



# *In silico* comparative analysis of tylenchid nematode pectate lyases

R.M. Moraes Filho<sup>1</sup> and L.S.S. Martins<sup>2,3</sup>

<sup>1</sup>Departamento de Agronomia da Universidade Federal Rural de Pernambuco, Recife, PE, Brasil

<sup>2</sup>Departamento de Biologia da Universidade Federal Rural de Pernambuco, Recife, PE, Brasil

<sup>3</sup>Programa de Pós-Graduação em Agronomia-Melhoramento Genético de Plantas, Universidade Federal Rural de Pernambuco, Recife, PE, Brasil

Corresponding author: R.M. Moraes Filho

E-mail: romulommfilho@yahoo.com.br

Genet. Mol. Res. 15 (3): gmr.15038402

Received January 7, 2016

Accepted February 26, 2016

Published August 18, 2016

DOI <http://dx.doi.org/10.4238/gmr.15038402>

Copyright © 2016 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution ShareAlike (CC BY-SA) 4.0 License.

**ABSTRACT.** Phytoparasitic nematodes can infect a wide range of crop plants, and cause billions of dollars of agricultural losses each year. These parasites represent the largest source of biotic stress experienced by plants. The order Tylenchida comprises the most important parasitic nematodes, particularly the root-knot and cyst nematodes. These parasitic organisms obtain nutrients to support their development through complex interactions with their hosts. Plant-parasitic nematodes secrete a mixture of cell-wall degrading enzymes to facilitate migration through the plant root. Enzymes are secreted that degrade the principal cell-wall components, cellulose, hemi-cellulose, or pectin. Pectate lyases are important parasitism factors in plant-parasitic nematodes. These enzymes degrade polygalacturonic acid, which is a fundamental constituent of pectin of host cell walls. Thus, pectate lyases permit the

penetration and colonization of plant host cells by parasites. Here, we analyzed 22 pectate lyase protein sequences from tylenchid nematode species. Our results revealed great variation in the isoelectric points of pectate lyases, and groups of acidic and alkaline proteins that may have distinct enzymatic activities were identified. Phylogenetic analysis also revealed the presence of two main groups of pectate lyases with distinct chemical properties. Seven conserved motifs were identified, but only five were present in all sequences. Results of the molecular docking analysis revealed differences in the predicted interaction sites in the pectate lyases from the two groups. These results may provide a theoretical basis for future studies of host plant resistance to nematode infection.

**Key words:** Computational analysis; Parasitism resistance; Homology modeling; Cell wall-degrading enzymes; Tylenchida

## INTRODUCTION

The phylum Nematoda comprises up to 25,000 described species. Almost all crop plants possess one or more species of nematodes in the form of ecto- or endo-parasites. The order Tylenchida is the most important parasitic nematode group, which causes billions of dollars of agricultural losses each year, particularly the devastating root-knot (*Meloidogyne* spp) and cyst (*Heterodera* and *Globodera* spp) nematodes. These parasitic organisms represent the largest source of biotic stress experienced by plants and can cause stunting, early senescence, and, in severe cases, total crop loss (Bird, 2004; Perry and Moens, 2011).

Phytonematodes have evolved complex parasitic relationships with their host plants, enabling them to obtain the necessary nutrients to support their development and reproduction (Hussey et al., 2002). These microscopic parasites penetrate the roots of host plants as second-stage juveniles and induce dramatic changes in selected root vascular cells. This forms elaborate feeding cells to permanently supply nutrients that enable the nematodes to develop into reproductive adults. Both cyst and root-knot nematodes secrete cell-wall degrading enzymes (CWDE) from the subventral glands to facilitate migration through the plant root. The secreted protein mixture consists of enzymes that degrade the main cell wall components (Williamson and Gleason, 2003; Davis et al., 2008; Harris and Stone, 2009; Perry and Moens, 2011). Phytonematodes produce specific sets of CWDEs based on the cell-wall composition of the host. Inability to degrade any cell-wall component may result in unsuccessful infection or in the survival of the host plant (Rai et al., 2015).

Pectin is an important structural component of primary cell walls and is found in high concentrations in the middle lamella between plant cells. Therefore, a secreted pectinase is necessary for the infection of roots by nematodes (Huang et al., 2005; Rai et al., 2015).

Pectate lyases (PL, EC 4.2.2.2) are a class of polysaccharidases that function as pathogenicity factors for many phytopathogenic organisms. These enzymes play essential roles in nematode infection and parasitism of plants by degrading polygalacturonic acid, which is a fundamental constituent of pectin of host cell walls, permitting penetration and colonization. At least five classes of PLs with different characteristics are observed among bacteria and fungi. PLs identified in phytonematodes are absent in non-parasitic nematodes

and almost all other animals. The similarity of PLs to class III PLs of fungi and bacteria at the amino acid level suggests that they were acquired by plant parasitic nematodes via horizontal gene transfer from prokaryotic organisms (Tardy et al., 1997; Smant et al., 1998; Davis et al., 2000; De Boer et al., 2002; Doyle and Lambert, 2002; Scholl et al., 2003).

*In silico* analysis can be of great value for predicting structures and functions of proteins, and has been used to characterize many proteins and enzymes from diverse eukaryotic and prokaryotic species (Feng et al., 2015; Han et al., 2015; Vatansever et al., 2015). To date, genes encoding PLs have been cloned from several species of tylenchid nematodes from *Heterodera*, *Globodera*, *Meloidogyne*, *Bursaphelenchus*, and *Aphelenchus* genera (Popeijus et al., 2000; De Boer et al., 2002; Doyle and Lambert, 2002; Kikuchi et al., 2004; Huang et al., 2005; Kudla et al., 2007; Karim et al., 2009).

Therefore, the aim of the present study was to characterize and compare PL protein sequences from tylenchid nematode species in order to identify conserved domains, to predict their tertiary structures, and to obtain knowledge on their functions and molecular interactions. This comparative analysis will provide valuable theoretical insights for future studies involving these proteins and host plant resistance to nematode infection.

## MATERIAL AND METHODS

### Database search and sequence retrieval

Three types of PL proteins from *Meloidogyne incognita* (Mi-PL1, AAS88579.1; Mi-PL2, AAQ97032.1; Mi-PL3, AY861685.1) were used as queries in the tool BLASTn from the National Center for Biotechnology Information (NCBI, <http://www.ncbi.nlm.nih.gov>) database. Other complete protein sequences of different tylenchid nematode species that had >30% identity with the queries were collected, totaling 22 PL sequences.

### Sequence analysis

Physico-chemical parameters of PL sequences were analyzed by ProtParam (<http://web.expasy.org/protparam>) (Gasteiger et al., 2005). Subcellular localizations were predicted by the CELLO2GO server (Yu et al., 2014). Signal peptide cleavage sites were predicted using the TOPCONS server (<http://topcons.cbr.su.se/>) (Tsirigos et al., 2015).

### Conserved motif analysis

Conserved motif structure was analyzed by using the MEME SUITE tool (Bailey et al., 2009) with the following parameters: maximum number of motifs to find = 7; minimum width of motif = 6; and maximum width of motif = 50.

### Phylogenetic analysis

Sequence alignment of PL proteins was performed with ClustalW algorithm implemented in Molecular Evolutionary Genetic Analysis (MEGA 6.06) (Tamura et al., 2011), with default parameters. The phylogenetic tree was constructed using the neighbor-joining method for 2000 bootstrap replicates.

## Protein-protein interactions

The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING 9.1) database (Franceschini et al., 2013) (<http://string-db.org/>) was used to predict potential interacting proteins. The database contains information from numerous sources, including experimental repositories, computational prediction methods, and public text collections.

## Tertiary structure prediction, evaluation, and validation of the model

3-D models of tylenchid nematode PLs were predicted using the Phyre2 server (<http://www.sbg.bio.ic.ac.uk/phyre2>) (Kelley et al., 2015) in multi-template intensive mode and visualized by UCSF Chimera package (Pettersen et al., 2004). Model quality was evaluated using the Molprobit server (<http://molprobit.biochem.duke.edu/>) (Chen et al., 2010) by Ramachandran plot analysis. Z-score was calculated using interactive ProSA-web server (<https://prosa.services.came.sbg.ac.at/prosa.php>) to recognize errors in 3-D structures, which indicated model quality and total energy deviation of the structure with respect to energy distribution derived from random conformations (Wiederstein and Sippl, 2007).

## Molecular docking and prediction of binding sites

Ligand binding sites were predicted using the Patch Dock server (<http://bioinfo3d.cs.tau.ac.il/PatchDock/>) (Schneidman-Duhovny et al., 2005).

## RESULTS

### Identification and characterization of PLs

In this study, 22 PL sequences from five families of the Tylenchida order (Anguinidae, Aphelenchidae, Aphelenchoididae, Heteroderidae, and Meloidogynidae) were retrieved from NCBI in FASTA format (Table 1). Five sequences were identified in the database as PL1, three as PL2, and one as PL3. The remaining sequences were identified as PL without a specific type. The TOPCONS server was used to predict post-translational modifications. This tool predicts the presence and location of signal peptide cleavage sites and transmembrane regions within protein sequences. TOPCONS analysis showed that all evaluated sequences possess N-terminal signal peptides (Table 2). The identified signal peptides were 21-24 aa long and occurred 0-6 aa before a conserved proline. In order to investigate structures and functions, physico-chemical parameters of PL sequences were analyzed. As shown in Table 2, after the removal of signal peptides, PL sequences varied in size from 226 aa (*Aphelenchus avenae*) to 257 aa (*Meloidogyne incognita* PL3) and from 23.8 kDa (*A. avenae*) to 32.2 kDa (*M. incognita* PL3). Isoelectric point (pI) was between 5.25 (*Bursaphelenchus mucronatus* 2) and 9.42 (*Globodera tabacum*). The estimated average hydropathicity was between -0.183 (*G. pallida*) and -0.904 (*M. enterolobii*). These results indicate that the evaluated PLs were hydrophilic. Analysis of subcellular location using the CELLO2GO server predicted that most of the evaluated PLs are extracellular proteins. Only the *A. avenae* (BAI44497.1) PL was identified as cytoplasmic.

**Table 1.** Twenty-two evaluated pectate lyases and their species of origin.

Species	GenBank accession No.	GenBank definition	Family	Common name
<i>Aphelenchus avenae</i>	BAI44497.1	pectate lyase	Aphelenchidae	Mycophagous nematode worm
<i>Bursaphelenchus mucronatus</i>	BAE48373.1	pectate lyase	Aphelenchoididae	Pine wood nematode
<i>Bursaphelenchus mucronatus</i>	BAE48375.1	pectate lyase	Aphelenchoididae	Pine wood nematode
<i>Bursaphelenchus xylophilus</i>	BAE48369.1	pectate lyase	Aphelenchoididae	Pine wilt nematode
<i>Bursaphelenchus xylophilus</i>	BAE48370.1	pectate lyase	Aphelenchoididae	Pine wilt nematode
<i>Diitylenchus destructor</i>	AFL48198.1	pectate lyase	Anguinidae	Potato rot nematode
<i>Globodera pallida</i>	AEA08853.1	pectate lyase 1	Heteroderidae	Cyst nematode
<i>Globodera rostochiensis</i>	AAF80746.1	pectate lyase 1	Heteroderidae	Cyst nematode
<i>Globodera rostochiensis</i>	AHW98765.1	pectate lyase 2	Heteroderidae	Cyst nematode
<i>Globodera tabacum</i>	AEA08813.1	pectate lyase 1	Heteroderidae	Cyst nematode
<i>Globodera virginiae</i>	AEA08827.1	pectate lyase 1	Heteroderidae	Cyst nematode
<i>Heterodera glycines</i>	ADW77534.1	pectate lyase	Heteroderidae	Cyst nematode
<i>Heterodera glycines</i>	AAM74954.1	pectate lyase 2	Heteroderidae	Cyst nematode
<i>Heterodera schachtii</i>	ABN14272.1	pectate lyase	Heteroderidae	Cyst nematode
<i>Heterodera schachtii</i>	ABN14273.1	pectate lyase	Heteroderidae	Cyst nematode
<i>Meloidogyne enterolobii</i>	ADN87334.1	pectate lyase	Meloidogynidae	Root-knot nematode
<i>Meloidogyne enterolobii</i>	ALB38961.1	pectate lyase 2	Meloidogynidae	Root-knot nematode
<i>Meloidogyne graminicola</i>	AID59201.1	pectate lyase	Meloidogynidae	Root-knot nematode
<i>Meloidogyne incognita</i>	AAS88579.1	pectate lyase 1	Meloidogynidae	Root-knot nematode
<i>Meloidogyne incognita</i>	AAQ97032.1	pectate lyase 2	Meloidogynidae	Root-knot nematode
<i>Meloidogyne incognita</i>	AY861685.1	pectate lyase 3	Meloidogynidae	Root-knot nematode
<i>Meloidogyne javanica</i>	AAL66022.1	pectate lyase	Meloidogynidae	Root-knot nematode

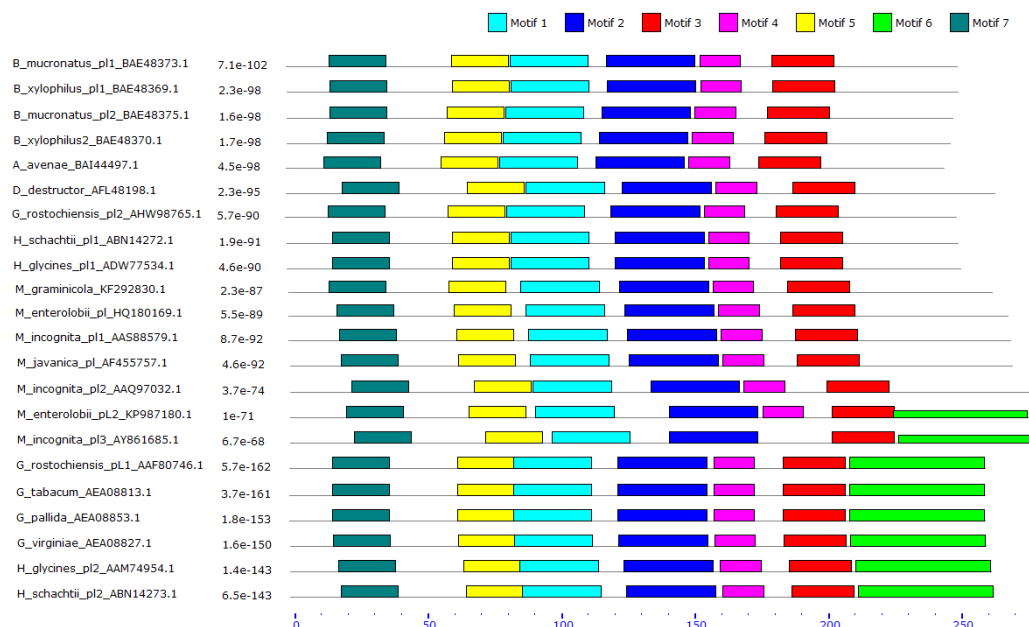
**Table 2.** Primary structure analysis, signal peptides, and subcellular location of pectate lyases.

Origin species	Protein size	MW (kDa)	pI	GRAVY	SPCS	ScL
<i>A. avenae</i> pl_BAI44497.1	247/226	26.0/23.8	8.94/8.93	-0.338/-0.532	A21-A22	Cy
<i>B. mucronatus</i> pl_BAE48373.1	252/231	26.7/24.4	8.43/7.74	-0.373/-0.530	A21-A22	Ex, Cy
<i>B. mucronatus</i> pl_BAE48375.1	250/229	26.0/23.9	5.11/5.25	-0.159/-0.350	A21-A22	Ex, Cy
<i>B. xylophilus</i> pl_BAE48369.1	252/231	26.6/24.5	8.14/8.45	-0.322/-0.573	A21-A22	Ex, Cy
<i>B. xylophilus</i> pl_BAE48370.1	249/228	25.9/23.7	5.66/5.97	-0.168/-0.368	A21-A22	Cy, Ex
<i>D. destructor</i> pl_AFL48198.1	266/244	28.7/26.3	5.64/5.25	-0.322/-0.539	A22-A23	Ex
<i>G. pallida</i> pl1_AEA08853.1	261/240	27.9/25.6	9.24/9.28	-0.007/-0.183	A21-A22	Ex
<i>G. rostochiensis</i> pl1_AAF80746.1	261/240	28.0/25.6	9.28/9.32	-0.020/-0.200	A21-A22	Ex
<i>G. rostochiensis</i> pl2_AHW98765.1	252/231	27.0/24.5	8.29/8.35	-0.077/-0.238	A21-A22	Ex
<i>G. tabacum</i> pl1_AEA08813.1	261/240	28.0/25.7	9.38/9.42	-0.071/-0.256	A21-A22	Ex
<i>G. virginiae</i> pl1_AEA08827.1	261/240	28.3/26.0	9.22/9.26	-0.111/-0.300	A21-A22	Ex
<i>H. glycines</i> pl_ADW77534.1	253/232	26.8/24.5	8.89/8.90	-0.226/-0.353	A21-A22	Ex
<i>H. glycines</i> pl2_AAM74954.1	263/242	28.0/25.7	9.21/9.21	-0.048/-0.243	A21-A22	Ex
<i>H. schachtii</i> pl_ABN14272.1	252/231	26.8/24.5	9.24/9.24	-0.263/-0.398	A21-A22	Ex
<i>H. schachtii</i> pl_ABN14273.1	264/243	28.2/25.9	9.34/9.30	-0.128/-0.240	A21-A22	Ex
<i>M. enterolobii</i> pl_ADN87334.1	270/248	29.8/27.4	6.30/5.89	-0.746/-0.904	A22-A23	Ex
<i>M. enterolobii</i> pl2_ALB38961.1	277/255	29.8/27.5	8.87/8.97	-0.137/-0.281	A22-A23	Ex
<i>M. graminicola</i> pl_AID59201.1	264/243	29.3/26.6	7.99/8.17	-0.461/-0.596	A21-A22	Ex, Cy
<i>M. incognita</i> pl1_AAS88579.1	271/249	29.6/27.2	6.06/5.74	-0.638/-0.786	A22-A23	Ex
<i>M. incognita</i> pl2_AAQ97032.1	280/256	30.6/27.7	9.02/8.86	-0.147/-0.282	A24-A25	Ex
<i>M. incognita</i> pl3_AY861685.1	279/257	30.5/28.0	9.30/9.25	-0.131/-0.293	A24-A25	Ex
<i>M. javanica</i> pl_AAL66022.1	271/249	29.8/27.3	5.83/5.55	-0.621/-0.773	A22-A23	Ex

GRAVY, grand average of hydropathicity; MW, molecular weight; pI, isoelectric point are reported as precursor values/modeled protein values. ScL, subcellular locations are reported as Cy (cytoplasmic) and Ex (extracellular).

## Multiple alignment and identification of conserved motifs

Conserved motifs of PLs were analyzed using the MEME Suite tool. A conserved motif is a sequence or pattern that occurs consistently in a group of related protein sequences. Based on the results, seven conserved motifs were discovered (Figure 1 and Table 3). Five of the seven motifs occurred in all analyzed sequences. Motif 4 was absent in the *M. incognita* PL3 sequence. Motif 6 was present in eight sequences.



**Figure 1.** Conserved motifs identified in evaluated pectate lyases (PLs) using the MEME SUITE tool. The position of each block indicates where a motif has matched the sequence. The width of the blocks indicates the width of the motif relative to the length of the sequence. The color and border of the blocks are used to identify the matching motif, as shown in the legend. The height of the blocks represents the significance of the match, with taller blocks being more significant.

**Table 3.** Conserved motifs identified using the MEME tool in 22 pectate lyase precursor sequences.

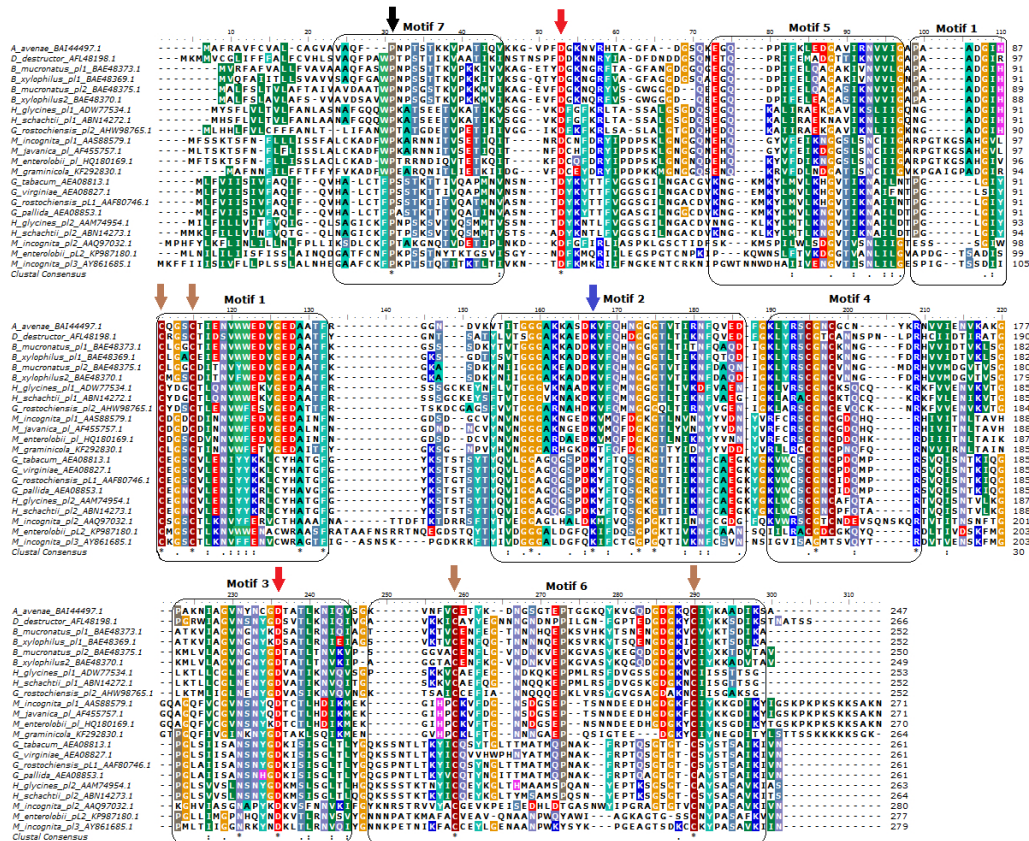
Motif	E-value	Width	Best possible match	Conserved domain
1	3.4e <sup>-326</sup>	29	PPAHGIHCEGSCSTLENVWVERVCEHAATF	Pectate lyase
2	2.3e <sup>-309</sup>	33	YQVTGGGARHAPDKVFOHSGRGLIHKNFCAED	Pectate lyase
3	9.5e <sup>-159</sup>	23	PGLVIAGVNSNYGDKATLSNIQI	-
4	3.9e <sup>-133</sup>	15	KLWRSCGNCPDQMPR	-
5	9.0e <sup>-144</sup>	21	EHQKYLMELEHGATIKNLIIG	-
6	1.2e <sup>-120</sup>	50	QKSPNLTLYICQEYHGLTHMATMQPNAKFRPTQSGTGCSYSTSAIKIV	-
7	5.1e <sup>-114</sup>	21	HAFQWPNPRSTETVPETMQV	-

In addition, to determine other conserved regions of proteins, we aligned all 22 PL sequences by ClustalW (Figure 2). Multiple alignment of PL sequences revealed conserved regions containing cysteine-rich motifs (Motifs 1, 4, and 6) and three conserved amino acid residues with charged side chains, two aspartate residues (position 53, between Motif 7 and Motif 1; position 256, Motif 3), and one lysine residue (position 167, Motif 2), which are potentially involved in catalysis (Shevchik et al., 1997; Popeijus et al., 2000).

### Phylogenetic analysis

Phylogenetic analysis among 22 PL sequences, revealed the presence of two main groups (Figure 3). Group 1 contained 13 PL sequences of 11 species and Group 2 contained nine





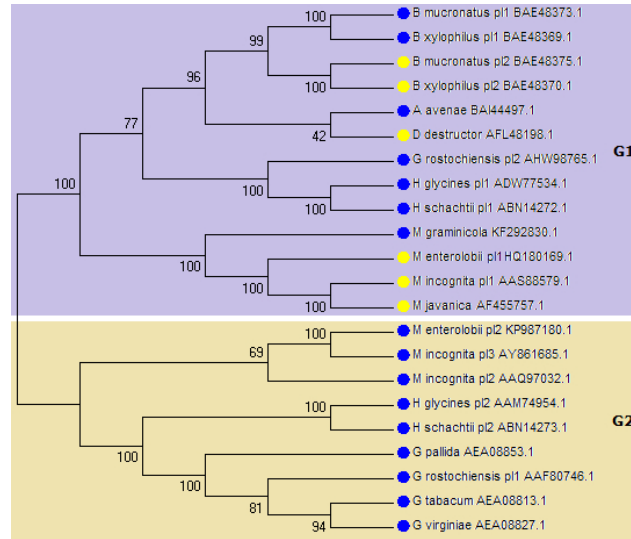
**Figure 2.** Sequence alignment of 22 nematode pectate lyase precursors. Sequences were aligned by ClustalW, and identical and similar residues are displayed in the same color. Motifs discovered by the MEME SUITE tool are indicated above the sequences. The black arrow indicates the conserved proline after N-terminal signal peptides, brown arrows indicate conserved cysteine residues, red arrows indicates conserved aspartate residues, and the blue arrow indicates the conserved lysine residue.

sequences of eight species. While G1 contained hydrophilic PLs that were both acidic and alkaline, G2 consisted of less hydrophilic alkaline PLs. These results suggest that G1 contains two families of PLs with different physico-chemical properties, while G2 contains PLs of similar characteristics.

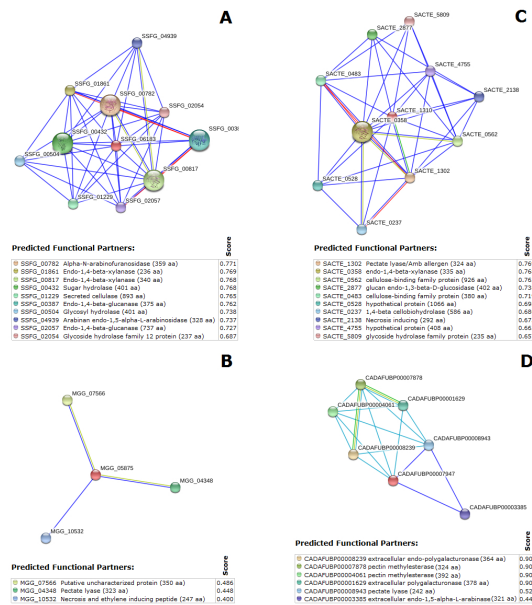
### Functional interaction network analysis

In order to predict protein interactions, the two *M. enterolobii* PLs were mapped to the STRING 9.1 tool as a sample of each identified group from the phylogenetic analysis. STRING is a database of known and predicted protein interactions. The interactions include direct (physical) and indirect (functional) associations. The most similar sequences to the query sequences were those from fungi and bacteria. *Magnaporthe oryzae* (MGG\_05875; PL, 240 aa) and *Streptomyces ghanaensis* (SSFG\_06183; secreted lyase 264 aa) sequences were used as models for *M. enterolobii* PL1. *Aspergillus fumigatus* (CADAFUBP00007947; PL, 254 aa) and *Streptomyces*

sp (SACTE\_1310; PL, 280 aa) were used as models for *M. enterolobii* PL2. The identified putative interaction partners for both bacterial and fungal models were all CWDE proteins (Figure 4).



**Figure 3.** Neighbor-joining phylogenetic tree of tylenchid nematode pectate lyases. The percentage of 2000 bootstrap replicates is given at each node. Based on the phylogenetic tree result, 22 PL protein sequences were divided into two groups with different colors. Yellow circles indicate acidic proteins and blue circles indicate alkaline proteins.



**Figure 4.** Putative interaction partners of *Meloidogyne enterolobii* pectate lyases (PLs). **A.** SSFG\_06183 and **B.** MGG\_05875 (bacterial and fungal models for *M. enterolobii* PL1). **C.** SACTE\_1310 and **D.** CADAUFUBP00007947 (bacterial and fungal models for *M. enterolobii* PL2). Stronger associations are represented by thicker lines.



## Tertiary structure prediction and molecular docking

Tertiary structure prediction of PLs was performed using the Phyre2 server. This server uses the alignment of hidden Markov models via HHsearch (Söding, 2005) to improve the accuracy of alignment and detection rate. It also incorporates the Poing tool (Jefferys et al., 2010), which is an *ab initio* folding simulation to model regions of the proteins with no detectable homology to known structures. Poing tool also combines multiple templates to improve model accuracy.

Verification of stereochemical quality of 22 models using Ramachandran plot analysis was performed by the Molprobit server (Table 4). Models showed that 79.45% (*M. enterolobii* PL1) to 90.27% (*B. xylophilus* PL2) of amino acid residues were in favored regions. ProSA-web (Protein Structure Analysis web) was used to recognition errors in the tertiary structure prediction of PLs. The Z-score was used to measure the energy, as it indicated overall quality of the model. Positive Z-score values show that the structure is not stabilized while zero and negative scores represent one of the ideal structures. All models showed negative Z-score values, ranging from -0.32 (*B. mucronatus* PL1) to -5.55 (*M. incognita* PL1).

**Table 4.** Validation of 22 pectate lyase (PL) models generated by the Phyre2 server.

Origin species	Z-score	Favored RP	Outliers
<i>A. avenae</i> pl_BAI44497.1	-4.28	89.73%	3.57%
<i>B. mucronatus</i> pl_BAE48373.1	-3.92	86.90%	1.75%
<i>B. mucronatus</i> pl_BAE48375.1	-0.32	88.71%	1.61%
<i>B. xylophilus</i> pl_BAE48369.1	-4.96	87.77%	1.75%
<i>B. xylophilus</i> pl_BAE48370.1	-3.48	90.27%	2.65%
<i>D. destructor</i> pl_AFL48198.1	-4.79	83.88%	5.37%
<i>G. pallida</i> pl1_AEA08853.1	-2.89	83.19%	7.56%
<i>G. rostochiensis</i> pl1_AAF80746.1	-4.11	86.97%	3.78%
<i>G. rostochiensis</i> pl2_AHW98765.1	-3.41	82.97%	5.68%
<i>G. tabacum</i> pl1_AEA08813.1	-3.92	86.55%	3.36%
<i>G. virginiae</i> pl1_AEA08827.1	-4.36	86.55%	4.62%
<i>H. glycines</i> pl_ADW77534.1	-4.19	86.96%	3.04%
<i>H. glycines</i> pl2_AAM74954.1	-3.72	89.17%	2.08%
<i>H. schachtii</i> pl_ABN14272.1	-4.33 (-4.08)	90.04% (98.34%)	1.24% (0)
<i>H. schachtii</i> pl_ABN14273.1	-3.13	86.03%	3.06%
<i>M. enterolobii</i> pl_ADN87334.1	-5.34	87.40%	2.44%
<i>M. enterolobii</i> pl2_ALB38961.1	-1.29	79.45%	6.72%
<i>M. graminicola</i> pl_AID59201.1	-4.5	86.72%	4.56%
<i>M. incognita</i> pl1_AAS88579.1	-5.55 (-5.53)	84.21% (97.17%)	7.29% (0)
<i>M. incognita</i> pl2_AAQ97032.1	-3.60 (-0.61)	79.82% (96.46%)	7.87% (0)
<i>M. incognita</i> pl3_AY861685.1	-1.22 (-0.01)	80.63% (96.84%)	5.93% (0)
<i>M. javanica</i> pl_AAL66022.1	-5.2	88.26%	4.45%

Values in parentheses are values after model refinement.

After validation, two models for each group identified in the phylogenetic analysis (Group 1, *Heterodera schachtii* PL1, *M. incognita* PL1; Group 2, *M. incognita* PL2 and *M. incognita* PL3) were selected for model refinement and molecular docking analysis. Due to the presence of Ramachandran outliers, model refinement was carried out with the KiNG software (Chen et al., 2009).

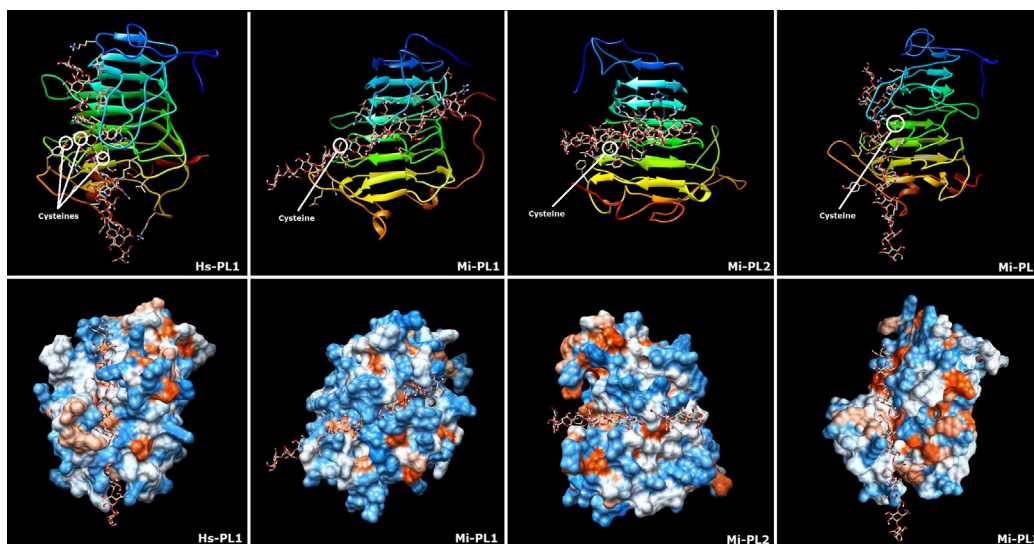
The objective of computational docking is to determine how two molecules will interact. Molecular docking is often employed to determine how receptors and ligands interact to form a binding pocket. *H. schachtii* PL1, *M. incognita* PL1, *M. incognita* PL2, and *M. incognita* PL3 were used to predict interaction sites with the pectic acid structure (Walkinshaw and Arnott, 1981; <http://glyco3d.cermav.cnrs.fr/rsrsrc/polysaccharide/pdb/pectic->

acid\_expanded.pdb) by the Patchdock server. Patchdock algorithms are inspired by image segmentation and object recognition techniques, which are used in computer vision. Given two molecules, their surfaces are divided into patches according to the surface shape. The patches are then filtered so that only patches with hot-spot residues are retained. Once the patches are identified, they are superimposed using a shape-matching algorithm (Schneidman-Duhovny et al., 2005).

Interaction refinement was performed using the Firedock server. FireDock is an efficient method for refinement and re-scoring of rigid-body protein-protein docking solutions. Results from patch dock are refined by Firedock, generating results ranked on the basis of global energy (Table 5). Figure 5 shows the solutions with best results of global energy values by the Firedock analysis.

**Table 5.** Results of pectate lyase (PL) molecular docking analysis.

Origin species	Score	Global energy	Interface area ( $\text{\AA}^2$ )	Catalytic cysteines
<i>H. schachtii</i> PL1	9208	-20.24	1329.4	145, 147, 150
<i>M. incognita</i> PL1	8500	-34.79	1190	149
<i>M. incognita</i> PL2	8546	-18.35	1104.4	141
<i>M. incognita</i> PL3	8310	-16.06	1027.4	135



**Figure 5.** Tertiary structure prediction of four selected pectate lyase (PL) sequences, predicted interactions with pectic acid (above), and different molecular pockets and binding sites identified (below). Circles indicate conserved cysteines with probable catalytic functions. Hydrophobicity is represented as a color gradient, with blue being the most hydrophilic, to white, to orange red for the most hydrophobic.

## DISCUSSION

Following the first report on the non-symbiotic degradation of plant cell-wall pectin by an animal (Popeijus et al., 2000), several PLs have been described for phytopathogenic and non-phytopathogenic nematodes (De Boer et al., 2002; Doyle and Lambert, 2002; Kikuchi et

al., 2004; Huang et al., 2005; Kudla et al., 2007; Karim et al., 2009). Most plant pathogens have different and independently regulated PLs, which possess 27-80% similarity. These isozymes have similar catalytic mechanisms, but recognize different methylated and non-methylated oligogalacturonate units (Herron et al., 2000; Huang et al., 2005). Nematodes infect plant species from diverse families, and an array of PLs with different chemical and enzymatic properties may be beneficial for the infection of a broad range of host plants (Huang et al., 2005).

Functional analysis revealed PLs that had very distinct physico-chemical properties but were similar at the amino acid level. Phylogenetic analysis showed that Group 1 contained both acidic and alkaline PLs. Similarly, PLs from bacteria are known to have distinct pI, different cell-wall maceration activities, and substrate specificities. PelA from *Dickeya dadantii* (Sin.: *Erwinia chrysanthemi*) has an acidic pI and poor maceration activity, while PelD and PelE are alkaline and have high maceration efficiency (Tardy et al., 1997). The pI is the pH at which a protein carries no net charge. At a pH below their pI, proteins carry a positive net charge and above their pI they carry a negative net charge. The pI of a protein is important because it represents the point at which the protein is least soluble, and therefore unstable (Shaw et al., 2001). Thus, the low pI value of PelA may contribute to its poor cell-wall macerating efficiency, due the acidic pH of plant tissues (Favey et al., 1992; Tardy et al., 1997). Despite its poor maceration ability, PelA is essential for the full pathogenicity of *D. dadantii* (Boccaro et al., 1988).

The fact that phytopathogenic organisms are often encountered in soil raises the possibility that the acidic PLs may have physiological functions that are important for survival outside the plant tissue (Favey et al., 1992).

MEME revealed seven conserved motifs in the 22 evaluated sequences. The three cysteine rich motifs (Motifs 1, 4, and 6) are potentially involved in catalytic activities. Motif 4 was not present in *M. incognita* PL3 due the absence of two cysteines that were present in all other sequences. Motif 6 was common in Group 2 PLs, but was not present in Group 1 PLs due to the absence of a 7-aa sequence before the first conserved cysteine. These differences suggest that Group 1 and Group 2 PLs may have distinct enzymatic activities.

STRING 9.1 revealed 13 and 16 putative interaction partners for Group 1 and Group 2 PLs, respectively. All identified partners were CWDEs from bacteria and fungi. Due to the structural complexity of the plant cell wall, plant-pathogenic organisms are expected to have a great diversity of CWDEs to facilitate the penetration of host plant cells.

Knowledge on the 3-D structure of proteins is of great importance to understand their function. Tertiary structures can be predicted from the amino acid sequences by using different techniques and methods, including homology modeling. This method compares the structure of protein sequences with those of known protein sequences to predict the protein structure. Since no crystal structures of PLs from nematodes are available, theoretic models were generated based on homology modeling using bacterial and fungal PL structures. After validation and refinement, the four models selected were used to predict possible interaction sites with the pectic acid structure. *In silico* docking analysis confirms the interaction of the modeled PL structures with pectic acid polymers (Table 5). The analysis also shows that *H. schachtii* PL1, *M. incognita* PL1, and *M. incognita* PL2 models have potential interactions between cysteines located in Motif 4 and the pectic acid ligand. *M. incognita* PL3 interacts at an alternative site, with a single cysteine located in Motif 2 interacting with the ligand molecule (Table 5 and Figure 5).

The present results confirm variations in the 3-D structure and molecular interactions

of analyzed PLs. Prediction of the 3-D structure of phytonematode PLs would provide valuable insights into the molecular basis of these proteins' functions. Since PLs are essential for pathogenicity, they may serve as novel targets to combat nematode infections.

### Conflicts of interest

The authors declare no conflict of interest.

### ACKNOWLEDGMENTS

Research supported by CNPq, CAPES, and UFRPE resources. R.M. Moraes Filho was supported by a Post-Doctoral Fellowship PNPd-CAPES.

### REFERENCES

- Bailey TL, Boden M, Buske FA, Frith M, et al. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37: W202-8. <http://dx.doi.org/10.1093/nar/gkp335>
- Bird DM (2004). Signaling between nematodes and plants. *Curr. Opin. Plant Biol.* 7: 372-376. <http://dx.doi.org/10.1016/j.pbi.2004.05.005>
- Boccardo M, Diolez A, Rouve M and Kotoujansky A (1988). The role of individual pectate lyases of *Erwinia chrysanthemi* strain 3937 in pathogenicity on Saintpaulia plants. *Physiol. Mol. Plant Pathol.* 33: 95-104. [http://dx.doi.org/10.1016/0885-5765\(88\)90046-X](http://dx.doi.org/10.1016/0885-5765(88)90046-X)
- Chen VB, Davis IW and Richardson DC (2009). KING (Kinemage, Next Generation): a versatile interactive molecular and scientific visualization program. *Protein Sci.* 18: 2403-2409. <http://dx.doi.org/10.1002/pro.250>
- Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, et al. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 66: 12-21. <http://dx.doi.org/10.1107/S0907444909042073>
- Davis EL, Hussey RS, Baum TJ, Bakker J, et al. (2000). Nematode parasitism genes. *Annu. Rev. Phytopathol.* 38: 365-396. <http://dx.doi.org/10.1146/annurev.phyto.38.1.365>
- Davis EL, Hussey RS, Mitchum MG and Baum TJ (2008). Parasitism proteins in nematode-plant interactions. *Curr. Opin. Plant Biol.* 11: 360-366. <http://dx.doi.org/10.1016/j.pbi.2008.04.003>
- De Boer JM, Davis EL, Hussey RS, Popeijus H, et al. (2002). Cloning of a putative pectate lyase gene expressed in the subventral esophageal glands of *Heterodera glycines*. *J. Nematol.* 34: 9-11.
- Doyle EA and Lambert KN (2002). Cloning and characterization of an esophageal-gland-specific pectate lyase from the root-knot nematode *Meloidogyne javanica*. *Mol. Plant Microbe Interact.* 15: 549-556. <http://dx.doi.org/10.1094/MPMI.2002.15.6.549>
- Favey SC, Bourson Y, Bertheau A, Kotoujansky, et al. (1992). Purification of the acidic pectate lyase of *Erwinia chrysanthemi* 3937 and sequence analysis of the corresponding gene. *J. Gen. Microbiol.* 138: 499-508. <http://dx.doi.org/10.1099/00221287-138-3-499>
- Feng BZ, Li PQ, Fu L and Yu XM (2015). Exploring laccase genes from plant pathogen genomes: a bioinformatic approach. *Genet. Mol. Res.* 14: 14019-14036. <http://dx.doi.org/10.4238/2015.October.29.21>
- Franceschini A, Szklarczyk D, Frankild S, Kuhn M, et al. (2013). STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 41: D808-D815. <http://dx.doi.org/10.1093/nar/gks1094>
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, et al. (2005). Protein identification and analysis tools on the EXPASY server. In: *The proteomics protocols handbook* (Walker JM, ed.). Humana Press, New York, 571-607.
- Han Y, Zheng QS, Wei YP, Chen J, et al. (2015). *In silico* identification and analysis of phytoene synthase genes in plants. *Genet. Mol. Res.* 14: 9412-9422. <http://dx.doi.org/10.4238/2015.August.14.5>
- Herron SR, Benen JAE, Scavetta RD, Visser J, et al. (2000). Structure and function of pectic enzymes: virulence factors of plant pathogens. *Proc. Natl. Acad. Sci. USA* 97: 8762-8769. <http://dx.doi.org/10.1073/pnas.97.16.8762>
- Huang G, Dong R, Allen R, Davis EL, et al. (2005). Developmental expression and molecular analysis of two *Meloidogyne incognita* pectate lyase genes. *Int. J. Parasitol.* 35: 685-692. <http://dx.doi.org/10.1016/j.ijpara.2005.01.006>
- Hussey R, Davis E and Baum T (2002). Secrets in secretions: genes that control nematode parasitism of plants. *Braz. J. Genetics and Molecular Research* 15 (3): gmr.15038402

- Plant Physiol.* 14: 183-194. <http://dx.doi.org/10.1590/S1677-04202002000300002>
- Jefferys BR, Kelley LA and Sternberg MJ (2010). Protein folding requires crowd control in a simulated cell. *J. Mol. Biol.* 397: 1329-1338. <http://dx.doi.org/10.1016/j.jmb.2010.01.074>
- Karim N, Jones JT, Okada H and Kikuchi T (2009). Analysis of expressed sequence tags and identification of genes encoding cell-wall-degrading enzymes from the fungivorous nematode *Aphelenchus avenae*. *BMC Genomics* 10: 525. <http://dx.doi.org/10.1186/1471-2164-10-525>
- Kelley LA, Mezulis S, Yates CM, Wass MN, et al. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10: 845-858. <http://dx.doi.org/10.1038/nprot.2015.053>
- Kikuchi T, Jones JT, Aikawa T, Kosaka H, et al. (2004). A family of glycosyl hydrolase family 45 cellulases from the pine wood nematode *Bursaphelenchus xylophilus*. *FEBS Lett.* 572: 201-205. <http://dx.doi.org/10.1016/j.febslet.2004.07.039>
- Kudla U, Milac AL, Qin L, Overmars H, et al. (2007). Structural and functional characterization of a novel, host penetration-related pectate lyase from the potato cyst nematode *Globodera rostochiensis*. *Mol. Plant Pathol.* 8: 293-305. <http://dx.doi.org/10.1111/j.1364-3703.2007.00394.x>
- Perry RN and Moens N (2011). Introduction to plant-parasitic nematodes: modes of parasitism. In: Genomics and molecular genetics of plant-nematode interactions (Jones J, Gheysen G and Fenoll C, eds.). Springer Science + Business Media, Dordrecht, 3-20.
- Pettersen EF, Goddard TD, Huang CC, Couch GS, et al. (2004). UCSF Chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25: 1605-1612. <http://dx.doi.org/10.1002/jcc.20084>
- Popeijus H, Overmars H, Jones J, Blok V, et al. (2000). Degradation of plant cell walls by a nematode. *Nature* 406: 36-37. <http://dx.doi.org/10.1038/35017641>
- Rai KM, Balasubramanian VK, Welker CM, Pang M, et al. (2015). Genome wide comprehensive analysis and web resource development on cell wall degrading enzymes from phyto-parasitic nematodes. *BMC Plant Biol.* 15: 187. <http://dx.doi.org/10.1186/s12870-015-0576-4>
- Schneidman-Duhovny D, Inbar Y, Nussinov R and Wolfson HJ (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucl. Acids Res.* 33: W363-W367.
- Scholl EH, Thorne JL, McCarter JP and Bird DM (2003). Horizontally transferred genes in plant-parasitic nematodes: a high-throughput genomic approach. *Genome Biol.* 4: 39.
- Shaw KL, Grimsley GR, Yakovlev GI, Makarov AA, et al. (2001). The effect of net charge on the solubility, activity, and stability of ribonuclease Sa. *Protein Sci.* 10: 1206-1215. <http://dx.doi.org/10.1110/ps.440101>
- Shevchik V, Robert-Baudouy J and Hugouvieux-Cotte-Pattat N. (1997). Pectate lyase PelI of *Erwinia chrysanthemi* 3937 belongs to a new family. *J. Bacteriol.* 179: 7321-7330.
- Smant G, Stokkermans JP, Yan Y, De Boer JM, et al. (1998). Endogenous cellulases in animals: isolation of beta-1, 4-endoglucanase genes from two species of plant-parasitic cyst nematodes. *Proc. Nat. Acad. Sci. U. S. A.* 95: 4906-4911.
- Södning J (2005). Protein homology detection by HMM-HMM comparison. *Bioinformatics* 21: 951-960. <http://dx.doi.org/10.1093/bioinformatics/bti125>
- Tamura K, Peterson D, Peterson N, Stecher G, et al. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731-2739. <http://dx.doi.org/10.1093/molbev/msr121>
- Tardy F, Nasser W, Robert-Baudouy J and Hugouvieux-Cotte-Pattat N (1997). Comparative analysis of the five major *Erwinia chrysanthemi* pectate lyases: enzyme characteristics and potential inhibitors. *J. Bacteriol.* 179: 2503-2511.
- Tsirigos KD, Peters C, Shu N, Käll L, et al. (2015). The TOPCONS web server for combined membrane protein topology and signal peptide prediction. *Nucleic Acids Res.* 43: 401-407. <http://dx.doi.org/10.1093/nar/gkv485>
- Vatanssever R, Filiz E and Ozyigit II (2015). *In silico* identification and comparative analysis of molybdenum (Mo) transporter genes in plants. *Braz. J. Bot.* 39: 1-13.
- Walkinshaw MD and Arnott S (1981). Conformation and interactions of pectins. I. X-ray diffraction analyses of sodium pectate in neutral and acidified forms. *J. Mol. Biol.* 153: 1055-1073. [http://dx.doi.org/10.1016/0022-2836\(81\)90467-8](http://dx.doi.org/10.1016/0022-2836(81)90467-8)
- Wiederstein M and Sippl MJ (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 35: W407-W410. <http://dx.doi.org/10.1093/nar/gkm290>
- Williamson VM and Gleason CA (2003). Plant-nematode interactions. *Curr. Opin. Plant Biol.* 6: 327-333. [http://dx.doi.org/10.1016/S1369-5266\(03\)00059-1](http://dx.doi.org/10.1016/S1369-5266(03)00059-1)
- Yu CS, Cheng CW, Su WC, Chang KC, et al. (2014). CELLO2GO: a web server for protein subCELLular LOcalization prediction with functional gene ontology annotation. *PLoS One* 9: e99368. <http://dx.doi.org/10.1371/journal.pone.0099368>