

Molecular characterisation and expression of two venom allergen-like protein genes in *Heterodera glycines*[☆]

Bingli Gao^a, R. Allen^a, Tom Maier^b, Eric L. Davis^c, Thomas J. Baum^b, Richard S. Hussey^{a,*}

^aDepartment of Plant Pathology, University of Georgia, Athens, GA 30602-7274, USA

^bDepartment of Plant Pathology, Iowa State University, 351 Bessey Hall, Ames, IA 50011, USA

^cDepartment of Plant Pathology, North Carolina State University, Box 7616, Raleigh, NC 27695-7616, USA

Received 25 June 2001; received in revised form 30 July 2001; accepted 30 July 2001

Abstract

Secretory proteins encoded by genes expressed in the oesophageal gland cells of plant-parasitic nematodes have key roles in nematode parasitism of plants. Two venom allergen-like protein cDNAs (designated *hg-vap-1* and *hg-vap-2*) were isolated from *Heterodera glycines* gland cell cDNA libraries. Both cDNAs hybridised to genomic DNA of *H. glycines* in Southern blots. The *hg-vap-1* cDNA contained an open reading frame encoding 215 amino acids with the first 25 amino acids being a putative secretion signal. The *hg-vap-2* cDNA contained an open reading frame encoding 212 amino acids with the first 19 amino acids being a putative secretion signal. Genes of *hg-vap-1* and *hg-vap-2* contained four introns, which ranged in size from 44 to 574 bp, and five exons ranging in size from 43 to 279 bp. In situ hybridisation analyses showed that mRNAs of both *vap* genes accumulated specifically in the subventral gland cells of *H. glycines* during parasitism. The gland cell-specific expression and presence of predicted secretion signal peptides in both VAPs suggest that these proteins are secreted from the nematode and may play a role in the infection of host plants by this parasite. © 2001 Australian Society for Parasitology Inc. Published by Elsevier Science Ltd. All rights reserved.

Keywords: AG5; Oesophageal gland cells; Nucleotide sequence; Plant-parasitic nematode; Secretory proteins; Cyst nematode

1. Introduction

Proteins similar to the allergen antigen 5 of extracellular proteins from hymenopteran insect venom (King et al., 1990; Fang et al., 1988) constitute a family of secreted proteins that have been found in animal-parasitic nematodes, e.g., *Ancylostoma caninum* (Hawdon et al., 1996, 1999), and *Onchocerca volvulus* (Tawe et al., 2000), the free-living nematode *Caenorhabditis elegans* (Bin et al., 1999), and the plant-parasitic nematode *Meloidogyne incognita* (Ding et al., 2000). Secretory venom allergen-like proteins are characterised by the highly conserved spacings of cysteine residues, particularly in the carboxy-terminal region (Hawdon et al., 1996). The best characterised nematode member of this family of secreted proteins is Ac-ASP-1, which is the major ES protein released by infective third-stage larvae of *A. caninum* stimulated to feed in vitro

(Hawdon et al., 1996). The secretion of Ac-ASP-1 coincides with the invasion of the host by the nematode larvae, indicating the protein may have a role during the early stages of parasitism. The biological function of ES venom allergen-like proteins from the other animal-parasitic nematodes is also unknown, although their expression is associated with parasitic nematode stages. In adult *Haemonchus contortus* an ES venom allergen-like protein was reported to be produced in the oesophagus, suggesting that this protein may be secreted during feeding (Schallig et al., 1997).

Three large, transcriptionally active oesophageal gland cells, one dorsal and two subventral, are the principal sources of the secretions used by Tylenchid nematodes during infection and parasitism of plants (Hussey and Grundler, 1998). Cloning parasitism genes encoding the secretory proteins expressed in oesophageal gland cells is the key to understanding the molecular basis of nematode parasitism of plants (Davis et al., 2000). Recently, micro-aspiration of *Heterodera glycines* oesophageal gland cell cytoplasm to provide expressed mRNA for RT-PCR was used to generate a suppression subtractive hybridisation cDNA library enriched in genes expressed preferentially in the oesophageal gland cells (Gao et al., 2001; Wang et

[☆] Nucleotide sequence data reported in this paper are available in the GenBank™ EMBL and DDBJ databases under the accession numbers AF343567 and AY028639 for the cDNAs and AF374388 and AY033601 for the genomic clones.

* Corresponding author. Tel.: +1-706-542-1254; fax: +1-706-542-1262.

E-mail address: hussey@arches.uga.edu (R.S. Hussey).

al., 2001). In analysing the gland cell suppression subtractive hybridisation cDNA library, a cDNA clone that had homology with secretory venom allergen-like proteins was isolated (Gao et al., 2001). We report here the cloning of a second venom allergen-like protein cDNA from *H. glycines*, isolation and analysis of putative full-length cDNAs and genomic clones of both of the *H. glycines* venom allergen-like proteins, and complete expression profiles of both genes throughout the *H. glycines* parasitic cycle.

2. Materials and methods

2.1. Nematode culture

Heterodera glycines was cultured on soybean (*Glycine max*) in the greenhouse. Preparasitic second-stage juveniles (J2) were collected by hatching eggs on sieves with 25- μ m openings suspended over deionised water in plastic bowls. Mixed stages of parasitic *H. glycines* were collected 8 days after inoculation by root blending and sieving (de Boer et al., 1999).

2.2. Isolation of cDNA clones

One cDNA (clone GSB9) encoding a venom allergen-like protein, designated *hg-vap-1* for *H. glycines* venom allergen-like protein (Sutherland, M., Westlund, B., Burnam, L., Sluder, A., Liu, L. 2001. Characterization of two *Caenorhabditis elegans* venom allergen-like proteins. Keystone Symposia Abstract, Taos, NM. January 20–25, 2001), was isolated from a *H. glycines* oesophageal gland cell cDNA library constructed by suppression subtractive hybridisation (Clontech Laboratories, Inc.) that used cDNA derived from microaspirated cytoplasm from oesophageal gland cells of *H. glycines* mixed parasitic stages as ‘tester’ and cDNA derived from the intestinal region of *H. glycines* parasitic stages as ‘driver’ (Diatchenko et al., 1996; Gao et al., 2001). A second putative full-length cDNA (clone 2A05) encoding a different venom allergen-like protein, designated *hg-vap-2*, was identified during random sequencing of clones from a long distance PCR cDNA library (Clontech Laboratories, Inc.) derived from gland cell cytoplasm of parasitic stages of *H. glycines* (Gao et al., 2001). The putative full-length cDNA of *hg-vap-1* was obtained by screening a nylon membrane macroarrayed with 6,130 gland cell long distance PCR cDNAs as previously described (Gao et al., 2001). Sequences of putative full-length *hg-vap-1* and *hg-vap-2* cDNAs, both in pGEM-T Easy vector, were obtained using T7 and SP6 primers in sequencing reactions.

2.3. Genomic clones

Genomic clones of both *hg-vap-1* and *hg-vap-2* were obtained by PCR using gene-specific primers and 2 μ l of a lysed Lambda Fix II *H. glycines* genomic DNA library (Yan et al., 1998) as template. Forward primer GSB9F (5'-

CTTTAATCCTCCAATCCGTCCG-3'), which was located in the 5' UTR region, and reverse primer GSB9R (5'-CCATTGTAATCCGAACACTTGCT-3'), which was located 40 bp upstream of the stop codon TAA, were used to PCR-amplify *hg-vap-1* genomic DNA. Similarly forward primer 2A05F1 (5'-ATGCACTTGATTAACCTAATCGCC-3'), which began with the start codon ATG, and reverse primer 2A05R (5'-CTTTGTACATTTGTGCTTTG CTTT-3'), which was located 20 bp upstream of the stop codon TGA, were used to amplify *hg-vap-2* genomic DNA. PCR products were cut from gels and purified with a QIAquick Gel Extraction Kit (Qiagen Inc.), ligated into pGEM-T Easy vector (Promega), and transformed into *Escherichia coli* DH5- α . Inserts were excised from plasmids and sequenced.

2.4. Southern hybridisation

Three hundred microliters of packed *H. glycines* J2 were frozen in liquid nitrogen and manually ground with a smooth metal mortar to extract genomic DNA as described by Ray et al. (1994). Ten micrograms of *H. glycines* genomic DNA digested with *Eco*RI were used per lane for electrophoresis in 1% agarose gels. Electrophoresed genomic DNA was transferred onto Hybond-N Nylon membranes (Amersham Pharmacia Biotech). Probes specific to *hg-vap-1* and *hg-vap-2* were prepared by PCR-amplification of full-length cDNAs from recombinant pGEM-T Easy vectors (Promega) using T7 and SP6 primers. PCR products were gel-separated, cut from the gel, purified with QIAquick Gel Extraction Kit (Qiagen Inc.), and then labelled with ^{32} P-dCTP using RTS RadPrime DNA Labelling System (GIBCO-BRL). High stringency hybridisation (overnight hybridisation at 68°C and washed with 0.1 SSC and 0.1% SDS at 68°C) was performed according to the standard procedure (Sambrook et al., 1989). Films were exposed overnight at –80°C prior to developing.

2.5. In situ hybridization

For *hg-vap-1*, Nested Primer 1 (5'-TCGAGCGGCCGC-CCGGGCAGGT-3') and Nested Primer 2R (5'-AGCG-TGGTCGCGGCCGAGGT-3') (Clontech) were used to PCR-amplify the cDNA fragment (495 bp) from suppression subtractive hybridisation library clone GSB9 (Gao et al., 2001). For *hg-vap-2*, forward primer 2A05F (5'-ACCGGATCTGATTATGACCGG-3') and reverse primer 2A05R (5'-CTTTGTACATTTGTGCTTTGCTTT-3') were used to amplify the cDNA fragment (236 bp) from long distance PCR library clone 2A05. The PCR products were used to synthesise digoxigenin-labeled sense and antisense cDNA probes (Boehringer Mannheim GmbH) by asymmetric PCR (Wang et al., 2001) amplification with Nested Primer 1 and Nested Primer 2R or 2A05F and 2A05R.

In situ hybridisation was performed as described by de Boer et al. (1998). Mixed stages of parasitic *H. glycines* were washed three times with 1 $\times$ PBS buffer and fixed in

10% formalin buffered in phosphate (pH 7.4) for 2 days at room temperature. The fixed nematodes were randomly cut on glass slides with a razor blade until approximately 90% of the nematodes were cut. Nematode sections were partially digested with 500 ng/ml proteinase-K (Boehringer Mannheim) at room temperature for 1 h. After pre-hybridisation, nematode sections in hybridisation solution were aliquoted to wells of a 96-well MultiScreen plate (Millipore). Denatured PCR digoxigenin-labelled DNA probes (1 μ l) were added to each well. Solutions were changed by aspiration through the well bottoms using the Millipore MultiScreen vacuum manifold. cDNA probes that hybridised within the nematode sections were detected by light microscopy.

3. Results

3.1. cDNAs

The putative full-length *hg-vap-1* cDNA contained 810 nucleotides (excluding the poly (dA) tail) with a putative open reading frame (ORF) of 648 bp (Fig. 1). The cDNA contained 90 bp of 5' UTR and a 72 bp of 3' UTR, which contained a polyadenylation signal (AATAAA). The ORF encoded a deduced protein of 215 amino acids with a calculated mol. wt. of 23,176 and a pI of 8.6, predicted by ProtParam tool (Gill and von Hippel, 1989). A putative signal sequence, predicted by Signal P (Nielsen et al., 1997), terminated immediately upstream of a protease cleavage site between amino acids Ala25 and Leu26. There was one putative *N*-glycosylation site at amino acid Asn52. The *hg-vap-1* cDNA did not contain a nematode spliced leader sequence at the 5' end.

The putative ORF of the *hg-vap-2* cDNA was 639 bp (Fig. 2). The putative full-length cDNA contained 795 nucleotides (excluding the poly (dA) tail), 36 bp of 5' UTR, and 120 bp of 3' UTR, which contained a polyadenylation signal (AATAAA). The ORF encoded a deduced protein of 212 amino acids with a calculated mol. wt. of 23,941 and a pI of 6.9. A putative signal sequence terminated immediately upstream of a protease cleavage site between amino acids Gln19 and Gln20. A putative *N*-glycosylation site occurred at amino acid Asn137. A nematode spliced leader sequence was not present at the 5' end of the *hg-vap-2* cDNA.

In Southern blot hybridisations at high stringency, both *hg-vap-1* and *hg-vap-2* cDNAs hybridized with *Eco*RI digested genomic DNA from *H. glycines*. *Hg-vap-1* cDNA hybridised to one major band (~3 kb) and two minor bands (~7 kb and ~1.5 kb) and *hg-vap-2* cDNA hybridised to two major (~7.5 kb and ~6 kb) (data not shown). Since the cDNA sequences show no *Eco*RI sites, the hybridisation of the probes with multiple bands suggests that there may be three or more *vap*-like genes in the genome of *H. glycines*.

3.2. Genomic clones of *hg-vap-1* and *hg-vap-2*

Hg-vap-1 and *hg-vap-2* genes were obtained by PCR amplification from a *H. glycines* genomic library (Yan et al., 1998). The lengths of the genomic DNA stretches containing *hg-vap-1* and *hg-vap-2* were 1555 and 880 bp, respectively (Fig. 3). Both PCR products were sequenced (GenBank accession numbers AF374388 and AY033601). The exon/intron boundaries of the genomic sequences were determined by aligning the genomic sequences with the corresponding cDNA sequences (Figs. 1–3). Both *hg-vap-1* and *hg-vap-2* contained four introns, which varied in length from 44 to 574 bp, with an average intron size of 143 bp. (Fig. 3). The signal sequences of both genes were interrupted by an intron. The size of the corresponding introns between *hg-vap-1* and *hg-vap-2* was very different, with introns for *hg-vap-1* being larger at three positions (e.g., first intron was 574 bp compared with 64 bp for *hg-vap-2*). The *cis*-splicing in the *hg-vap-1* and *hg-vap-2* genes followed the 'GU-AG rule' (Blumenthal and Steward, 1997) with the exception of the third intron in *hg-vap-1*, which used GC as a 5'-splicing donor sequence instead of GU. The exon size for the two genes ranged from 43 to 279 bp, with an average exon size of 153 bp. In contrast to the introns, the size of the corresponding exons between *hg-vap-1* and *hg-vap-2* was very similar.

3.3. Sequence similarities

BLAST searches (Benson et al., 1998) revealed the predicted Hg-VAP-1 and Hg-VAP-2 amino acid sequences had highest similarities with extracellular venom allergen-like proteins from *M. incognita* (Mi-MSP-1; AF013289, 114/e⁻²⁴), *A. caninum* (Ac-ASP-1, AF079521, 118/e⁻²⁵ and Ac-ASP-2, AF089728, 117/e⁻²⁵), and *C. elegans* (C39E9.2, T19852, 131/e⁻³⁰ and F49E11.6, T22438, 130/e⁻²⁹). Alignment of sequences was made with a multiple alignment program (Huang, 1994). The local alignment of Hg-VAP-1 and Hg-VAP-2 with the venom allergen-like proteins from nematodes revealed significant sequences were conserved (Fig. 4). Hg-VAP-1 and Hg-VAP-2 contain the eight conserved cysteine residue spacings (except Cys163 differs by one spacing from the conserved Cys167 in Ac-ASP-1 and Ac-ASP-2) and the sequence HWSQ (Hg-VAP-1 aa 144–147) or HWTQ (Hg-VAP-2 aa 143–146), which is a conserved variation of the HYTQ sequence found in most members of the venom allergen-like proteins (Hawdon et al., 1996). Hg-VAP-1 contains four additional cysteine residues and Hg-VAP-2 contains two additional ones. Cys85, Cys162, Cys163, and Cys175 are known to be involved in disulfide bonds in other cysteine-rich proteins (Hawdon et al., 1999). Both Hg-VAP-1 and Hg-VAP-2 sequences align with the C-terminus of Ac-ASP-1.

Comparison of the Hg-VAP-1 and Hg-VAP-2 deduced amino acid sequences minus the putative signal sequence indicated that the sequences were 45% identical and 58%

GSB9F		
1	gattcataatctgtccgcttttaatcctccaatccgctccgtctccattcatttccaaatta	60
61	ttcaaatcctacaataacaatacataaaaaATGGAGCATTTCCTTCTAAATTTGTTCTC	120
1	M E H F P S K F V L	10
121	TGCTTGCTCATTGCATTTCATCGCCGTTCCATGCTCTGTTTTGGCACTTTCTGCTGGCAGC	180
11	C L L I A F I A V P C S V L A L S A G S	30
	intron splice site 1	
181	CGCGTGTGATGCTCGCGTGCCACAACAATTACCGCTCACAAGTGGCCAAAGGCACTGCG	240
31	R V S M L A C H N N Y R S Q L A K G T A	50
241	GACAACAAATCGGGCAAATGCCGGCTGGCAGCAATTTGATCCAACAAAAATATGATTTCG	300
51	D N K S G K M P A G S N L I Q Q K Y D S	70
301	AATATTGAGTCGGTGGCGCAGAAGTGGGCCAACGGATGCTCAATGTCCCACTCGGGTTCA	360
71	N I E S V A Q K W A N G C S M S H S G S	90
361	TCAGGGATGGGCGAGAATTTGTACATGACTAGTTCATCGACAATTTCGGAAGCTGATGCG	420
91	S G M G E N L Y M T S S S T I S E A D A	110
	intron splice site 2	
421	CTGAAACAAGCTTGTGACATGTGGTGGGCGGAGTTGAAGCAGTTCGGCTTTCAGTCAAGT	480
111	L K Q A C D M W W A E L K Q F G F Q S S	130
481	CTTGTGCTGGACATGGCTCAGTTCAACAAGGGCATCGGACATTGGAGCCAGCAAGCATGG	540
131	L V L D M A Q F N K G I G H W S Q A W	150
	intron splice site 3	
541	GCCTCCACTGCTCAAATTGGCTGTGCCATGGCACGTTGTCCGTCCTCCACCTGGAAGACA	600
151	A S T A Q I G C A M A R C P S S T W K T	170
601	TGGGTTGTGTGCAACTACAAGGAGGCTGGCAACTTCTTGAACCAGCCGGTCTACAAAAAA	660
171	W V V C N Y K E A G N F L N Q P V Y K K	190
	intron splice site 4 GSB9R	
661	GGTGCCGCGTGCAGCAAGTGTTCGGATTACAATGGGGCGACCTGTGACAGCGCCAGCGGA	720
191	G A A C S K C S D Y N G A T C D S A S G	210
721	CTGTGCAAATTGCCATAAgtgaggggagaatgcttggaagaaggatttttttaatttgatt	780
211	L C K L P *	216
781	tatgaatattgtagaatgggagtgaaataatgggcattaattgaaaaaaaaaaaaaaaaaaaa	840

Fig. 1. Complete nucleotide and deduced amino acid sequences of the *Heterodera glycines* *hg-vap-1* cDNA. The predicted 25 amino acid secretion signal sequence is indicated in bold type. The gene-specific primers used for amplification of genomic DNA are underlined. A putative *N*-linked glycosylation site (Asn52) is double underlined. The TAA stop codon is marked with an asterisk and the putative polyadenylation signal sequence (aataaa) is overlined. The intron splice sites are shown with arrows. The cDNA fragment used in situ localisation is from cDNA position 385 to the 3'-end.

similar (Table 1). When compared with other venom allergen-like protein family members from selected nematode species, Hg-VAP-1 was found to be 30–40% identical and 44–55% similar and Hg-VAP-2 was 32–38% identical and 43–51% similar. The highest level of identity and similarity was with the VAP homologues in *C. elegans* (Table 1).

3.4. Gland cell-specific *hg-vap-1* and *hg-vap-2* expression

To determine the developmental expression pattern of *hg-vap-1* and *hg-vap-2* and the specific tissues in which they are expressed, different life stages of *H. glycines* were hybri-

dised with digoxigenin-labelled cDNA probes. In situ hybridisation showed the digoxigenin-labelled antisense cDNA probes of *hg-vap-1* and *hg-vap-2* specifically hybridised with mRNA within the subventral oesophageal gland cells of *H. glycines* (Figs. 5 and 6). The hybridisation signal was usually intense in the cytoplasm in the region of the gland cells with the nucleus and was absent in the gland cell extensions. No hybridisation with mRNA was observed within sections of *H. glycines* with the control sense cDNA probes of either *hg-vap-1* or *hg-vap-2*. *Hg-vap-1* transcripts were localised within the subventral gland cells of unhatched second-stage juveniles indicating synthesis of

2A05F1

1 aaaccataaattacaacttaaatacaagcaaaatcaaATGCACTTGATTAACTTAATCGCC 60
 1 M H L I N L I A 8

61 CTCTTTTTCATGCTTTTTCGGCCCATCCGTCCAGCAATACACAAAAACGCCAACGAATGAG 120
 9 L F F M L F G▲ P S V Q Q Y T K T P T N E 28
 intron splice site 1

121 GACAAAGAAGCGGCTGTCAATTGTGACAAACAAATTCGGATCGCAATTGGCCCTGGGCAAT 180
 29 D K E A A V N C H N K F R S Q L A L G N 48

181 GCCGACAATAAATTGGGCGGCAACAAAATGCCAAAGGCGGGCAACATGCGTAAGTTTGAA 240
 49 A D N K L G G N K M P K A G N M R K F E 68

241 TGGGACGAAAACCTTGCCAAACTTGCGGATGAATGGGCAACAAATGCACATTATCGCAC 300
 69 W D E N L A K L A D E W A N K C T L S H 88

301 TCGTGGAACGGCTGGGCAGGCGAAAAATTTGGCAATGAATGGCGGAACATTTTCGAACAAG 360
 89 S W N G W A G E N L A M N G G T F S N K▲ 108
 intron splice site 2

361 GATGGCTTCGAGTACGCTTGCGGTCGCTGGTGGGACGAACTGAACCGTTACGGGTTC AAC 420
 109 D G F E Y A C G R W W D E L N R Y G F N 128

2A05F

421 CCGGATCTGATTATGACCGGGGAAAACTTCAGTGGCATCGGCCATTGGACTCAGATGGCG 480
 129 P D L I M T G E N F S G I G H W T Q▲ M A 148
 intron splice site 3

481 TGGGCCGACACCGACCGAATTGGCTGTGCCATGGCACAAAAC TGCCCAAATACCAATTGG 540
 149 W A D T D R I G C A M A Q N C P N T N W 168

541 AAAACATATGTGGTCTGCTGGTATTACACGGGTGGTAATTACTTTGGTGC GCCTGTCTAT 600
 169 K T Y V V C W Y Y T G▲ G N Y F G A P V Y 188
 intron splice site 4

2A05R

601 TTGGCCCCGGGAGCCGTGCAGCAAATGC AAAGCAAATGTGACAAAGCCACTGGA 660
 189 L A R E P C S K C K A N D K C D K A T G 208

661 CTTTGCTCTCAATGAaagatgaaatattatgaaaattgaattgggataattactaattat 720
 209 L C S Q * 213

721 tgtatctgtatgttttagtatggtttagtgtttcgggtataattttcacttattgggtcttaa 780

781 ttaacaatgaataaataaatagttttctaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 834

Fig. 2. Complete nucleotide and deduced amino acid sequences of the *Heterodera glycines* *hg-vap-2* cDNA. The predicted 19 amino acid secretion signal sequence is indicated in bold type. The gene-specific primers used for amplification of genomic DNA (2A05F1 and 2A05R) and the fragment used for in situ hybridisation (2A05F and 2A05R) are underlined. A putative N-linked glycosylation site (Asn137) is double underlined. The TGA stop codon is marked with an asterisk and the putative polyadenylation signal sequence (aataaa) is overlined. The intron splice sites are shown with arrows.

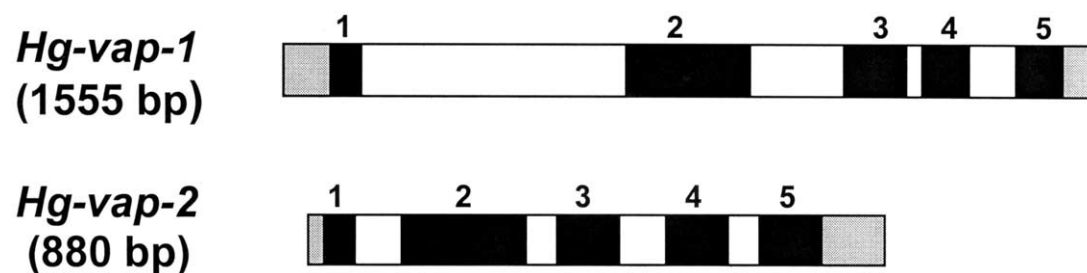


Fig. 3. Organisation of *Heterodera glycines* extracellular venom allergen-like protein genes, *hg-vap-1* and *hg-vap-2*. UTRs (grey boxes), exons (dark boxes), and introns (open boxes) are drawn schematically to indicate their relative positions and sizes.

Hg-VAP-1	1	MEHFPSKFVLCLLIAFIAVP-----CSVLALSAGSRVSMILACHN-----
Hg-VAP-2	1	----MHLINLIALFFMLFGP----SVQYTKTPINEDKEAAVNCHN-----
Mi-MSP-1	1	-MSNKLIIISILILITIIYT-----VVNSLTVPEQNAVVDICIN-----
C39E9.2	1	-----MHTSILLLLTALCAG-----AYAQFSANGQAAILNVHN-----
F49E11.6	1	-----MKSILLLVCCIIAG-----VFSQFTSTGQQAIVDAHN-----
Ac-ASP-1	1	-MFSPVIVSVIFTIAFCDA SPARDGFGCSNSGITDKDROAFIDFHNNARRRVAKGVEDSN
Ac-ASP-2	1	-----MLVLVPLIALIIVSVHGNMRCGNNGMTDEARQKFLDVHN-----
VAAP-1	1	---MLFFVIFSIITIAVVAG-----FECPGGRLTPQQRKDIVRQNN-----
Hg-VAP-1	40	-----NYRSQ LAKGTA--DNKSGKMPAGSNLIQQKYDSNIESVAQKWANGCSMSHSG-
Hg-VAP-2	39	-----KFRSQA LAGNADNKLGGNKMPKAGNMRKFEDENLAKLADEWANKCTLSHSWN
Mi-MSP-1	36	-----KYRSQ LANGKTK-NKNGGNFPGKDILEVSYSKDLEKSAQRWANKCIFDHNGT
C39E9.2	33	-----TLRSRI AKGTY--VAKGTAKPAASDMLKMKWDATVAASQAAYANKCPTGHSQA
F49E11.6	32	-----KLRSSI AKGTY--VAKGTTQKSGSNMRKIKWDATVATSAQNYANTCPTGHSQG
Ac-ASP-1	240	TFLSVHNEFRSSVARGLEP-DALGGNAPKAAKMLKMVYDCEVEASAIIRHGKNCVYQHSHG
Ac-ASP-2	42	-YR-----SMVAKGQAK-DAISGNAPKAAKMKMIYDCNVESTAMQNAKKCVFAHSH-
VAAP-1	38	-----KFRSLLIHGK LK-NRNGTYMPRGKNMILLKWSCOLENSAQRWANQCVFHSR
Hg-VAP-1	90	-----SSGMGENLYMT--SSSTISEADALKQACDMWWAELKQEGFQSSSLVLDMQAQFNK---
Hg-VAP-2	92	-----GWAGENLAMN---CGTFSNKDGF EYACGRWWDENLRYGFNPDLIMTGENFSG---
Mi-MSP-1	88	DLYSGGKFYGENLYLDGDFEHKNITQLMIDACNAWWGESTTDDGVPPSWINNFLTNDKEN
C39E9.2	84	A----GLGENLYWYWT--SATITNIDQFGATGSAAWKEKFQDYGWS--SNTLSMSLFNT---
F49E11.6	83	S----GYGENLYWYWT--SCTIGNLDTFGPAASSSWESEFQQYQWGT--SNTLDMNTFNT---
Ac-ASP-1	299	ED---RPGLGENIYKT--SVLKFDKNKAAKQASQLWWNELKEFGVPSNVLTITALLNRP--
Ac-ASP-2	92	-----RKGVGENIWM S--TARQMDKAQAAQQASDGFSELA KYGVQGENKLTITQLWNR---
VAAP-1	90	NQRQ--CIGENVYAYWSSSESVEKL RNTAGTEAGKS WWSSEL PKLYKQNP SNLTDVAR---
Hg-VAP-1	141	-----GIGHWSQA WASTAQIGCAMARCPSSTWKT---WVVCNYKEAGNELNQPVYKKG
Hg-VAP-2	141	-----IGHWTQMAWADTD RIGCAMAQNCPTN---WKTYVVCWYYTG GNYFGAPVYLXR
Mi-MSP-1	148	DEKFEAVGHWTQMAWAKTYQIGCALKVCHKPDCNG---NLIDCRYYPGGNGMGSP IYQQG
C39E9.2	135	-----GIGHATQMAWAKTNLIGCGVKNC GKDTN--GFNKVTVVCQYKPGQNYLNQNIYTSG
F49E11.6	134	-----GIGHATQMAWANTFAIGCGVKNC GKDP SNGYNKVAVVCQYKTPGNYLNQPIYQQG
Ac-ASP-1	353	---GMQIGHYTQMAWDTT YKLGCAVVF CNDFTF-----GVCQYGP GGNMGHVIYTMG
Ac-ASP-2	143	---GVMIGHYTQMVWQESYKLG CYVEWCSSMTY-----GVCQYSPQGNM MNSLIYEKG
VAAP-1	146	---QGVLFHTQMAWGKTHKIGCGIATNCDGGR---TLIATCHYSPAGNMLKELIYELG
Hg-VAP-1	192	AACS--KCS DYN GATCD SAGLCKLP-----
Hg-VAP-2	192	EPCS---KCKANDKCDKATGLCSQ-----
Mi-MSP-1	205	KPAS---GCGKAGPSTKYSG LCKPDPHQNN
C39E9.2	189	TTCS---KCPSGTSC EAATGLCA-----
F49E11.6	189	TTCA---ACPSGTACDSS--GLCA-----
Ac-ASP-1	403	QPCS---QCSPGATCSVTEGLCSAP-----
Ac-ASP-2	193	NPCTKDSDCGSNASC SAGEALCVVRG----
VAAP-1	198	EPCKTDSDCNTKKCAKKS--GLCRK-----

Fig. 4. Sequence alignment among predicted amino acid sequences of Hg-VAP-1 and Hg-VAP-2 for *Heterodera glycines* with proteins from other nematodes; Hg-VAP-1 (GenBank Accession no. AF343567) from *H. glycines*; Hg-VAP-2 (AY028639) from *H. glycines*; Mi-MSP-1, a venom allergen AG5-like protein from *M. incognita* (AF013289); C39E9.2, hypothetical protein from *C. elegans* (T19852); F49E11.6, hypothetical protein from *C. elegans* (T22438); Ac-ASP-1, Ancylostoma-secreted protein precursor from *Ancylostoma caninum* (Q16937); Ac-ASP-2, secreted protein ASP-2 precursor from *A. caninum* (AF089728); VAAP-1, venom allergen antigen-like protein from *Brugia malayi* (AF334661). Residues identical to those of Hg-VAP-1 or Hg-VAP-2 and those that are similar are shaded. Residues conserved in all proteins are marked with a black shadow.

Hg-VAP-1 prior to hatching (Fig. 5A). The hybridisation signal for both *hg-vap-1* and *hg-vap-2* was consistently strong within the subventral gland cells in the preparasitic and parasitic second-stage juveniles (Figs. 5B and 6A) and usually weak within the gland cells in late parasitic stages (Fig. 5C and 6B). The expression profiles for both *hg-vap-1*

and *hg-vap-2* mirror the degeneration of the subventral gland cells in *H. glycines* during the late stages of parasitism. The high stringency conditions used during the in situ hybridisation and the relatively low identity between the nucleotide sequences of the two DNA probes precludes that both probes were hybridising with the same transcripts.

Table 1

Percent identity and similarity of *Heterodera glycines* venom allergen-like proteins (Hg-VAP-1 and Hg-VAP-2) to other VAPs from different nematode species

Protein ^a	Identity/Similarity	
	Hg-VAP-1	Hg-VAP-2
Hg-VAP-2	45/58	
Mi-MSP-1	30/44	37/49
C39E9.2	39/54	38/48
F49E11.6	40/55	38/50
Ac-ASP-1 (C-terminal region)	35/50	35/48
Ac-ASP-2	34/53	32/51
VAAP-1	31/47	32/43

^a Nematode species and protein accession numbers are provided in the legend to Fig. 4. Amino acid groups considered similar in the similarity calculations in BLASTP were as follows: (S, T); (A, S); (F, Y, W); (L, V, I, M); (K, R); (Q, E, K); (R, Q); (N, S); (N, D); (K, D).

4. Discussion

The most evolutionary advanced adaptations for plant parasitism by nematodes are the products of parasitism genes expressed in the oesophageal gland cells and secreted through the stylet (feeding structure) into host tissues. These stylet secretions have a direct role in the infection and parasitism of plants, and temporal changes in the secretory proteins occur during the parasitic cycle (Davis et al., 2000). Understanding what makes a nematode a plant parasite will require characterising the different extracellular (secreted) proteins produced in the oesophageal gland cells during the parasitic cycle. In the present study, we identified two secretory venom allergen-like protein genes that are specifically expressed within the subventral oesophageal gland cells of unhatched and hatched second-stage juveniles and in the early parasitic stages of the soybean cyst nematode, *H. glycines*. The subventral gland cells are the most active oesophageal glands in infective second-stage juveniles of *H. glycines* and after the initial stages of parasitism, the subventral gland cells decrease in transcriptional activity to become non-functional in the adult female (de Boer et al., 1999). This is the first report of genes encoding venom allergen-like proteins being expressed in the oesophageal gland cells of a nematode.

Two types of venom allergen-like proteins occur in nematodes: a short single domain type (~220 amino acids, e.g., Ac-ASP-2) and a longer 2 domain type (~425 amino acids, e.g., Ac-ASP-1) (Bin et al., 1999; Hawdon et al., 1999). Hg-VAP-1 and Hg-VAP-2 are representative of the single domain venom allergen-like protein, the most common type found in nematodes. For example, only two members of the over 20 member venom allergen-like protein family in *C. elegans* are of the two domain type (Sutherland, M., Westlund, B., Burnam, L., Sluder, A., Liu, L. 2001. Characterization of two *Caenorhabditis elegans* venom allergen-

like proteins. Keystone Symposia Abstract, Taos, NM. January 20–25, 2001).

The biological functions of extracellular venom allergen-like proteins in nematodes remain obscure. The release of Ac-ASP-1 by infective third-stage larvae of *A. caninum* that were stimulated to feed in vitro, suggests that Ac-ASP-1 might have a role in the establishment of parasitism (Hawdon et al., 1996). The apparent production of an ES venom allergen-like protein in the oesophagus of adult *H. contortus* suggests that this protein may be secreted during feeding (Schallig et al., 1997). The expression analysis of

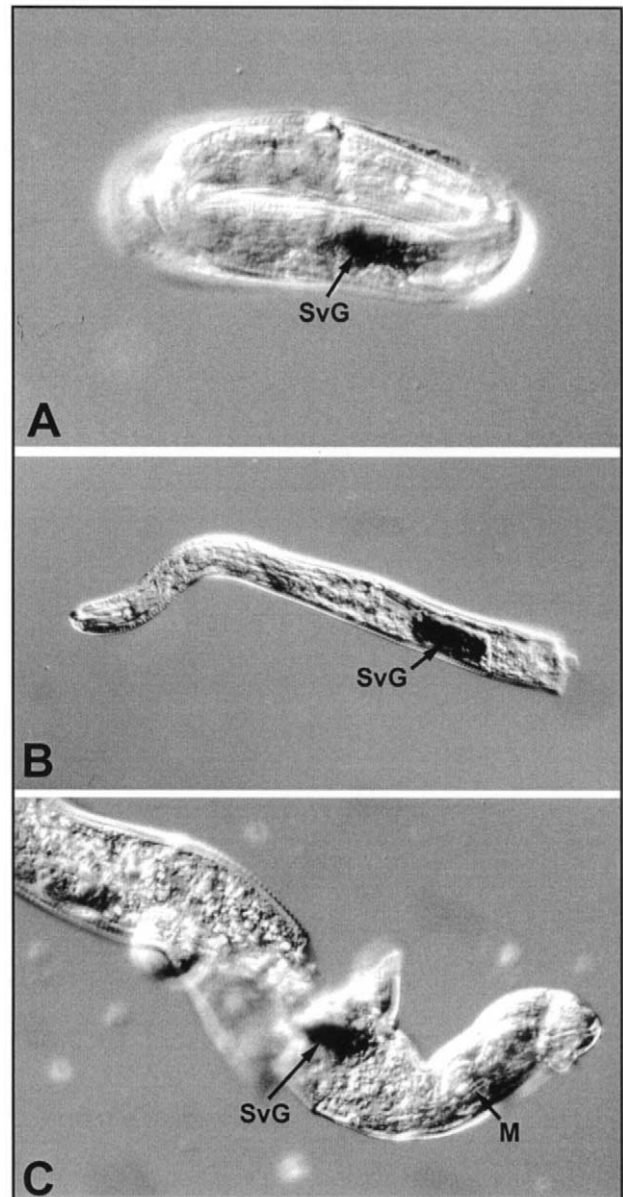


Fig. 5. Hybridisation of a digoxigenin-labelled antisense cDNA probe (dark staining) of *hg-vap-1* to transcripts expressed exclusively within the subventral oesophageal gland cells of *Heterodera glycines*. (A) Second-stage juvenile in an egg. (B) Preparasitic second-stage juvenile. (C) Late parasitic stage. SvG, subventral gland cells; M, metacorpus.

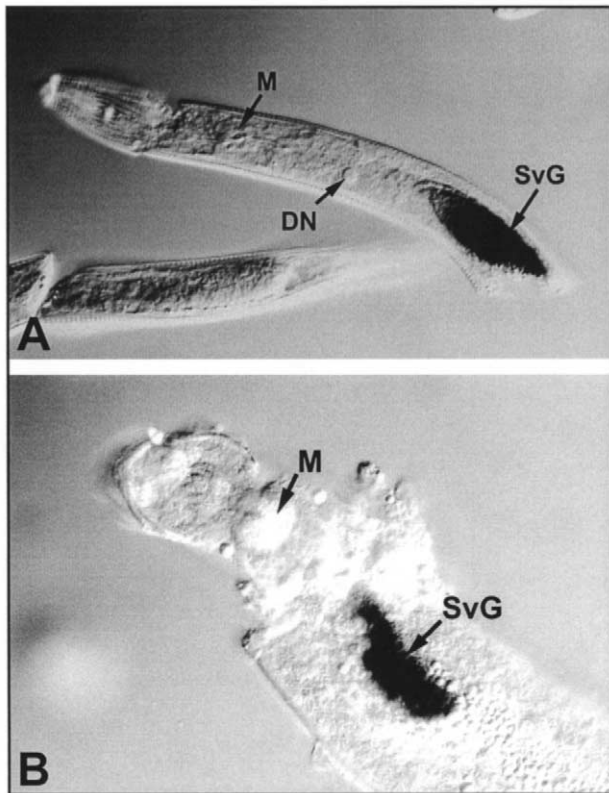


Fig. 6. Hybridisation of a digoxigenin-labelled antisense cDNA probe (dark staining) of *hg-vap-2* to transcripts expressed exclusively within the subventral oesophageal gland cells of *Heterodera glycines*. (A) Preparasitic second-stage juvenile. (B) Late parasitic stage. SvG, subventral gland cells; M, metacorpus; DN, dorsal gland cell nucleus.

two 2-domain *vap* genes identified in *C. elegans* showed expression of one gene in the amphidial sheath cells and the other gene in the excretory duct cell (A. Sluder, personal communication). Furthermore, in immunofluorescence assays, polyclonal antibodies to rAc-ASP-1 bound specifically to the amphids of second-stage juveniles of *M. incognita* (Hussey and Hotez, unpublished results). In the current study, however, the in situ localisation of *hg-vap-1* and *hg-vap-2* mRNA in the subventral gland cells suggests that their encoded secreted products may have a role in the interaction of *H. glycines* with its host plants.

The first parasitism genes encoding oesophageal gland cell secretory proteins to be cloned from plant-parasitic nematodes were β -1,4-endoglucanase (cellulase) genes from cyst nematodes (Smant et al., 1998; Yan et al., 1998). These cellulase genes are developmentally expressed in the subventral gland cells in *H. glycines* (de Boer et al. 1999) and one cellulase has been shown to be secreted within soybean roots by infective juveniles during the intracellular migration phase of the infection process (Wang et al., 1999). Although other putative parasitism genes have been cloned and characterised from plant-parasitic nematodes, the functions of their encoded secretory products in parasitism of plants remain unknown (Davis et al., 2000).

Since membrane-associated proteins also have signal peptides, the presence of a putative signal peptide alone is not conclusive proof that a protein functions as an extracellular protein. However, PSORT II analysis (Nakai and Horton, 1999) predicted Hg-VAP-1 and Hg-VAP-2 to be extracellular and, thus, both proteins should be considered to have potential roles in *H. glycines* parasitism of plants. Further demonstration of a role for Hg-VAP-1 and Hg-VAP-2 in plant parasitism will require immunolocalisation of the extracellular proteins in secretory granules in the nematode and evidence of their secretion by *H. glycines* in planta.

Direct molecular analysis of genes encoding secretory proteins expressed in the oesophageal gland cells should provide a complete profile of parasitism genes expressed in the gland cells during the parasitic cycle and rapidly advance our understanding of the molecular interactions during nematode parasitism of plants. In our investigation, *hg-vap-1* and *hg-vap-2* were cloned by directly microaspirating the oesophageal gland cell cytoplasm from different parasitic stages of *H. glycines* and generating cDNA libraries.

Acknowledgements

Support for this research was provided by the United Soybean Board (Project No. 9214), the Iowa Soybean Promotion Board, the Iowa Agriculture and Home Economics Experiment Station (J-19492; Project No. 3381), by North Carolina Agricultural Research Service, by Hatch Act and State of Iowa, and by state and Hatch Funds allocated to the Georgia Agricultural Experiment Stations.

References

- Benson, D.A., Boguski, M.S., Lipman, D.J., Ostell, J., Ouellette, B.F.F., 1998. GenBank. Nucleic Acids Res. 26, 1–7.
- Bin, A., Hawdon, J., Qiang, S., Hainan, R., Huiqing, Q., Wei, H., Shu-hua, X., Tiehua, L., Xing, G., Zheng, F., Hotez, P., 1999. *Ancylostoma* secreted protein 1 (ASP-1) homologues in human hookworms. Mol. Biochem. Parasitol. 99, 143–9.
- Blumenthal, T., Steward, K., 1997. RNA processing and gene structure. In: Riddle, D.L., Blumenthal, T., Meyer, B.J., Priess, J.R. (Eds.). *C. elegans* II, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 117–46.
- Davis, E.L., Hussey, R.S., Baum, T.J., Bakker, J., Schots, A., Rosso, M.-N., Abad, P., 2000. Nematode parasitism genes. Annu. Rev. Phytopathol. 38, 365–96.
- De Boer, J.M., Yan, Y., Smant, G., Davis, E.L., Baum, T.J., 1998. In-situ hybridization to messenger RNA in *Heterodera glycines*. J. Nematol. 30, 309–12.
- De Boer, J.M., Yan, Y., Wang, X., Smant, G., Hussey, R.S., Davis, E.L., Baum, T.J., 1999. Developmental expression of secretory β -1,4-endoglucanases in the subventral esophageal glands of *Heterodera glycines*. Mol. Plant Microbe Interact. 12, 663–9.
- Diatchenko, L., Lau, Y.-F.C., Campbell, A.P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E.D.,

- Siebert, P.D., 1996. Suppression subtractive hybridization: A method for generating differentially regulated tissue-specific cDNA probes and libraries. *Proc. Natl. Acad. Sci. U.S.A.* 93, 6025–30.
- Ding, X., Shields, J., Allen, R., Hussey, R.S., 2000. Molecular cloning and characterization of a venom allergen AG5-like cDNA from *Meloidogyne incognita*. *Int. J. Parasitol.* 30, 77–81.
- Fang, K.S.Y., Vitale, M., Fehlner, P., King, T.P., 1988. CDNA cloning and primary structure of a white-faced hornet venom allergen, antigen 5. *Proc. Natl. Acad. Sci. U.S.A.* 85, 895–9.
- Gao, B., Allen, R., Maier, T., Davis, E.L., Baum, T.J., Hussey, R.S., 2001. Identification of putative parasitism genes expressed in the esophageal gland cells of the soybean cyst nematode, *Heterodera glycines*. *Mol. Plant-Microbe Interact.*, 14 In Press.
- Gill, S.C., von Hippel, P.H., 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182, 319–26.
- Hawdon, J.M., Jones, B.F., Hoffman, D.R., Hotez, P.J., 1996. Cloning and characterization of *Ancylostoma*-secreted protein. *J. Biol. Chem.* 271, 6672–8.
- Hawdon, J.M., Narasimhan, S., Hotez, P.J., 1999. *Ancylostoma* secreted protein 2: Cloning and characterization of a second member of a family of nematode secreted proteins from *Ancylostoma caninum*. *Mol. Biochem. Parasitol.* 99, 149–66.
- Huang, X., 1994. On global sequence alignment. *Computer Appl. Biosci.* 10, 227–35.
- Hussey, R.S., Grundler, F.M., 1998. Nematode parasitism of plants. In: Perry, R.N., Wright, D.J. (Eds.). *Physiology and biochemistry of free-living and plant parasitic nematodes*, CAB International Press, England, pp. 213–43.
- King, T.P., Moran, D., Wang, D.F., Kochoumian, L., Chait, B.T., 1990. Structural studies of a hornet venom allergen antigen 5, *Dol m V* and its sequence similarity with other proteins. *Protein Seq. Data Anal.* 3, 263–6.
- Nakai, K., Horton, P., 1999. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem. Sci.* 24, 34–35.
- Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* 10, 1–6.
- Ray, C., Abbott, A.G., Hussey, R.S., 1994. *Trans*-splicing of a *Meloidogyne incognita* mRNA encoding a putative esophageal gland protein. *Mol. Biochem. Parasitol.* 68, 93–101.
- Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. *Molecular cloning: a laboratory manual*. 2nd ed, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Schallig, H.D.F.H., van Leeuwen, M.A.W., Verstrepen, B.E., Cornelissen, A.W.C.A., 1997. Molecular characterization and expression of two putative protective excretory secretory proteins of *Haemonchus contortus*. *Mol. Biochem. Parasitol.* 88, 203–13.
- Smant, G., Stokkermans, J.P.W.G., Yan, Y., de Boer, J.M., Baum, T.J., Wang, X., Hussey, R.S., Gommers, F.J., Henrissat, B., Davis, E.L., Helder, J., Shots, A., Bakker, J., 1998. Endogenous cellulases in animals: isolation of β -1,4-endoglucanase genes from two species of plant-parasitic cyst nematodes. *Proc. Natl. Acad. Sci. U.S.A.* 95, 4906–49011.
- Tawe, W., Pearlman, E., Unnasch, T.R., Lustigman, S., 2000. Angiogenic activity of *Onchocerca volvulus* recombinant proteins similar to vespinal venom antigen 5. *Mol. Biochem. Parasitol.* 109, 91–99.
- Wang, X., Allen, R., Ding, X., Goellner, M., Maier, T., de Boer, J., Baum, T., Hussey, R.S., Davis, E.L., 2001. Signal peptide-selection of cDNA cloned directly from the esophageal gland cells of the soybean cyst nematode *Heterodera glycines*. *Mol. Plant-Microbe Interact.* 14, 536–44.
- Wang, X., Meyers, D., Yan, Y., Baum, T., Smant, G., Hussey, R., Davis, E., 1999. In planta localization of a β -1,4-endoglucanase secreted by *Heterodera glycines*. *Mol. Plant-Microbe Interact.* 12, 64–67.
- Yan, Y., Smant, G., Stokkermans, J., Qin, L., Helder, J., Baum, T., Schots, A., Davis, E., 1998. Genomic organization of four β -1,4-endoglucanase genes in plant-parasitic cyst nematodes and its evolutionary implications. *Gene* 220, 61–70.