Available online at www.sciencedirect.com
SCIENCE DIRECT®
Experimental Parasitology 110 (2005) 123–133
Experimental Parasitology
www.elsevier.com/locate/yexpr

Isolation and characterization of a cDNA encoding a serine proteinase from the root-knot nematode *Meloidogyne incognita*

Rodrigo da Rocha Fragoso [a,b], João Aguiar Nogueira Batista [a], Osmundo Brilhante Oliveira Neto [a,b], Maria Fátima Grossi de Sá [a,*]

[a] *EMBRAPA, Recursos Genéticos e Biotecnologia, Brasília, DF, Brazil*
[b] *Departamento de Biologia Celular, Universidade de Brasília, Brasília, DF, Brazil*

Received 23 August 2004; received in revised form 14 February 2005; accepted 16 February 2005
Available online 23 March 2005

Abstract

This report describes the first serine proteinase gene isolated from the sedentary nematode *Meloidogyne incognita*. Using degenerate primers, a 1372 bp cDNA encoding a chymotrypsin-like serine proteinase (*Mi-ser1*) was amplified from total RNA of adult females by RT-PCR and 5′ and 3′ rapid amplification of cDNA ends. The deduced amino acid sequence of *Mi-ser1* encoded a putative signal peptide and a prodomain of 22 and 33 amino acids, respectively, and a mature proteinase of 341 amino acids with a predicted molecular mass of 37,680 Da. Sequence identity with the top serine proteinases matches from the databases ranged from 23 to 27%, including sequences from insects, mammals, and other nematodes. Southern blot analysis suggested that *Mi-ser1* is encoded by a single or few gene copies. The pattern of developmental expression analyzed by Northern blot and RT-PCR indicated that *Mi-ser1* was transcribed mainly in females. The domain architecture composed of a single chymotrypsin-like catalytic domain and the detection of a putative signal peptide suggested a digestive role for Mi-ser1.

Index Descriptors and Abbreviations: RKN, root-knot nematode; RACE, rapid amplification of cDNA ends; EST, expressed sequence tag; RT-PCR, reverse transcription and polymerase chain reaction

Keywords: cDNA cloning; *Meloidogyne*; Nematode; Serine proteinase

1. Introduction

Plant parasitic nematodes are economically important pests for several crops worldwide. Three endoparasitic sedentary genera—*Globodera*, *Heterodera* (cyst nematodes), and *Meloidogyne* (root-knot nematodes, RKN)—cause enormous economic losses to world agriculture (Sasser, 1980) estimated at US $100 billion annually (Lilley et al., 1999). Among the species involved, *Meloidogyne incognita*, the southern root-knot nematode, is considered to be the most important because it is the most widely distributed phytonematode species (Sasser et al., 1983; Trudgill, 1995) and has an extensive host range (Ehwaeti et al., 1999; Jepson, 1987). As a result, this species is probably the major pathogen responsible for damage to plants worldwide (Trudgill and Blok, 2000).

The current agricultural practices of crop rotation and chemical applications are usually expensive, inefficient, and environmentally hazardous (Shomaker and Been, 1999). Indeed, the use of highly toxic nematicides may result in environmental contamination or human intoxication (Jeyaratnam, 1990). A better solution has been found in plant-resistant genotypes, and several genes have been included in genetic breeding programs of perennial and annual plants in attempts to improve

* Corresponding author.
E-mail address: fatimasa@cenargen.embrapa.br (M.F. Grossi de Sá).

0014-4894/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.exppara.2005.02.010

the control of these nematodes by environmentally safe practices (Roberts, 1992). However, the breeding of resistant plants by classic genetics is limited by the reduced number of naturally resistant sources and sexual incompatibility among different species. Thus, plant transformation through genetic engineering is a promising alternative for the development of nematode-resistant plants (Atkinson et al., 2001).

Proteinases are important targets for intervention because of their fundamental roles in many metabolic processes and in host–parasite interactions (Beynon and Bond, 1994). Proteinases of the four major classes are present in parasitic helminths (Sakanari, 1990) and several roles have been suggested for these enzymes in the mechanisms of parasitism, including tissue invasion, evasion of the host immune system, and nourishment of the parasite (Coombs and Mottram, 1997). Anti-feeding strategies in natural host–parasite interactions involve the production of enzymatic inhibitors, such as seen in plants that are resistant or tolerant to insects or pathogens (Carlini and Grossi-de-Sa, 2002). The accumulation of serine proteinase inhibitors detected in subterranean organs of solanaceous plants could contribute to the plant defense against RKN (Trudgill and Blok, 2000).

Serine and cysteine proteinases predominate in plant nematode intestines (Lilley et al., 1996) whereas aspartic, cysteine, and metallo-proteinases predominate in animal nematode intestines (Chappell and Dresden, 1986; Longbottom et al., 1997; Redmond et al., 1997). All of the three sedentary endoparasitic genera show marked proteolytic activity. A protein homogenate of females of the potato cyst nematode *Globodera pallida* contains a major serine and a minor cysteine proteinase (Koritsas and Atkinson, 1994). The different species of RKN vary in proteinase composition and affinity, as shown by proteolytic assays using crude extracts from *Meloidogyne hapla*, *Meloidogyne javanica* and *M. incognita* (Michaud et al., 1996). Studies with *Heterodera glycines* resulted in the isolation of genes encoding three serine proteinases, *hgsp-I*, *hgsp-II*, and *hgsp-III* (Lilley et al., 1997) and two cysteine proteinases, *hgcp-I* and *hgcp-II* (Urwin et al., 1997a).

The induction of resistance to nematodes in host plants through the expression of proteinase inhibitors has been described (Urwin et al., 1997b, 1998). Thus, transgenic *Arabidopsis thaliana* expressing a mutant of a rice cysteine proteinase inhibitor (Oc-I ΔD86) were resistant to *Heterodera schachtii* and *M. incognita* (Urwin et al., 1997b). Major effects in the parasites included reduced female size and fecundity as a result of undernourishment. Enhanced resistance was obtained by using a dual gene construct containing a serine proteinase inhibitor, CpTI (cowpea trypsin inhibitor from *Vigna unguiculata*), and Oc-I ΔD86 (Urwin et al., 1998).

As part of an effort to develop nematode-resistant transgenic plants based on an anti-feeding strategy, we have isolated the genes encoding proteinases of *M. incognita*. A cDNA encoding a serine proteinase from *M. incognita* females was isolated using RT-PCR and RACE and named *Mi-ser1*. Disruption of the proteinase activity encoded by *Mi-ser1* may represent a target for *M. incognita* controlling.

2. Material and methods

2.1. Growth and collection of M. incognita

Nematodes were propagated on tomato (*Lycopersicon esculentum*) roots and harvested by standard protocols (Hussey and Barker, 1973) at three life stages: eggs, second stage juveniles (J2), and females. The roots were triturated in 0.5%(v/v) sodium hypochlorite and the egg-containing fraction was separated using 400-mesh sieves and suspended in kaolin prior to centrifugation at 2500*g* for 10 min. The pellet was resuspended in 50%(w/v) sucrose, centrifuged at 2500*g* for 1 min, and the eggs were collected from the supernatant on a 500-mesh sieve. Juveniles were obtained from eggs hatched in plastic trays with distilled water and were concentrated by centrifugation at 2500*g* for 30 min. Females were obtained from roots incubated for 12 h in 40%(v/v) pectinase and then triturated. The 100-mesh fraction was precipitated with kaolin and further purified as described for the eggs, except that 40%(w/v) sucrose was used.

2.2. Reverse transcription-polymerase chain reaction (RT-PCR) amplifications

Reverse transcription of total RNA (2 μg) from eggs, J2, and females of *M. incognita* was done with an oligo(d(T))-anchor primer and AMV-RT (Boehringer Mannheim), according to the manufacturer's instructions. PCR amplifications were done using the degenerate primers Ser1 (5′ACTGCTGCHCAYTG3′) and Ser2inv (5′GGRCCACCAGAGTCRCC3′), based on conserved motifs present beside the histidine and serine catalytic residues of serine proteinases. The PCR system (final volume, 25 μl) contained 2.5 U *Taq* DNA polymerase (Invitrogen), 2.0 mM $MgCl_2$, 200 μM dNTPs, and 400 nM of each primer. The PCR program consisted of an initial denaturation at 94 °C for 1 min, 30 cycles of amplification at 94 °C for 20 s, 42 °C for 20 s, and 72 °C for 30 s, followed by a final elongation step at 72 °C for 2 min. An aliquot of the PCR product was then used for a second round of amplification using the same conditions.

2.3. 5′ and 3′ rapid amplification of cDNA ends

The 5′ end of the cDNA fragment derived from the RT-PCR was amplified by 5′ rapid amplification of

cDNA ends (RACE) using a 5′/3′ RACE kit (Boehringer Mannheim) according to the manufacturer's instructions. The antisense primer Mi-sera (5′GTGTGCAGT TGGCCGCT3′) was used for reverse transcription and the antisense primers Mi-serb (5′AACAGCAGGCGG TCGGTC3′) and Mi-serc (5′GCGACGCATGTCCA GTAC3′) were used for the first and second round of a nested PCR, respectively. The first-round PCR conditions were a denaturation step at 94 °C for 1.5 min, 30 cycles of 94 °C for 45 s, 42 °C for 45 s, and 72 °C for 1.5 min, and a final elongation step at 72 °C for 5 min. The same conditions were used in the second PCR round, except that the annealing temperature was 52 °C. A second 5′ RACE was done using Mi-serc for reverse transcription and two additional antisense primers, Mi5serd (5′GGTCCCACCGTACGCCC3′) and Mi5sere (5′CTGCTTAAGACAGCTGG3′), for the nested PCR amplification. The reaction conditions were the same as described above, except that the annealing temperature was 45 °C in the first PCR round and 55 °C in the second round. The complete cDNA sequence was obtained by 3′ RACE using the nested sense primers derived from the 5′ RACE sequence, Mi3serf (5′CAGC GGCCAACTGCACAC3′) and Mi3serg (5′GGGAA CACCATCGAATG3′). PCR amplifications were done with a denaturation step at 94 °C for 1.5 min, 30 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1.5 min, and a final elongation at 72 °C for 5 min.

2.4. DNA sequencing and sequence analysis

The amplified cDNA fragments were cloned into pGEM-T Easy (Promega) and sequenced in an automated ABI DNA sequencer. Computer analyses of the sequences were done using the GCG software package (Genetics Computer Group, University of Wisconsin). Databank comparisons were done using the BLASTx software (Altschul et al., 1990) from the NCBI databank (http://www.ncbi.nlm.nih.gov). The Conserved Domain Database search (CDD-Search) from the NCBI site was used to compare motif identity and similarity with known conserved domains (Marchler-Bauer et al., 2003). Identification of the signal peptide sequence, including cleavage-site localization, was done using the SignalP software (http://www.cbs.dtu.dk/services/SignalP/) of the Prediction Server of the Center for Biological Sequence Analysis, BioCentrum-DTUTechnical, University of Denmark (Nielsen et al., 1997). Sequence alignments were done using CLUSTAL W software (Thompson et al., 1994) and were edited with the BOXSHADE software (http://www.ch.embnet.org/software/BOX_form.html); the dendrograms were edited with the TreeView software. The molecular mass and predicted p*I* of the deduced protein were determined by the Protein Machine software available at the Expasy site (http://us.expasy.org/tools/).

2.5. Southern blotting analysis

Genomic DNA from *M. incognita* eggs was isolated according to Sambrook et al. (1989) and digested (8 μg per digestion) with *Eco*RI, *Hin*dIII, *Xba*I, and *Nsi*I. The DNA digests were separated on a 0.8% agarose gel and transferred to a Hybond N^+ nylon membrane (Amersham Pharmacia Biotech, UK) using standard procedures (Sambrook et al., 1989). The *Mi-ser1* cDNA was labeled with α-[^{32}P]dCTP to a high specific activity using the Megaprimer DNA-labeling kit (Amersham Pharmacia Biotech) and used to probe the blot. The membrane was washed to a stringency of 0.1× SSC with 0.2% SDS at 65 °C.

2.6. Analysis of expression by Northern blot and RT-PCR

Total RNA from *M. incognita* eggs, J2, and females was purified using an RNeasy RNA extraction kit (Qiagen) according to the manufacturer's protocol and stored at −80 °C. Gel electrophoresis (Fourney et al., 1998) of total RNA from each developmental stage (22 μg per lane) and the hybridizations were done according to standard procedures (Sambrook et al., 1989). The membrane was washed to a stringency of 0.1× SSC at 65 °C. The signal intensities were compared using optical densities determined by Zero-Dscan software (Scanalytics) from digital images obtained with a Duoscan T1200 (AGFA) transparency scanner.

RT-PCR was done using internal primers specific for actin or *Mi-ser1* and, as a template, a 1:20 dilution of reverse transcribed cDNA from each developmental stage. cDNAs were obtained as described for the RT-PCR amplifications. A control was done using the same set of primers used in the RT-PCR and genomic DNA as template.

2.7. Database searching and sequence retrieval for nematode serine proteinases

The nucleotide sequence of *Mi-ser1* was used as a query in BLAST searches of the *Caenorhabiditis elegans* genome at WormBase (http://www.wormbase.org) and of the plant-parasite nematode dbEST at the Nematode Center (http://www.nematode.net). A GenBank search at NCBI (http://www.ncbi.nlm.nih.gov) was also done using as a query a combination of words that defined nematode species, developmental stage, and gene annotation.

3. Results

3.1. Cloning of a cDNA encoding a putative serine proteinase from M. incognita females

Preliminary results using in vitro assays with fluorogenic substrates and specific inhibitors revealed high

serine proteinase activity in J2 larvae and females of *M. incognita* (data not shown). Together with other reports describing serine proteinases in plant parasitic nematodes (Koritsas and Atkinson, 1994; Lilley et al., 1996) and the isolation of serine proteinase genes from *H. glycines* (Lilley et al., 1997), these results prompted us to search for homologous serine proteinase genes in *M. incognita* by using a generic PCR approach and degenerate primers based on conserved nucleotide sequences present in the genes of these enzymes.

RT-PCR amplification of total RNA from *M. incognita* females using degenerate primers from regions of conserved amino acids surrounding the active site of the His57 and Ser195 residues (Beynon and Bond, 1994) resulted in a major band of approximately 550 bp. Sequence analysis of seven clones from this fragment revealed a single sequence, homologous to other serine proteinases. Through 5′ and 3′ RACE, fragments located upstream and downstream from the central fragment were amplified and a final 1372-bp cDNA, named *Mi-ser1*, was amplified (Fig. 1). A final RT-PCR was done with specific primers corresponding to the regions surrounding the start and stop codons and amplified a fragment of approximately 1200 bp, the sequence of which was identical to the *Mi-ser1* ORF assembled from the RT-PCR and RACE fragments. This result confirmed that *Mi-ser1* corresponded to a unique, contiguous sequence and was not a hybrid resulting from several distinct sequence combinations.

An initial 5′ RACE resulted in the amplification of a cDNA fragment with no start codon in-frame with the ORF detected in the RT-PCR product, indicating that the isolated cDNA was incomplete at its 5′ end. To obtain the 5′ end of the cDNA, a second 5′ RACE was done using new antisense primers located closer to the 5′ end of the amplified fragment, and resulted in an 88-bp extension beyond the first 5′ RACE. Analysis of the extended sequence revealed a 7-bp thymidine stretch (positions 82–88) immediately upstream from the 5′ end of the fragments from the first 5′ RACE. This result indicated that in the first attempt to amplify the 5′ end the complementary sequence of this segment had competed with the homopolymeric adenosine tail added at the 3′ end of the cDNA and was preferentially amplified.

Southern blotting was done to determine the gene copy number of *Mi-ser1* (Fig. 2). Samples were digested with *Eco*RI and *Hin*dIII, which do not cleave the cDNA, and *Xba*I and *Nsi*I, which cleave the cDNA once at positions 1014 and 630, respectively. The digestions with *Eco*RI, *Nsi*I, and *Xba*I are consistent with a single copy gene. The fainter bands in the digestions with *Eco*RI and *Nsi*I could correspond to hybridization with a homologue gene. However, digestion with *Hin*dIII yielded three bands, including a fragment smaller than 1 kb. One interpretation for this result is the presence of internal cuts for *Hin*dIII in introns. Assuming *Mi-ser1* is a single copy gene, the digestion with *Xba*I should produce two bands, while our hybridization resulted in a single band. However, our electrophoresis of the digested DNA did not resolve fragments smaller than 500 bp, and it is possible that a smaller fragment may have been lost from the gel. In summary, the overall simple hybridization

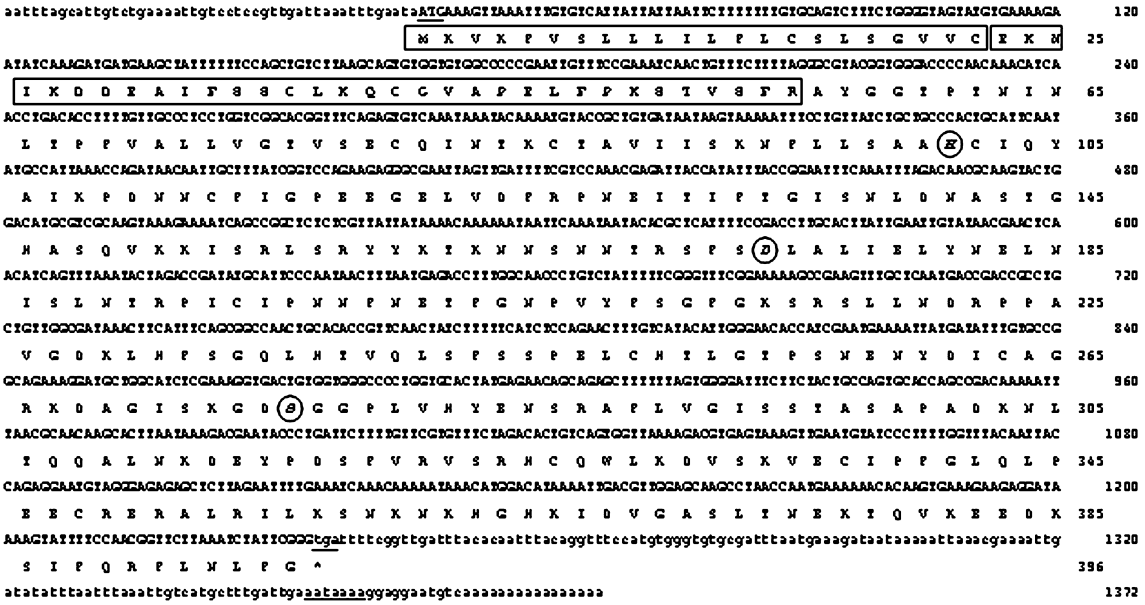

Fig. 1. Nucleotide sequence of the serine proteinase *Mi-ser1* cDNA isolated from *M. incognita* females and the predicted amino acid sequence. The coding sequence is shown in uppercase letters and the 5′ and 3′ untranslated sequences in lowercase letters. The start and stop codons are underlined and in bold. A putative polyadenylation signal at the 3′ UTR is underlined. The signal peptide (pre-region) is boxed and the prodomain (pro-region) is boxed and in bold. The residues of the catalytic triad are circled. This sequence has been submitted to the GenBank/EMBL databases under Accession No. AY714229.

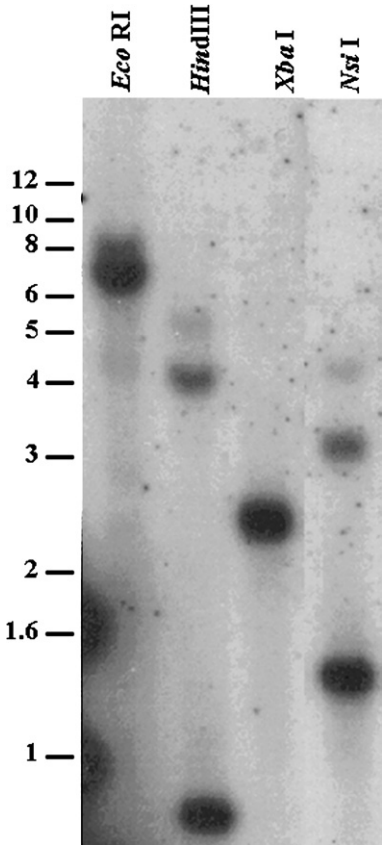

Fig. 2. Southern blot analysis of *Mi-ser1*. Genomic DNA (8 μg) from *M. incognita* was digested with *Eco*RI, *Hin*dIII, *Xba*I or *Nsi*I and probed with an α-^{32}P-labeled cDNA fragment encompassing the complete ORF of *Mi-ser1*. DNA size markers are indicated in kilobase.

pattern suggests *Mi-ser1* is encoded by a single or few gene copies in the haploid genome of *M. incognita*.

3.2. Characterization of the deduced amino acid sequence of Mi-ser1

The *Mi-ser1* cDNA had a single 1188-bp ORF encoding a putative 396-amino acid protein (Fig. 1). Comparison of the deduced Mi-ser1 protein sequence with other sequences from the databanks using BLASTp showed 23–27% identity to a multitude of serine proteinases, including chymotrypsins, trypsins, plasminogens, coagulation IX factors, enteropeptidases, enterokinases, elastases, kallikreins, and arginine esterase precursors, all from animal species (data not shown). Despite the large repertoire of physiological functions of these proteins, all of them share a closely related serine proteinase domain (Krem et al., 1999) and belong to the chymotrypsin-like clan of serine proteinases (Hedstrom, 2002). Domain analysis of Mi-ser1 using the CDD-Search (Conserved Domain Database) yielded an *e* value of 1e−27 relative to a trypsin-like serine proteinase (cd00190) and 2e−17 relative to trypsin (pfam00089). An analysis using the CDART (Conserved Domain Architecture Retrieval Tool) indicated that Mi-ser1 possesses a single chymotrypsin domain.

Sequence analysis of Mi-ser1 revealed a putative signal peptide of 22 residues, a prodomain of 33 residues, and a mature proteinase of 341 residues (Fig. 1), with a predicted molecular mass of 37,680 Da. Using SignalP software (Nielsen et al., 1997), the N-terminal of Mi-ser1 was found to share high similarity with several eukaryote signal peptides, and had a putative cleavage site at residues 22–23 (VVC–EK). The N-terminal residues of mature serine proteinases are highly conserved, with the R–IVGG motifs being the most commonly found in vertebrate and invertebrate trypsins (Muharsini et al., 2001; Valaitis et al., 1999). Comparison of the N-terminal of the mature Mi-ser1 proteinase with other serine proteinases indicated a zymogen activation site at residues 55–56 (R–AYGG).

Non-redundant nematode sequences of serine proteinases from different families were used for the multiple alignment (Fig. 3). The amino acid residues of the catalytic triad of serine proteinases are found in conserved motifs and Mi-ser1 catalytic triad residues were identified in these consensus regions (His46, Asp119, and Ser221). However, in contrast to most serine proteinases that normally have three conserved disulfide bridges (Jiang and Kanost, 2000; Mulenga et al., 2001), only two were identified in Mi-ser1 (Fig. 3). The disulfide bridge that was absent in Mi-ser1, corresponding to the third pair in the figure, has a key role in active site formation when the residues determining substrate specificity are placed together in the oxyanion hole (Yoshida et al., 1998).

Conserved residues of S1 site specificity (marked by losangs in Fig. 3), lining the substrate binding pocket that contribute to the trypsin specificity (Asp, Gly, and Gly) or to the chymotrypsin specificity (Ser, Gly, and Gly), were not identified in Mi-ser1 (Czapinska and Otlewski, 1999; Krem et al., 1999, Perona and Craik, 1997). The primary specificity residue in Mi-ser1 was Gly215 which indicates other specificity rather than trypsin or chymotrypsin (Bangyeekhun et al., 2001; Chiou et al., 1998; Kawamoto et al., 1999; Mazumdar-Leighton et al., 2000; Zhu et al., 2000). In addition, the Ala244 and Asp262 may tune the S1 site entrance which suggests a binding to small aliphatic residues at P1 site of the substrate, a characteristic of elastases (Hedstrom, 2002).

3.3. Expression of Mi-ser1 mRNA in the developmental stages of M. incognita

The cDNA fragment containing the *Mi-ser1* ORF was used to probe the developmental expression of the gene during the nematode life cycle. Northern blot analysis (Fig. 4A) revealed a single band of approximately 1.4 kb, in accordance with the size of the amplified *Mi-ser1* cDNA. Expression was detected mainly in eggs and

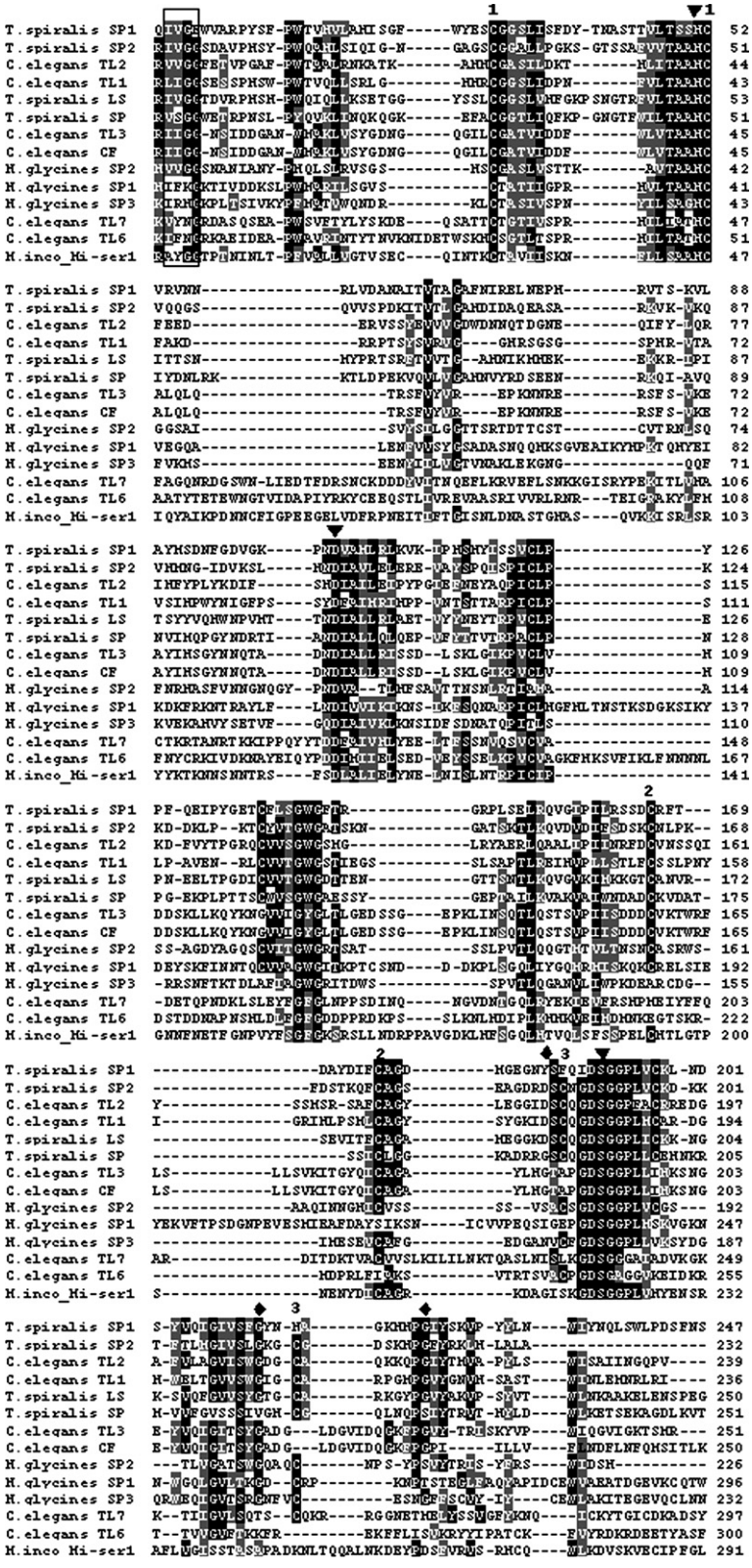

Fig. 3. Alignment of Mi-ser1 from *M. incognita* with other nematode serine proteinases using CLUSTAL W. The sequences are from *Trichinella spiralis* (SP— AAK16516; SP1— AAK31787; SP2— AAD09211; LS— AAK16520), *Caenorhabditis elegans* (TL1— AAA68746; TL2— AAB00662; TL3— AAB09110; TL6— AAB93344; TL7— AAB54144; CF— NP_500999), and *Heterodera glycines* (SP1—CAA74206; SP2—CAA74205; SP3—CAA74204). Residues are numbered beginning from the predicted N-terminal end of the mature proteins. The furthermost C-terminal residues of the sequences are not included in the alignment. The activation site is boxed and the predicted disulfide bonds are numbered. The catalytic triad (▼), substrate specificity residues (◆), and sequence alignment gaps (—) are also shown.

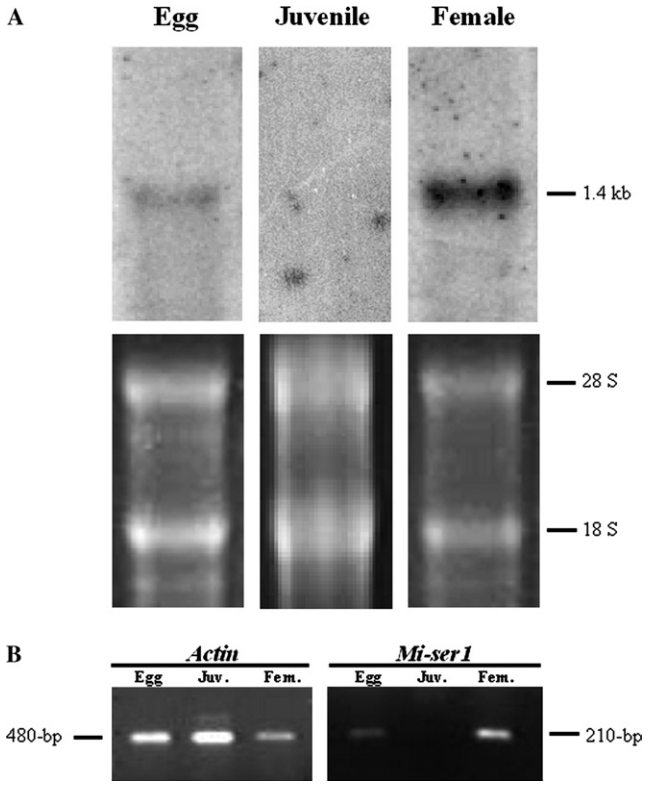

Fig. 4. Developmental expression of *Mi-ser1*. (A) Northern blot of total RNA (22 g) from eggs, J2 juveniles, and females of *M. incognita* hybridized with α-[^{32}P]dCTP-labeled *Mi-ser1* cDNA. The amount of RNA loaded per lane, normalized by ethidium bromide staining of the 28S and 18S rRNAs, is shown at the bottom. (B) RT-PCR amplification patterns obtained with specific primers for *Mi-ser1* and actin.

females. Normalization of the amount of RNA loaded for each sample using the intensity of the rRNAs, and comparison between females and eggs, showed a hybridization signal that was 6.9-fold higher in females than in eggs, indicating a higher level of expression in adult females. Similar results were obtained in the RT-PCR analysis (Fig. 4B). In the densitometric analysis, a very weak band was also detected in J2, with a signal intensity that was 26- and 3.8-fold lower than in females and eggs, respectively. However, we cannot exclude the possibility that this result corresponds to a contamination of the J2 sample with eggs. Amplification of genomic DNA with the same set of primers used in the RT-PCR resulted in the amplification of a higher molecular weight fragment (data not shown), confirming the specificity of the RT-PCR amplifications from the cDNA templates.

3.4. In silico analysis of serine proteinase ESTs from Meloidogyne spp

So far, the Parasitic Nematode Project (http://www.nematode.net/) has generated approximately 400,000 ESTs (Parkinson et al., 2003), of which nearly 73,000 are from RKN and nearly 20,000 are from *M. incognita*. We used this dataset to expand the characterization of *Mi-ser1*. Although few proteinase genes from plant parasitic nematodes have been isolated and characterized, functional genomic analyses of many nematodes species are currently underway and a large number of nematode ESTs are available in databanks. To obtain clues on the expression profile of nematode proteinases, we searched for serine proteinases in the RKN EST databanks. As a control, the same analysis was also done for actin ESTs and the results are shown in Table 1. Among the RKN, the number of ESTs available from eggs and J2 is much larger than from females, which represent only around 15% of the available sequences. Three ESTs encoding serine proteinases were identified in egg cDNA libraries from *M. incognita*, six in *M. hapla*, four in *M. arenaria*, and five in *M. chitwoodi*, giving a total of 18 sequences. Considering the total number of egg ESTs from all species, the mean was one serine proteinase EST per 2027 ESTs. No serine proteinase sequences were found in J2 or female libraries, however only a low number of ESTs from females were available for comparison.

Sequence analysis of the 18 RKN serine proteinase ESTs revealed at least nine different sequences and three gene clusters (Fig. 5). The largest cluster comprised 11 ESTs, including five from *M. chitwoodi* and two each from *M. incognita*, *M. hapla*, and *M. arenaria*. The identity between the sequences varied from 58 to 91%, with a mean of 74.5%. When compared to other serine proteinases, the best match was 41% identity over a 115 amino acid overlap with HGSP-III from *H. glycines*. Another cluster was represented by two ESTs, one from *M. hapla* and the other from *M. arenaria*, with 85.6% identity in its amino acid sequences. BLAST search results with the *M. arenaria* ESTs revealed 38% identity over a 91 amino acid overlap with *H. glycines* HGSP-I.

One EST from *M. arenaria* eggs (rm47d11.y1) encoded a chymotrypsin-like serine proteinase that showed 99% identity (*e* value of 6.6e−96) with *Mi-ser1* over a 148 amino acid overlap. This 447-nt EST was

Table 1
Number of ESTs found by the NCBI search tool in nematode developmental stages

	M. incognita			*M. hapla*			*M. arenaria*		*Meloidogyne* spp.		
	Egg	J2	Fem	Egg	J2	Fem	Egg	J2	Egg	J2	Fem
Total	7314	7556	4428	9783	9668	5001	3366	1652	36,486	24,932	10,849
Actin	24 (0.328)	62 (0.821)	4 (0.090)	56 (0.572)	757 (7.830)	7 (0.140)	9 (0.267)	4 (0.242)	114 (0.340)	852 (3.417)	11 (0.101)
Serine proteinase	3 (0.041)	—	—	6 (0.061)	—	—	4 (0.119)	—	18 (0.049)	—	—

Numbers in parentheses represent the percentage of each search result relative to the total number of ESTs.

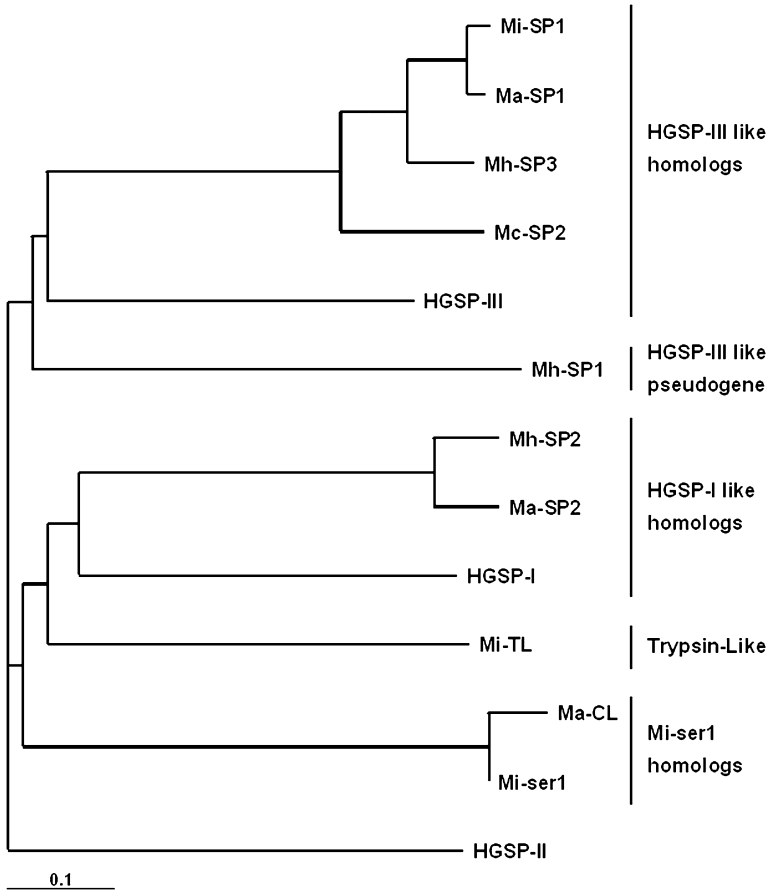

Fig. 5. Dendrogram of the alignment of the predicted protein sequence of *Mi-ser1* and selected ESTs from *Meloidogyne* spp., and the *H. glycines* serine proteinases. The alignment and the dendrogram were constructed using CLUSTAL W and edited with TreeView. The first two letters in each sequence name indicate the species of origin. The scale bar represents 0.1 amino acid substitutions per site.

identical to *Mi-ser1*, except for a single nucleotide that resulted in the substitution of an alanine by threonine. Another gene represented by a single EST from *M. incognita* showed 32% identity (over a 134 amino acid overlap) to a *Drosophila melanogaster* trypsin. The remaining three ESTs from *M. hapla* showed high identity to HGSP-III over a small stretch in the region surrounding the serine catalytic residue, but had little identity elsewhere; they also had several stop codons in the three frames and probably corresponded to a pseudogene.

In contrast to the high amino acid sequence identity within the clusters, the identity among sequences in different clusters was low (mean identity of 17%). The identity between the *M. incognita* ESTs Mi-SP1 and Mi-TL (Fig. 5) was 16% and, when compared to Mi-ser1, showed a mean identity of only 10.5%.

4. Discussion

In plants infected by nematodes, there is enhanced expression of defense genes such as trypsin inhibitors, peroxidases, chitinases, lipoxygenases, and extensins at the sites of nematode feeding (Gheysen and Fenoll, 2002). Increased expression of trypsin inhibitors was observed in *L. esculentum* (Williamson and Hussey, 1996) and *A. thaliana* (Vercauteren et al., 2001) infected with *M. incognita*, suggesting a critical role for serine proteases in plant–nematode interactions.

Despite reports showing that cysteine and serine proteinases are the major proteolytic enzymes in plant parasitic nematodes and the importance of proteinases in host–parasite interactions and in essential physiological processes of parasites (Koritsas and Atkinson, 1994; Lilley et al., 1996, 1997; Michaud et al., 1996), little is still known about the role of these endoproteolytic enzymes in the parasite life cycle. Since few proteinase genes have been cloned and characterized from plant parasitic nematodes (Atkinson et al., 2003), we chose to isolate and characterize serine proteinase cDNAs from *M. incognita* as an initial step to identifying potential targets for the development of nematode-resistant transgenic plants based on anti-feeding strategies.

The sequence reported here, *Mi-ser1*, is the first full-length serine proteinase gene cloned and characterized from RKNs. Although detection of a putative signal peptide indicated that the predicted protein sequence of

Mi-ser1 was complete, comparison of the 5′ end of the *Mi-ser1* cDNA sequence with the classic spliced leader and some variant forms described from *C. elegans* (Ross et al., 1995) and *M. javanica* (Koltai et al., 1997) did not reveal any significant similarity, indicating that the cloned cDNA could still be incomplete. A spliced leader is a common feature at the 5′ end of about 80% of *C. elegans* mRNAs (Nilsen, 1993), and the corresponding sequence has been used as a primer for the amplification of nematode proteinase genes (Neveu et al., 2003). Alternatively, the spliced leader of *M. incognita* could differ from those previously described and may represent a novel variant or simply be absent. However, even if the 5′ untranslated region of *Mi-ser1* were incomplete, the presence of two stop codons in-frame with the *Mi-ser1* ORF immediately upstream of the putative start codon indicated that translation did not begin further upstream.

In many insect pests, serine proteinases are encoded by multigene families. In the *D. melanogaster* genome, these enzymes constitute the second largest gene family, with 377 different sequences, whereas in the *C. elegans* genome serine proteinases are encoded by just 18 genes (Rubin et al., 2000). If a lower number of serine proteinase genes was a general feature of nematodes, this may make the inhibition of these enzymes easier in plant parasitic nematodes when compared with many insects that have a larger repertoire of these genes. Southern blot analysis indicated that *Mi-ser1* was encoded by a single or a few copies in the *M. incognita* genome. However, the EST databank searches revealed the existence of serine proteinase genes grouped in clusters among the RKNs. There is a low sequence identity between clusters, but high sequence identity within each cluster, suggesting that the serine proteinase diversity in RKNs appeared before species speciation. The high identity among the serine proteinase homologues in different species of RKN suggests that inhibitors against a specific serine proteinase will be effective against its homologues in a broad range of species within the genus.

The tentative classification and functional annotation of Mi-ser1 was based on the high-scoring database matches from the BLAST search and on the identification of key residues determinant of specificity. Serine proteases, which constitute almost one-third of all known proteolytic enzymes, are classified in 10 clans according to protein structure, fold and mechanics (Hedstrom, 2002). Clan A is the major one and groups 301 known serine proteases that share a common origin with chymotrypsin. This clan possesses 10 families of serine proteases, which retain the order His/Asp/Ser on the catalytic site (Barrett and Rawlings, 1995). The principal family is the S1, also referred as chymotrypsin-like proteases (Hedstrom, 2002). All serine proteases found in the BLAST search with Mi-ser1 belong to the family S1, presenting a chymotrypsin fold characterized by two perpendicular β-barrel domains, each formed by six antiparallel β-strands, and the C-terminal α-helix (Czapinska and Otlewski, 1999; Krem et al., 1999; Perona and Craik, 1997), where the active site cleft is located between these two barrels (Hedstrom, 2002). However, this approach may not be completely reliable, particularly since the top database matches gave low pairwise identities and important residues predictive of chymotrypsin or trypsin specificity were not found in Mi-ser1. Sequence analysis suggested that Mi-ser1 may have an elastase specificity, but the precise classification and functional annotation of Mi-ser1 will depend on more enzymatic and biochemical data.

Since the in situ localization of *Mi-ser1* was not investigated here, we have no information on the histological expression of *Mi-ser1*. The structure of the mature Mi-ser1 protein consisted of only a trypsin domain and a putative signal peptide, indicating that Mi-ser1 is secreted and could be involved in midgut digestion (Beynon and Bond, 1994). The expression of *Mi-ser1* was developmentally regulated and showed to be higher in females than in eggs and almost undetectable in J2. The Northern blot and RT-PCR results were supported by the in silico analysis of serine proteinase ESTs from *Meloidogyne* spp. in which ESTs encoding serine proteinases were found in eggs but not in J2. A precise comparison with females was not possible because the number of available female EST sequences was much lower.

These data agree with previous reports that showed high levels of proteinases expression in the feeding stages of parasitic nematodes (Atkinson et al., 2003). In phytoparasitic sedentary nematodes, females represent the parasitic stage that needs to digest plant proteins to obtain free amino acids and energy, while the J2 infective stage is a non-feeding stage (Wyss et al., 1992) that obtains energy by the consumption of its own lipid reserves (Reversat, 1981). This fact is supported by the analysis of J2 ESTs in which genes involved in the glyoxylate pathway were detected in high numbers (McCarter et al., 2003). The expression of serine proteinases in eggs is less clear. It is possible that the consumption of protein stores is necessary for larval growth. Alternatively, serine proteinases have a key role in several developmental processes that may be necessary for larval growth.

The results described here indicate that Mi-ser1 is a major serine proteinase in *M. incognita* females and suggest it may be important for female nourishment and, consequently, growth and reproduction. Further characterization of Mi-ser1 and the proteolytic system of *M. incognita*, combined with the use of a proteinase inhibitor-based approach should contribute to the development of transgenic plants resistant to parasitic nematodes, perhaps by using anti-feeding strategies to control the spatial and temporal expression of specific inhibitors against *M. incognita* proteinases. In this way,

further work is being directed towards the heterologous expression of Mi-ser1 and the use of the recombinant protein to select specific inhibitors through the screening of mutagenized bacteriophage libraries of serine proteinase inhibitors by phage display.

Acknowledgments

This work was supported by grants from the Brazilian government EMBRAPA, CNPq, and CAPES.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of Molecular Biology 215, 403–410.

Atkinson, H.J., Green, J., Cowgill, S., Levesley, A., 2001. The case for genetically modified crops with a poverty focus. Trends in Biotechnology 19, 91–96.

Atkinson, H.J., Urwin, P.E., McPherson, M.J., 2003. Engineering plants for nematode resistance. Annual Review of Phytopathology 41, 615–639.

Bangyeekhun, E., Cerenius, L., Soderhall, K., 2001. Molecular cloning and characterization of two serine proteinase genes from the crayfish plague fungus, *Aphanomyces astaci*. Journal of Invertebrate Pathology 77, 206–216.

Barrett, A.J., Rawlings, N.D., 1995. Families and clans of serine peptidases. Archives of Biochemistry and Biophysics 318, 247–250.

Beynon, R.J., Bond, J.S., 1994. Proteolytic Enzymes: a Practical Approach. IRL Press, Oxford University Press.

Carlini, C.R., Grossi-de-Sa, M.F., 2002. Plant toxic proteins with insecticidal properties. A review on their potentialities as bioinsecticides. Toxicon 40, 1515–1539.

Chappell, C.L., Dresden, M.H., 1986. *Schistosoma mansoni*: proteinase activity of 'hemoglobinase' from the digestive tract of adult worms. Experimental Parasitology 61, 160–167.

Chiou, S.J., Vanden Broeck, J., Janssen, I., Borovsky, D., Vandenbussche, F., Simonet, G., De Loof, A., 1998. Cloning of the cDNA encoding Scg-SPRP, an unusual Ser-protease-related protein from vitellogenic female desert locusts (*Schistocerca gregaria*). Insect Biochemistry and Molecular Biology 28, 801–808.

Coombs, G.H., Mottram, J.C., 1997. Parasite proteinases and amino acid metabolism: possibilities for chemotherapeutic exploitation. Parasitology 114, S61–S80.

Czapinska, H., Otlewski, J., 1999. Structural and energetic determinants of the S-1-site specificity in serine proteases. European Journal of Biochemistry 260, 571–595.

Ehwaeti, M.E., Fargette, M., Phillips, M.S., Trudgill, D.L., 1999. Host status differences and their relevance to damage by *Meloidogyne incognita*. Nematology 1, 421–432.

Fourney, R., Miyakoshi, J., Day, R.I., Paterson, M., 1998. Northern blotting: efficient RNA staining and transfer. Focus 10, 5–7.

Gheysen, G., Fenoll, C., 2002. Gene expression in nematode feeding sites. Annual Review of Phytopathology 40, 191–219.

Hedstrom, L., 2002. Serine protease mechanism and specificity. Chemical Reviews 102, 4501–4523.

Hussey, R.S., Barker, K.R., 1973. A comparison of methods of collecting inocula of *Meloidogyne* spp., including a new technique. Plant Disease Report 57, 1025–1028.

Jepson, B.S., 1987. Identification of Root-knot Nematodes (*Meloidogyne* Species). CAB Int., Wallingford, UK.

Jeyaratnam, J., 1990. Acute pesticide poisoning: a major global health problem. World Health Statistics Quarterly 43, 139–144.

Jiang, H., Kanost, M.R., 2000. The clip-domain family of serine proteinases in arthropods. Insect Biochemistry and Molecular Biology 30, 95–105.

Kawamoto, S., Mizuguchi, Y., Morimoto, K., Aki, T., Shigeta, S., Yasueda, H., Wada, T., Suzuki, O., Jyo, T., Ono, K., 1999. Cloning and expression of Derf 6, a serine protease allergen from the house dust mite, *Dermatophagoides farinae*. Biochimica et Biophysica Acta 1454, 201–207.

Koltai, H., Chejanovsky, N., Raccah, B., Spiegel, Y., 1997. The first isolated collagen gene of the root-knot nematode *Meloidogyne javanica* is developmentally regulated. Gene 196, 191–199.

Koritsas, V.M., Atkinson, H.J., 1994. Proteinases of females of the phytoparasite *Globodera pallida* (potato cyst nematode). Parasitology 109, 357–365.

Krem, M.M., Rose, T., Di Cera, E., 1999. The C-terminal sequence encodes function in serine proteases. Journal of Biological Chemistry 274, 28063–28066.

Lilley, C., Urwin, P.E., Atkinson, H.J., Mcpherson, M.J., 1997. Characterization of cDNAs encoding serine proteinases from the soybean cyst nematode *Heterodera glycines*. Molecular and Biochemical Parasitology 89, 195–207.

Lilley, C.J., Devlin, P., Urwin, P.E., Atkinson, H.J., 1999. Parasitic nematodes, proteinases and transgenic plants. Parasitology Today 15, 414–417.

Lilley, C.J., Urwin, P.E., McPherson, M.J., Atkinson, H.J., 1996. Characterization of intestinally active proteinases of cyst-nematodes. Parasitology 113, 415–424.

Longbottom, D., Redmond, D.L., Russell, M., Liddell, S., Smith, W.D., Knox, D.P., 1997. Molecular cloning and characterisation of a putative aspartate proteinase associated with a gut membrane protein complex from adult *Haemonchus contortus*. Molecular and Biochemical Parasitology 88, 63–72.

Marchler-Bauer, A., Anderson, J.B., DeWeese-Scott, C., Fedorova, N.D., Geer, L.Y., He, S., Hurwitz, D.I., Jackson, J.D., Jacobs, A.R., Lanczycki, C.J., Liebert, C.A., Liu, C., Madej, T., Marchler, G.H., Mazumder, R., Nikolskaya, A.N., Panchenko, A.R., Rao, B.S., Shoemaker, B.A., Simonyan, V., Song, J.S., Thiessen, R.A., Vasudevan, S., Wang, Y., Yamashita, R.A., Yin, J.J., Bryant, S.H., 2003. CDD, a curated Entrez database of conserved domain alignments. Nucleic Acids Research 31, 383–387.

Mazumdar-Leighton, S., Babu, C.R., Bennett, J., 2000. Identification of novel serine proteinase gene transcripts in the midguts of two tropical insect pests, *Scirpophaga incertulas* (Wk.) and *Helicoverpa armigera* (Hb.). Insect Biochemistry and Molecular Biology 30, 57–68.

McCarter, J.P., Mitreva, M.D., Martin, J., Dante, M., Wylie, T., Rao, U., Pape, D., Bowers, Y., Theising, B., Murphy, C.V., Kloek, A.P., Chiapelli, B.J., Clifton, S.W., Bird, D.M., Waterston, R.H., 2003. Analysis and functional classification of transcripts from the nematode *Meloidogyne incognita*. Genome Biology 4, R26.

Michaud, D., Cantin, L., Bonade-Bottino, M., Jouanin, L., Vrain, T.C., 1996. Identification of stable plant cystatin/nematode proteinase complexes using mildly denaturing gelatin/polyacrylamide gel electrophoresis. Electrophoresis 17, 1373–1379.

Muharsini, S., Dalrymple, B., Vuocolo, T., Hamilton, S., Willadsen, P., Wijffels, G., 2001. Biochemical and molecular characterization of serine proteases from larvae of *Chrysomya bezziana*, the Old World Screwworm fly. Insect Biochemistry and Molecular Biology 31, 1029–1040.

Mulenga, A., Sugimoto, C., Ingram, G., Ohashi, K., Misao, O., 2001. Characterization of two cDNAs encoding serine proteinases from the hard tick *Haemaphysalis longicornis*. Insect Biochemistry and Molecular Biology 31, 817–825.

Neveu, C., Jaubert, S., Abad, P., Castagnone-Sereno, P., 2003. A set of genes differentially expressed between avirulent and virulent *Meloidogyne incognita* near-isogenic lines encode secreted proteins. Molecular Plant–Microbe Interactions 16, 1077–1084.

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.

Nilsen, T.W., 1993. Trans-splicing of nematode premessenger RNA. Annual Review of Microbiology 47, 413–440.

Parkinson, J., Mitreva, M., Hall, N., Blaxter, M., McCarter, J.P., 2003. 400000 nematode ESTs on the Net. Trends in Parasitology 19, 283–286.

Perona, J.J., Craik, C.S., 1997. Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. Journal of Biological Chemistry 272, 29987–29990.

Redmond, D.L., Knox, D.P., Newlands, G., Smith, W.D., 1997. Molecular cloning and characterisation of a developmentally regulated putative metallopeptidase present in a host protective extract of *Haemonchus contortus*. Molecular and Biochemical Parasitology 85, 77–87.

Reversat, G., 1981. Consumption of food reserves by starved second-stage juveniles of *Meloidogyne javanica* under conditions including osmobiosis. Nematologica 27, 207–214.

Roberts, P.A., 1992. Current status of the availability, development, and use of host plant resistance to nematodes. Journal of Nematology 24, 213–227.

Ross, L.H., Freedman, J.H., Rubin, C.S., 1995. Structure and expression of novel spliced leader RNA genes in *Caenorhabditis elegans*. Journal of Biological Chemistry 270, 22066–22075.

Rubin, G.M., Yandell, M.D., Wortman, J.R., Gabor Miklos, G.L., Nelson, C.R., Hariharan, I.K., Fortini, M.E., Li, P.W., Apweiler, R., Fleischmann, W., Cherry, J.M., Henikoff, S., Skupski, M.P., Misra, S., Ashburner, M., Birney, E., Boguski, M.S., Brody, T., Brokstein, P., Celniker, S.E., Chervitz, S.A., Coates, D., Cravchik, A., Gabrielian, A., Galle, R.F., Gelbart, W.M., George, R.A., Goldstein, L.S., Gong, F., Guan, P., Harris, N.L., Hay, B.A., Hoskins, R.A., Li, J., Li, Z., Hynes, R.O., Jones, S.J., Kuehl, P.M., Lemaitre, B., Littleton, J.T., Morrison, D.K., D.K, Mungall, C., O'Farrell, P.H., Pickeral, O.K., Shue, C., Vosshall, L.B., Zhang, J., Zhao, Q., Zheng, X.H., Lewis, S., 2000. Comparative genomics of the eukaryotes. Science 287, 2204–2215.

Sakanari, J.A., 1990. Anisakis—from the platter to the microfuge. Parasitology Today 6, 323–327.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning. A Laboratory Manual, second ed. Cold Spring Harbor Laboratory Press, New York.

Sasser, J.N., 1980. Root-knot nematodes: a global menace to crop production. Plant Disease 64, 36–41.

Sasser, J.N., Eisenback, J.D., Carter, C.C., 1983. The international *Meloidogyne* project—its goals and achievements. Annual Review of Phytopathology 21, 271–288.

Shomaker, C.H., Been, T.H., 1999. A model for infestation foci of potato cyst nematodes, *Globodera rostochiensis* and *G. pallida*. Phytopathology 89, 583–590.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673–4680.

Trudgill, D.L., 1995. Origins of root-knot nematodes (*Meloidogyne* spp. Nematoda) in relation to their cultural control. Phytoparasitica 23, 191–194.

Trudgill, D.L., Blok, V.C., 2000. Apomictic, polyphagous root-knot nematodes: exceptionally successful and damaging biotrophic root pathogens. Annual Review of Phytopathology 39, 53–77.

Urwin, P.E., Lilley, C.J., McPherson, M.J., Atkinson, H.J., 1997a. Characterization of two cDNAs encoding cysteine proteinases from the soybean cyst nematode *Heterodera glycines*. Parasitology 114, 605–613.

Urwin, P.E., Lilley, C.J., McPherson, M.J., Atkinson, H.J., 1997b. Resistance to both cyst and root-knot nematodes conferred by transgenic *Arabidopsis* expressing a modified plant cystatin. Plant Journal 12, 455–461.

Urwin, P.E., McPherson, M.J., Atkinson, H.J., 1998. Enhanced transgenic plant resistance to nematodes by dual proteinase inhibitor constructs. Planta 204, 472–479.

Valaitis, A.P., Augustin, S., Clancy, K.M., 1999. Purification and characterization of the western spruce budworm larval midgut proteinases and comparison of gut activities of laboratory-reared and field-collected insects. Insect Biochemistry and Molecular Biology 29, 405–415.

Vercauteren, I., Van Der Schueren, E., Van Montagu, M., Gheysen, G., 2001. *Arabidopsis thaliana* genes expressed in the early compatible interaction with root-knot nematodes. Molecular Plant–Microbe Interactions 14, 288–299.

Williamson, V.M., Hussey, R.S., 1996. Nematode pathogenesis and resistance in plants. Plant Cell 8, 1735–1745.

Wyss, U., Grundler, F.M.W., Munch, A., 1992. The parasitic behaviour of second-stage juveniles of *Meloidogyne incognita* in roots of *Arabidopsis thaliana*. Nematologica 38, 98–111.

Yoshida, S., Taniguchi, M., Suemoto, T., Oka, T., He, X., Shiosaka, S., 1998. cDNA cloning and expression of a novel serine protease, TLSP. Biochimica et Biophysica Acta 1399, 225–228.

Zhu, Y.C., Kramer, K.J., Dowdy, A.K., Baker, J.E., 2000. Trypsinogen-like cDNAs and quantitative analysis of mRNA levels from the Indianmeal moth, *Plodia interpunctella*. Insect Biochemistry and Molecular Biology 30, 1027–1035.