and that the resulting models are therefore reliable. As shown in Fig. 9, both SnOLP and SnOSML

contain the two principal domains seen for PR-5d protein from *N. sylvestris* [18].

## DISCUSSION

Here, we described the first report of the isolation of PR5-like genes from the *S. nigrum* L. var. *americanum* genome. By using PCR and specific primers, we isolated two distinct DNA sequences containing putative coding regions for two new PR-5 genes from *S. nigrum*. In the first case, a complete ORF coding for a precursor of a PR5-like protein was obtained including an amino-terminal signal peptide and a carboxy-terminal propeptide. The calculated pI value of 7.28 for the mature polypeptide indicated a neutral isoform, which was named SnOLP. Additionally, a partial sequence for a gene coding another PR5-like protein was obtained, this time a basic isoform (calculated pI value of 8.89) named SnOSML protein. The presence of various isoforms with diverging pI values has been reported for several PR proteins, including the PR-5 protein group where three classes have been found in some plants genome; acid, basic and neutral isoforms [17]. According to the available evidence PR proteins are synthesized as precursors with an N-terminal signal peptide which mediates the transport of PR proteins through the secretory pathway [3]. In general, the basic and neutral isoforms, which are retained in the vacuole of the cell, have an additional carboxy-terminal extension of between 8 and 21 amino acids, that is absent in the acidic extracellular form, but there are some exceptions. These short carboxy-terminal propeptides have been shown to contain information for the vacuolar targeting of several PR proteins in plants and may be removed during or after transport to the plant vacuole [25, 27, 44].

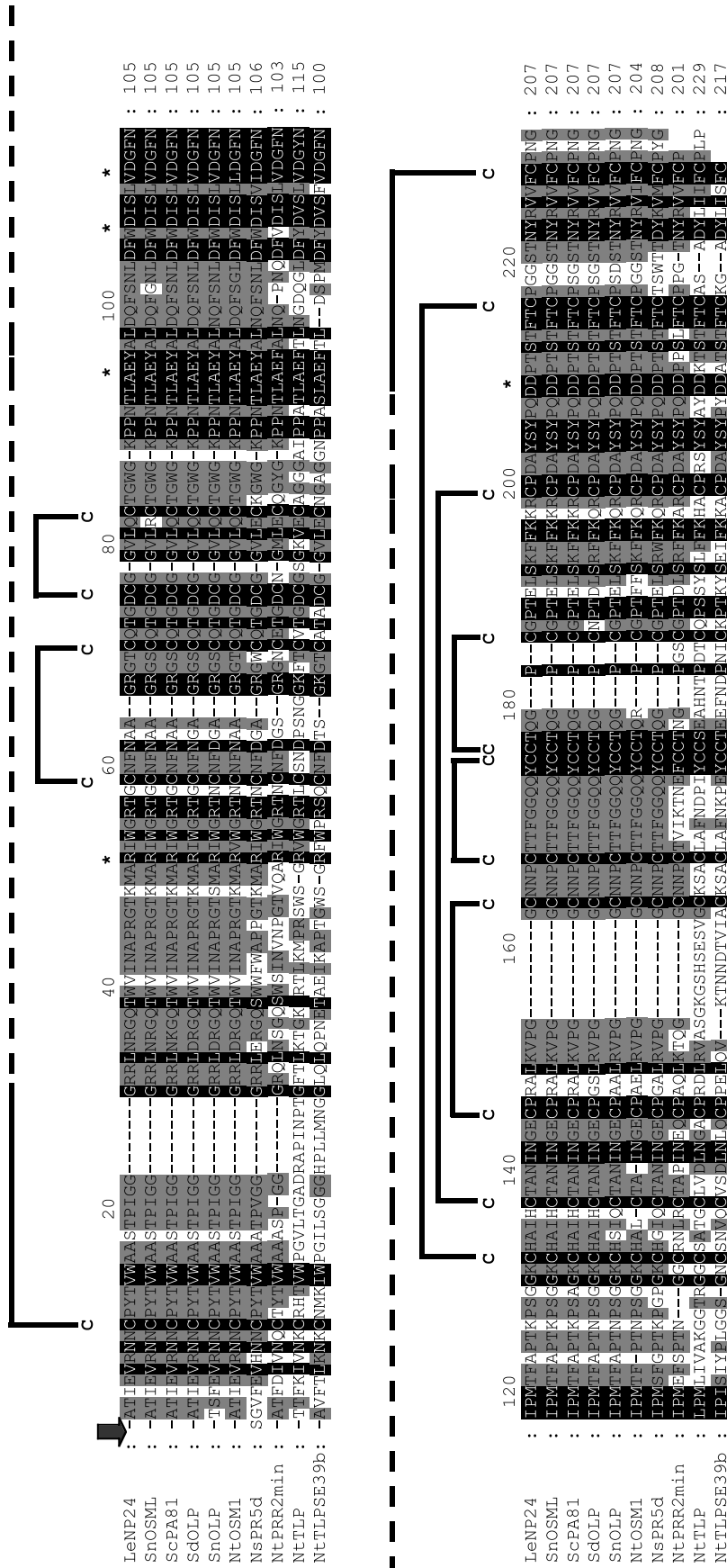


Fig. 6. Comparison of the mature deduced amino acid sequence of SnOLP and SnOSML, with eight solanaceous PR-5 proteins. The alignment was made using the PileUp program of the UWGGG package. GeneDoc software was used for homology shading. Identical amino acid residues and enable substitution groups are shaded in black when were present in all compared sequences and in gray when present in seven out of 10. The positions of the eight disulfide bonds (C-C) are marked. The cleavage site between the signal peptide and mature protein is indicated by the arrow. Asterisks show the key amino acid residues contributing to acidic cleft of tobacco PR-5d protein [18]. Codes, species, annotations and GenBank accession numbers are as follows; LeNP24 (*L. esculentum* NP24; GI6093527); ScOSML81 (*S. commersonii* OSML81; GI1709497); SdOLP (*Solanum dulcamara* OLP; GI10445203), NtOSM1 (*N. tabacum* Osmotin; GI486673); NsPR5d (*N. sylvestris* PR5d; GI13895534); NtPRR2min (*N. tabacum* PR-R2minor; GI131017) NtTLP (*N. tabacum* TLP; GI4239765; last three C-terminal amino acids not shown), NtTLPSE39b (*N. tabacum* TLPSE39b; GI4586372).

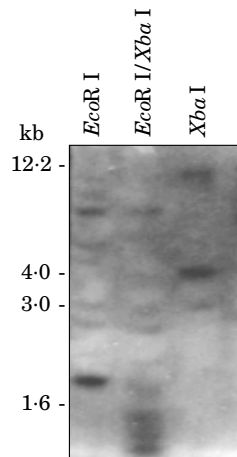


FIG. 7. Southern hybridization analysis of *S. nigrum* genomic DNA digested with the *EcoR* I and/or *Xba* I restriction enzymes (20  $\mu$ g DNA per lane). The blot was hybridized with the 744 bp fragment of the *SnOLP* gene probe; washing was carried out under conditions of high stringency.

#### Occurrence of PR5-like genes in *Solanum nigrum*

The Southern hybridization analysis of genomic DNA confirmed the presence of multiple PR5-like genes in *S. nigrum* var. *americanum*. When the genomic DNA digested with *EcoR* I restriction enzyme was hybridized with the ORF *SnOLP* probe, eight bands were produced, indicating that the predicted *SnOLP* and *SnOSML* PR5-like proteins are probably part of a multigene family. PR-5 multigene families (mgfs) of PR have already been shown for several *Solanum* species [55, 64] and oat [23] and other PR protein groups are also encoded in mgfs. The reason why plants produce a large number of diverse PRs is not clear. One likely cause could be evolutionary pressure towards the protection of the plant against different pathogens and abiotic stresses, since different PR-5 genes are activated by different signals (e.g. ABA, ethylene, auxin, infection by tobacco mosaic virus, salinity, lack of water, cold, u.v. light, wounding, and fungal/oomycete infection) [31, 36, 45, 64]. Which of these signals activates the cloned *S. nigrum* L. var. *americanum* genes remain to be studied.

The number and size of the bands obtained by Vleeshouwers *et al.* [55] with *S. nigrum* L. total DNA hybridization using tobacco osmotin gene probe [27] are different from the hybridization pattern we obtained with *S. nigrum* probed with the PR5-like *SnOLP* gene, although *EcoR* I was used to cut the DNA in both cases. This might indicate that divergent *S. nigrum* L. var. *americanum* genotypes were used. The technical procedures we used to clone the *SnOLP* and *SnOSML* PR5-like genes, i.e. PCR and hybridization with heterologous probe, represent a fast and efficient method that could be used

to identify other PR5-like genes, including those from the multigene family present in *S. nigrum* L. var. *americanum*.

#### Phylogenetic approach on Solanaceae family PR-5 proteins

PR-5 proteins have now been identified in the genomes of various plant species, including monocotyledons and dicotyledons [2, 11, 12, 28, 44, 60]. They have also been seen in nematodes [16]. In the Solanaceae, a phylogenetic analysis of PR-5 proteins (Fig. 8) consistently grouped the sequences into three clusters containing, interestingly, different members. In the first cluster there are four *N. tabacum* proteins (NtPR-R1 major, NtPR-R2 minor, NtTLPSE39b and NtTLP; Fig. 8), none of which possess a carboxy-terminal peptide in their precursor form, and all of which, therefore, are secreted into the extracellular space. Three are acidic (NtPR-R1 major, NtPR-R2 minor and NtTLPSE39b), the last basic (NtTLP). Notably, evolutionary distances on the phylogenetic tree show this group to be the most divergent in terms of sequence. They share only 43–73 % (mean 57 %) sequence identity with homologues of branches 2 and 3. This finding may be correlated with their extracellular location; as part of the first line of defence of plants against pathogens these proteins must be able to mutate at a rate sufficient to match the pathogens' attempts to acquire resistance. The phenomenon of host-pathogen co-evolution is well described [5, 8, 9, 49]. The second group contains three sequences ScOSM15, LePR-5 and NtPR5d which are acidic, basic and neutral, respectively and share 93–96 % sequence identity. This graphically supports the recent assertion that division of PR proteins based on their isoelectric point is not a useful scheme of classification [34]. The third major branch contains 13 other PR-5 homologues. Strikingly, although evolutionary distances between them are relatively small, and they share a mean sequence identity of 93 %, an enormous range of functions has been ascribed to members of the group. These proteins are induced by temperature stress (e.g. SdOLP), osmotic potential (e.g. NtOSM1-3), salt stress (e.g. LeNP24) or in response to attack by viruses (e.g. LeP23) or fungi (e.g. CaTLP). Hence, either relatively small sequence changes confer various activities or, perhaps more likely, functional annotation of members of the family is seriously incomplete, with a gene originally assigned a certain function, additionally possessing other unsuspected roles.

#### Computer-based analysis of the three-dimensional structure of the proteins

Analysis of the PR-5 protein structure, in combination with sequence analysis of antifungal and non-antifungal homologues, implicated two separate structural features in activity—a region of negative electrostatic potential resulting from the presence of four acidic residues



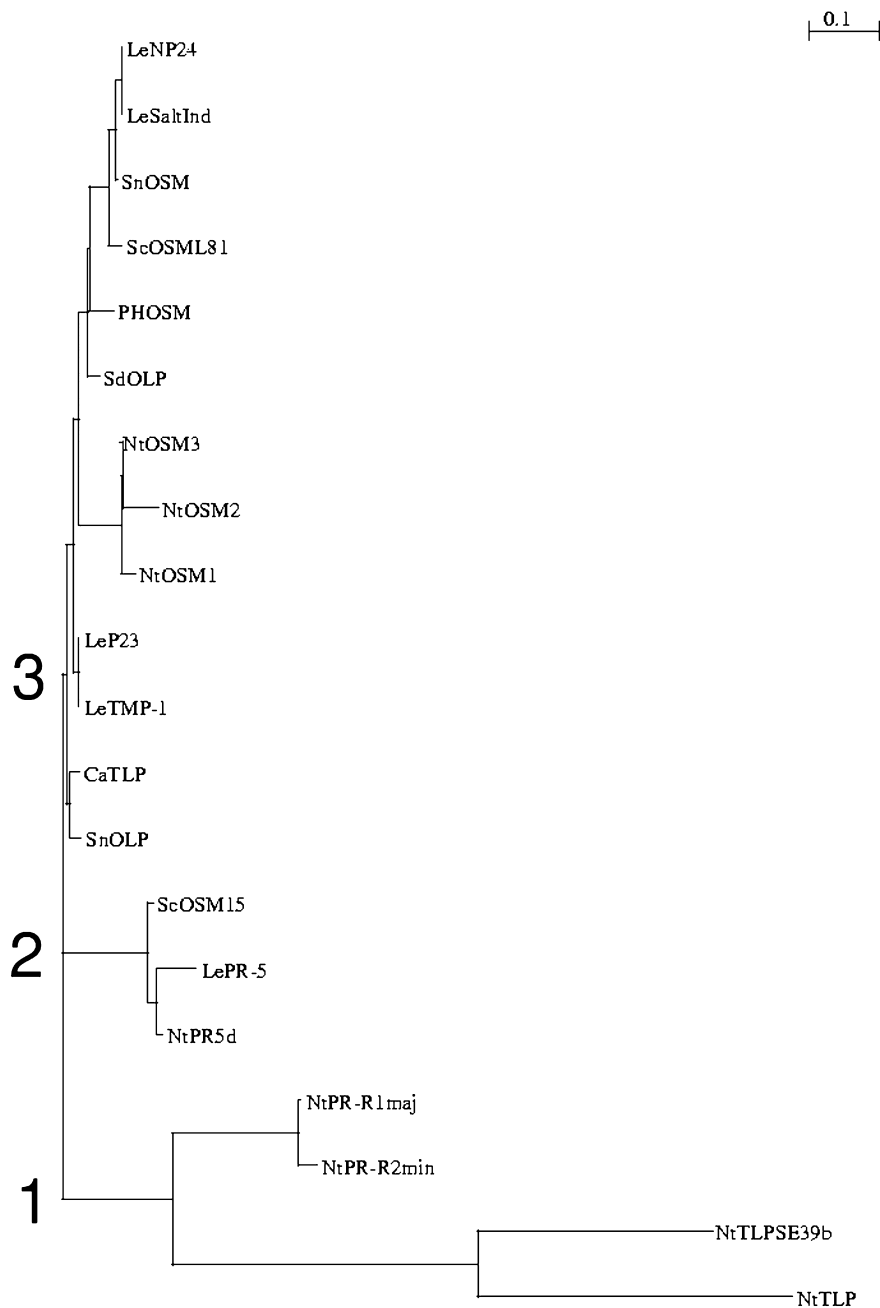


FIG. 8. A neighbour-joining tree of PR-5 homologues from Solanaceae constructed as detailed in Materials and Methods. Sequence codes not mentioned in the legend to Fig. 6 are the following; LeSaltInd (*L. esculentum* Salt induction protein; GI82110); PHOSM (*Petunia x hybrida* Osmotin; GI14165167); NtOSM3 (*N. tabacum* PR protein osmotin; GI1167854); NtOSM2 (*N. tabacum* Osmotin; GI119900); LeP23 (*L. esculentum* viroid-induction P23 protein; GI19314); LeTMP-1 (*L. esculentum* viroid-induction TPM-1 protein; GI418127); CaTLP (*Capsicum annuum* TLP; GI15419836); ScOSML15 (*S. commersonii* OSML15; GI1709496); LePR-5 (*L. esculentum* PR-5 protein; GI17414370); NtPR-R1maj (*N. tabacum* PR-R1 major protein; GI131015); (*N. tabacum* TLPSE39b; GI4586372); NtTLP (*N. tabacum* TLP; GI4239765).

grouped at the bottom of the interdomain cleft, and a hydrophobic patch [18]. This later originally contained two phenylalanine residues (F90 and F95 in SnOLP numbered), but a neighbouring leucine (L93) can also be added to this group which is similarly conserved as a

hydrophobic residue in antifungal proteins. These potentially functional components are identically positioned in both the SnOLP and SnOSML models and the PR-5 protein structure (Fig. 9). Furthermore, no other substitutions in the presumed active site cleft that might

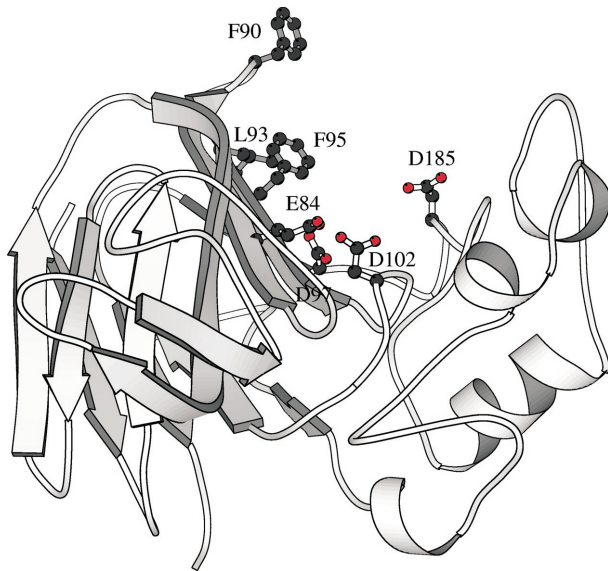


FIG. 9. Molscript [19] diagram of the model of SnOLP showing the two putative functional patches—the negatively-charged cleft base as a consequence of the four shown acidic residues, and the three residues forming the hydrophobic patch (shaded gray).

sterically or electrostatically inhibit activity are visible. Thus, these *S. nigrum* homologues potentially have all the known structural elements presumed necessary for activity. Although, the protein product of the PR-5 genes have never been reported in *S. nigrum*, overall sequence similarity and the conservation of key presumed functional elements strongly indicate that these genes may encode functional proteins.

### CONCLUDING REMARKS

PR-5 genes are currently used as markers of plant defence reactions associated with the systemic acquired resistance (SAR). Since SAR is associated with both local and systemic increased expression of many genes encoding PR proteins and it involves an enhanced state of resistance against a broad spectrum of pathogens [57], the availability of PR-5 genes from black nightshade is of great value to study these defence responses. Moreover, the PR genes cloned could also be used as probes to isolate PR5-like genes by traditional methods of screening or to determinate the PR-5 gene expression levels under different stress conditions. Considering the remarkable resistance of *S. nigrum* against the economically important oomycete *P. infestans* and the probable involvement of solanaceous PR-5 proteins in defence against this oomycete, PR-5 proteins from *S. nigrum* L. var. *americanum* appear extremely attractive candidates for control of fungi and oomycetes using transgenic plants. Whether the isolated PR-5 genes are indeed involved in non-host

resistance of *S. nigrum* L. var. *americanum* toward *P. infestans* remains to be determined and is the subject of current research.

We thank Eliane A. Gomes (Laboratory of Sequencing, Embrapa Genetic Resources and Biotechnology) for help with the sequencing, Marilia Santos Silva (Department of Virology, Wageningen University) for suggestions, and Antonia R. Figueira (Department of Phytopathology, Lavras University) for the use of the Laboratory of Molecular Biology. We acknowledge the financial support of EMBRAPA and CNPq.

### References

1. Abad LR, D'Urzo MP, Liu D, Narasimhan ML, Reuveni M, Zhu JK, Niu X, Singh NK, Hasegawa PM, Bressan RA. 1996. Antifungal activity of tobacco osmotin has specificity and involves plasma membrane permeabilization. *Plant Science* **118**: 11–23.
2. Anzlovar S, Serra MD, Demastia M, Menestrina G. 1998. Membrane permeabilizing of pathogenesis-related protein linusitin from flax seed. *Molecular Plant–Microbe Interactions* **7**: 610–617.
3. Bol JF, Linthorst HJM, Cornelissen BJC. 1990. Plant pathogenesis-related proteins induced by virus infection. *Annual Review of Phytopathology* **28**: 113–138.
4. Colon LT, Eijlander R, Bulding DJ, van Ijzendoorn MT, Pieters MMJ, Hoogendoorn J. 1993. Resistance to potato late blight (*Phytophthora infestans* (Mont.) de Bary) in *Solanum nigrum*, *S. villosum* and their sexual hybrids with *S. tuberosum* and *S. demissum*. *Euphytica* **66**: 55–64.
5. Dybdahl M, Lively C. 1998. Host–parasite coevolution: evidence for rare advantage and time-lagged selection in a natural population. *Evolution* **52**: 1057–1066.
6. Faus I, Patiño C, del Río JL, del Moral C, Barroso HS, Rubio V. 1996. Expression of a synthetic gene encoding the sweet-tasting protein thaumatin in *Escherichia coli*. *Biochemical and Biophysical Research Communications* **229**: 121–127.
7. Felsenstein J. 1989. PHYLIP—Phylogeny Inference Package (Version 3.2). *Cladistics* **5**: 164–166.
8. Frank SA. 1992. Models of plant–pathogen coevolution. *Trends in Genetics* **8**: 213–219.
9. Frank SA. 1993. Coevolutionary genetics of plants and pathogens. *Evolutionary Ecology* **7**: 45–75.
10. Hammond-Kosack K, Jones JDG. 2000. Response to plant pathogens. In: Buchanan B, Gruissem R, Jones R, eds. *Biochemistry and Molecular Biology of Plants*. Rockville, MD: American Society of Plant Physiologists, 1102–1156.
11. Hejgaard J, Jacobsen S, Svendsen I. 1991. Two antifungal thaumatin-like proteins from barley grain. *FEBS Letters* **291**: 127–131.
12. Hu X, Reddy ASN. 1997. Cloning and expression of a PR5-like protein from *Arabidopsis*: inhibition of fungal growth by bacterially expressed protein. *Plant Molecular Biology* **34**: 949–959.
13. Ibeas JI, Lee H, Damsz B, Prasad DT, Pardo JM, Hasegawa PM, Bressan RA, Narasimhan ML. 2000. Fungal cell wall phosphomannans facilitate the toxic

- activity of a plant PR-5 protein. *The Plant Journal* **23**: 375–383.
14. **Kamoun S, Huitema E, Vleeshouwers VGAA.** 1999. Resistance to oomycetes: a general role for the hypersensitive response? *Trends in Plant Science* **4**: 196–200.
  15. **Karplus K, Barrett C, Hughey R.** 1998. Hidden markov models for detecting remote protein homologies. *Bioinformatics* **14**: 846–856.
  16. **Kitajima S, Sato F.** 1999. Plant pathogenesis-related proteins: molecular mechanisms of gene expression and protein function. *The Journal of Biochemistry* **125**: 1–8.
  17. **Koiwa H, Sato F, Yamada Y.** 1994. Characterization of accumulation of tobacco PR-5 proteins by IEF-immunoblot analysis. *Plant Cell Physiology* **35**: 821–827.
  18. **Koiwa H, Kato H, Nakatsu T, Oda J, Yamada Y, Sato F.** 1999. Crystal structure of tobacco PR-5d protein at 1.8 Å resolution reveals a conserved acidic cleft structure in antifungal thaumatin-like proteins. *Journal of Molecular Biology* **286**: 1137–1145.
  19. **Kraulis PJ.** 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *Journal of Applied Crystallography* **24**: 946–950.
  20. **LaRosa PC, Singh NK, Hasegawa PM, Bressan RA.** 1989. Stable NaCl tolerance of tobacco cells is associated with enhanced accumulation of osmotin. *Plant Physiology* **91**: 855–861.
  21. **Ledeboer AM, Verrips CT, Dekker BMM.** 1984. Cloning of the natural gene for the sweet-tasting plant protein thaumatin. *Gene* **30**: 23–32.
  22. **Li R, Wu N, Fan Y, Song B.** 1999. Transgenic potato plants expressing osmotin gene inhibits fungal development in inoculated leaves. *Chinese Journal of Biotechnology* **15**: 71–75.
  23. **Lin K, Bushnell WR, Szabo LJ, Smith AG.** 1996. Isolation and expression of a host response gene family encoding thaumatin-like proteins in incompatible oat-stem rust fungus interactions. *Molecular Plant-Microbe Interactions* **9**: 511–522.
  24. **Liu D, Raghothama KG, Hasegawa PM, Bressan RA.** 1994. Osmotin overexpression in potato delays development of disease symptoms. *Proceedings of the National Academy of Sciences, U.S.A.* **91**: 1888–1892.
  25. **Liu D, Rhodes D, D'Urzo MP, Xu Y, Narasimhan ML, Hasegawa PM, Bressan RA, Abad L.** 1996. *In vivo* and *in vitro* activity of truncated osmotin that is secreted into the extracellular matrix. *Plant Science* **121**: 123–131.
  26. **Malehorn DE, Borgmeyer JR, Smith CE, Shah DM.** 1994. Characterization and expression of an antifungal zeamatin-like protein (*zlp*) gene from *Zea mays*. *Plant Physiology* **106**: 1471–1481.
  27. **Melchers LS, Sel-Buurlage MB, Vloemans SA, Woloshuk CP, Van Roekel JSC, Pen J, van der Elzen PJM, Cornelisen BJC.** 1993. Extracellular targeting of the vacuolar tobacco proteins AP24, chitinase and  $\beta$ -1,3-glucanase in transgenic plants. *Plant Molecular Biology* **21**: 583–593.
  28. **Morris SW, Vernooij B, Titartn S, Starrett M, Thomas S, Wiltse CC, Frederiksen RA, Bhandhufalck A, Hulbert S, Uknes S.** 1998. Induced resistance responses in maize. *Molecular Plant-Microbe Interactions* **11**: 643–658.
  29. **Narasimhan ML, Damsz B, Coca MA, Ibeas JI, Yun DJ, Pardo JM, Hasegawa PM, Bressan RA.** 2001. A plant defense response effector induces microbial apoptosis. *Molecular Cell* **8**: 921–930.
  30. **Neale AD, Wahleithner JA, Lund M, Bonnett HT, Kelly A, Meeks-Wagner DR, Peacock WJ, Dennis ES.** 1990. Chitinase,  $\beta$ -1,3 glucanase, osmotin, and extensin are expressed in tobacco explants during flower formation. *The Cell* **2**: 673–684.
  31. **Nelson DE, Raghothama KG, Singh NK, Hasegawa PM, Bressan RA.** 1992. Analysis of structure and transcriptional activation of an osmotin gene. *Plant Molecular Biology* **19**: 577–588.
  32. **Newton SS, Duman JG.** 2000. An osmotin-like cryoprotective protein from the bitter-sweet nightshade *Solanum dulcamara*. *Plant Molecular Biology* **44**: 581–589.
  33. **Nielsen H, Engelbrecht J, Brunak S, von Heijne G.** 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering* **10**: 1–6.
  34. **Osmond RI, Hrmova M, Fontaine F, Imberty A, Fincher GB.** 2001. Binding interactions between barley thaumatin-like proteins and (1,3)-beta-D-glucans. Kinetics, specificity, structural analysis and biological implications. *European Journal of Biochemistry* **268**: 4190–4199.
  35. **Pontier D, Balague C, Roby D.** 1998. The hypersensitive response. A programmed cell death associated with plant resistance. *Comptes Rendus de l'Académie des Sciences Paris* **321**: 721–734.
  36. **Raghothama KG, Maggio A, Narasimhan ML, Kononowicz AK, Wang G, D'Urzo MP, Hasegawa PM, Bressan RA.** 1997. Tissue-specific activation of the osmotin gene by ABA, C<sub>2</sub>H<sub>4</sub> and NaCl involves the same promoter region. *Plant Molecular Biology* **34**: 393–402.
  37. **Roberts WK, Selitrennikoff CP.** 1990. Zeamatin, an antifungal protein from maize with membrane-permeabilizing activity. *Journal of General Microbiology* **136**: 1771–1778.
  38. **Rogers SO, Bendich AJ.** 1988. Extraction of DNA from plant tissues. In: Gelvin SB, Schilperoort RA, eds. *Plant Molecular Biology Manual*. Dordrecht: Kluwer Academic Publishers, Section 6, 1–21.
  39. **Romano E.** 1998. Análise da Integração do DNA pela técnica Southern blot. In: Brasileiro ACM, Carneiro VTC, eds. *Manual de Transformação Genética de Plantas*. Brasília, DF, Brasil: Embrapa-SPI/Embrapa-Cenargen, 205–222.
  40. **Ross H.** 1986. Potato breeding—problems and perspectives. In: Horn W, Röbbelen G, eds. *Journal of Plant Breeding supplement 13. Advances in Plant Breeding*. Berlin, Hamburg: Verlag Paul Parey, 132 pp.
  41. **Ryals JA, Neuenschwarnder UH, Willits MG, Molina A, Steiner H, Hunt MD.** 1996. Systemic acquired resistance. *The Plant Cell* **8**: 1809–1819.
  42. **Sali A, Blundell TL.** 1993. Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology* **234**: 779–815.
  43. **Salzman RA, Tikhonova I, Bordlon BP, Hasegawa PM, Bressan RA.** 1998. Coordinate accumulation of antifungal proteins and hexoses constitutes a developmentally controlled defense response during fruit ripening in grape. *Plant Physiology* **117**: 465–472.
  44. **Sato F, Koiwa H, Sakai Y, Kato N, Yamada Y.** 1995. Synthesis and secretion of tobacco neutral PR-5 protein by transgenic tobacco and yeast. *Biochemical and Biophysical Research Communications* **211**: 909–913.
  45. **Sato F, Kitajima S, Koyama T, Yamada Y.** 1996. Ethylene-induced gene expression of osmotin-like protein, a neutral isoform of tobacco PR-5, is mediated by the AGCCGCC *cis*-sequence. *The Plant Cell Physiology* **37**: 249–255.

46. Singh NK, Braker CA, Hasegawa PM, Handa AK, Hermodson MA, Pfankoch E, Regnier FE, Bressan RA. 1987. Characterization of osmotin. *Plant Physiology* **85**: 529–536.
47. Southern EM. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**: 503–517.
48. Takemoto D, Katsumi F, Doke N, Kawakita K. 1997. Identification of chitinase and osmotin-like protein as actin-binding proteins in suspension-cultures potato cells. *Plant Cell Physiology* **38**: 441–448.
49. Thompson JN, Burdon JJ. 1992. Gene-for-gene coevolution between plants and parasites. *Nature* **360**: 121–125.
50. Van der Wel H, Loeve K. 1972. Isolation and characterization of thaumatin I and II. The sweet-tasting protein from *Thaumatococcus danielli*. *European Journal of Biochemistry* **31**: 221–225.
51. Van Loon LC, Gerritsen YAM, Ritter CE. 1987. Identification, purification and characterization of pathogenesis-related proteins from virus-infected Samsun NN tobacco leaves. *Plant Molecular Biology* **9**: 593–609.
52. Van Loon LC, Pierpoint WS, Boller T, Conejero V. 1994. Recommendations for naming plant pathogenesis-related proteins. *Plant Molecular Biology Reporter* **12**: 245–264.
53. Van Loon LC, Van Strien EA. 1999. The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 proteins. *Physiological and Molecular Plant Pathology* **55**: 85–97.
54. Vigers AJ, Roberts WK, Selitrennikoff CP. 1991. A new family of plant antifungal proteins. *Molecular Plant–Microbe Interactions* **4**: 315–323.
55. Vleeshouwers VGAA, Van Dooijeweert W, Govers F, Kamoun S, Colon LT. 2000. Does basal PR gene expression in *Solanum* species contribute to non-specific resistance to *Phytophthora infestans*? *Physiological and Molecular Plant Pathology* **57**: 35–42.
56. Vleeshouwers VGAA, Van Dooijeweert W, Govers F, Kamoun S, Colon LT. 2000. The hypersensitive response is associated with host and nonhost resistance to *Phytophthora infestans*? *Planta* **210**: 853–864.
57. Ward ER, Uknes SJ, Williams SC, Dincher SS, Wiederhold DL, Alexander DC, Ahl-Goy P, Métraux J-P, Ryals JA. 1991. Co-ordinate gene activity in response to agents that induce systemic acquired resistance. *The Plant Cell* **3**: 1085–1094.
58. Woloshuk CP, Meulenhoff JS, Sela-Buurlage M, van der Elsen PJM, Cornelissen BJC. 1991. Pathogen-induced proteins with inhibitory activity toward *Phytophthora infestans*. *The Plant Cell* **3**: 619–628.
59. Xu Y, Chang P-FL, Liu D, Narashimhan ML, Raghothama KG, Hasegawa PM, Bressan RA. 1994. Plant defense genes are synergistically induced by ethylene and methyl jasmonate. *The Plant Cell* **6**: 1077–1085.
60. Yen HE, Edwards GE, Grimes HD. 1994. Characterization of a salt-responsive 24-kilodalton glycoprotein in *Mesembryanthemum crystallinum*. *Plant Physiology* **105**: 1179–1187.
61. Yun DJ, Zhao Y, Pardo JM, Narasimhan ML, Damsz B, Lee M, Abad LR, D'Urzo MP, Hasegawa PM, Bressan RA. 1997. Stress proteins on the yeast cell surface determine resistance to osmotin, a plant antifungal protein. *Proceedings of the National Academy of Sciences, U.S.A.* **94**: 7082–7087.
62. Zhu B, Chen THH, Li PH. 1993. Expression of an ABA-responsive osmotin-like protein gene during the induction of freezing tolerance in *Solanum commersonii*. *Plant Molecular Biology* **21**: 729–735.
63. Zhu B, Chen THH, Li PH. 1995. Expression of three osmotin-like protein genes in response to osmotic stress and fungal infection in potato. *Plant Molecular Biology* **28**: 17–26.
64. Zhu B, Chen THH, Li PH. 1995. Activation of two osmotin-like protein genes by abiotic stimuli and fungal pathogen in transgenic potato plants. *Plant Physiology* **108**: 929–937.
65. Zhu B, Chen THH, Li PH. 1996. Analysis of late-blight disease and freezing tolerance in transgenic potato plants expressing sense and antisense genes for an osmotin-like protein. *Planta* **198**: 70–77.