

Cloning and characterization of a venom allergen-like protein gene cluster from the pinewood nematode *Bursaphelenchus xylophilus*

Shifeng Lin, Heng Jian*, Haijuan Zhao, Dan Yang, Qian Liu

Department of Plant Pathology, China Agricultural University, Beijing 100193, China

ARTICLE INFO

Article history:

Received 15 March 2010

Received in revised form 8 October 2010

Accepted 13 October 2010

Keywords:

Venom allergen-like protein

Gene cluster

Bursaphelenchus xylophilus

Esophageal gland cells

Secretions

Parasitism genes

ABSTRACT

Pinewood nematode (PWN) is the causal agent of the pine wilt disease. Previous studies have suggested that secretions from the esophageal glands of PWN play an important role in pathogenicity. A cluster of three venom allergen-like protein genes and one pseudogene, *Bx-vap-1*, *Bx-vap-2*, *Bx-vap-3* and *Bx-vap-P*, were identified within a 3.7-kb region. Additionally, three putative modification, transport and regulatory protein genes were also detected in the same flanking region of the *Bx-vap* gene cluster. Genes *vap-1*, *-2* and *-3* are functional and encode three major allelic variants of PWN venom allergen-like proteins. But *Bx-vap-P* is an untranscribed pseudogene. Genes *vap-1*, *-2* and *-3* produce predicted products of 204, 206 and 203 amino acid residues, respectively, including the putative signal peptide sequence at the amino termini. In situ mRNA hybridization analysis showed that the transcripts of genes *vap-1*, *-2* and *-3* accumulated exclusively within the esophageal gland cells of *Bursaphelenchus xylophilus*.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

The pinewood nematode (PWN), *Bursaphelenchus xylophilus*, is the causal agent of pine wilt disease, which causes great devastation to many conifer trees. The PWN has spread from its native continent of North America (Rutherford et al., 1990) to Japan (Mamiya, 1988), Korea (Yi et al., 1989) and China (Cheng et al., 1983) in Asia, and Portugal in Europe (Mota et al., 1999); the damage is most extensive in its non-native geographic range. It is the most severe in Japan and China, where the pine wilt disease not only caused substantial economic losses to pine species without resistance to the nematode, but also has seriously damaged the landscape and environment by destroying entire forests. Despite the damage caused by *B. xylophilus* throughout the world, our understanding of the mechanisms of pathogenicity by *B. xylophilus* remains sketchy, especially at the molecular level. An expressed sequence tag (EST) project on *B. xylophilus* (Kikuchi et al., 2007) has been carried out, leading to the discovery of several genes encoding cell-wall-degrading enzymes, such as cellulases (endo- β -1,4-glucanase), β -1,3-glucanases, pectate lyase and expansin-like protein (Kikuchi et al., 2004, 2005, 2006, 2009). Recently, two subtractive *B. xylophilus* cDNA libraries specific to the dispersal 4th larval stage (D4S) and pine-grown propagative mixed stage (PGPS) were constructed, and some genes specifically related to the host–nematode interaction were cloned (Kang et al., 2009).

Venom allergen-like proteins (VAP) are a class of cysteine-rich secretory proteins (CRISPs) found in both plants and animals. The first characterized nematode member of this protein family was Ac-ASP-1, which is the major protein released only by infective L3 of *Ancylostoma caninum* activated to feed in vitro and probably plays a role in infection and the transition to parasitism (Hawdon et al., 1996). Another venom allergen-like protein identified from adult *Haemonchus contortus* is produced in the esophagus, suggesting that this protein may be secreted during feeding (Schallig et al., 1997). In plant–parasitic nematodes, venom allergen-like protein genes were isolated from *Meloidogyne incognita* (Ding et al., 2000; Wang et al., 2007), *Heterodera glycines* (Gao et al., 2001) and *Ditylenchus africanus* (Haegeman et al., 2009). Because all such transcripts accumulated exclusively within the subventral esophageal gland cells of these parasites, these venom allergen-like protein genes may play an important role in the infection of host plants by these parasites.

Secretory proteins expressed in the esophageal glands of plant–parasitic nematodes play an essential role in nematode infection and parasitism of plants (Hussey, 1989). Cloning and characterization of parasitism genes is the key to understanding the molecular basis of nematode parasitism of plants. To clone the gland genes, we have initiated a project to identify esophageal-gland genes that encode secreted proteins. A specific cDNA library was constructed from the esophageal gland region of *B. xylophilus* by modified solid-phase subtractive hybridization, and two cDNA clones similar to venom allergen-like proteins were identified in the *B. xylophilus* gland-region-specific cDNA library. Southern analysis and

* Corresponding author. Fax: +86 10 62731102.

E-mail address: hengjian@cau.edu.cn (H. Jian).

nucleotide sequence analyses revealed that the two venom allergen-like protein genes were tandemly arranged in genomic DNA of *B. xylophilus* and that their flanking sequences contain another venom allergen-like protein gene and a pseudogene. In this paper, we report the isolation and molecular analysis of full-length cDNA and genomic clones of the four *B. xylophilus* venom allergen-like protein genes, the localization of their transcripts and the gene expression profiles in the life cycle of *B. xylophilus*. Additionally, we also describe the detection of three modification, transport and regulatory protein genes in the same flanking region of the *Bx-vap* gene cluster.

2. Materials and methods

2.1. Biological material

Both *B. xylophilus* and *Botrytis cinerea* were kindly provided by Prof. Liu Xingzhong, Institute of Microbiology, Chinese Academy of Sciences. *B. xylophilus* was cultured on *B. cinerea* grown on potato dextrose agar (PDA) medium, extracted from *B. cinerea* hyphae with Baermann funnels for 12 h at 25 °C, and then washed 5 times in M9 buffer to remove remaining *B. cinerea* mycelium.

2.2. Isolation of two venom allergen-like protein genes

Based on the method previously described (Lambert et al., 1999), 500 nematodes at different parasitic stages were cut into gland regions (one fourth of body length from the anterior end of the nematode) and tail regions (a half of body length from the posterior end of the nematode) by eye scalpel under stereomicroscope. A gland-region-specific cDNA library was constructed by the modified solid-phase subtractive hybridization method described in detail elsewhere (Hu et al., 2006). In brief, subtractive hybridization (three rounds) was performed by using second-strand cDNA derived from gland regions of mixed-stage *B. xylophilus* as the 'tester' and first-strand cDNA derived from tail regions of mixed-stage *B. xylophilus* as the 'driver', and a library was constructed from this subtracted cDNA using pGEM-T easy vector (Promega, Madison, WI, USA) in *Escherichia coli* DH5 α . Two different venom allergen-like protein genes, designated *Bx-vap-1* and *Bx-vap-2*, were identified from clones B056 and B083 during random sequencing of clones from the gland-region-specific cDNA library. The full-length cDNA sequences of *Bx-vap-1* and 2 were subsequently obtained by 5'-rapid amplification of cDNA ends (5'-RACE) using the first-strand cDNA generated from mixed-stage *B. xylophilus* and the reverse gene-specific primers VAP1r1/VAP1r2 and VAP2r1/VAP2r2 with the BD SMART RACE cDNA amplification kit (Clontech, Palo Alto, CA, USA) according to the manufacturer's instructions. The amplified PCR products were cloned into pGEM-T easy vector (Promega, Madison, WI, USA) for sequencing and delineation of the full-length *Bx-vap-1* and 2 cDNAs.

2.3. Southern blot hybridization

Ten micrograms of genomic DNA isolated from *B. xylophilus* was digested respectively with *EcoR* I and *BamH* I. Genomic DNA isolated from *B. cinerea* was used as the control. The DNA fragments were separated on a 0.8% agarose gel and transferred onto a Hybond N+ nylon membrane (Roche Applied Science, Indianapolis, IN, USA). Hybridization probes were made by amplification of full-length cDNA from the original plasmid of *vap-1* or -2 using pairs of gene-specific primers (VAP1s/VAP1a or VAP2s/VAP2a). PCR products were purified with the Gel Extraction Mini Kit (Watson BioTechnologies Inc., Shanghai, China) and then random

primed labeled with digoxigenin using the DIG High Prime DNA Labeling and Detection Starter Kit I (Roche Applied Science). The probes were hybridized to the membrane at 42 °C overnight, and two high stringency washes with 0.5 $\times$ SSC/0.1% SDS solution at 68 °C were performed according to the standard procedure (Sambrook et al., 1989). The hybridization signals were detected with the DIG High Prime DNA Labeling and Detection Starter Kit I (Roche Applied Science).

2.4. Amplification of the VAP gene cluster and flanking regions

To determine whether *Bx-vap-1* and -2 are mutually linked, a regular PCR strategy with various primer combinations was used to amplify the intergenic region. A PCR fragment spanning *Bx-vap-1* and -2 with the primer combination VAP2s1 and VAP1a1 was obtained and sequenced. To obtain the nucleotide sequence flanking the DNA fragment, long distance-inverse PCR (LD-IPCR) was performed. The genomic DNA of *B. xylophilus* was completely digested by either *BamH* I or *EcoR* I, purified and self-ligated with T4 DNA ligase (Promega) at 4 °C for 48 h. Then the circular fragments were used as the templates for LD-IPCR with LA Taq (TaKaRa, Otsu, Japan) using the relevant inverse primer combination VAP1s1 and VAP2a1, thereby amplifying the flanking sequences adjacent to the known gDNA fragments of the two venom allergen-like protein genes. The LD-IPCR products were inserted into a pGEM-T easy vector for sequencing.

Sequence analyses revealed that the flanking regions of *Bx-vap-1* and -2 contain one entire sequence (designated *Bx-vap-3*) encoding one venom allergen-like protein, one truncated sequence (designated *Bx-vap-P*) similar to the venom allergen-like protein gene, and three entire sequences (designated *Bx-abc-1*, *Bx-rbp-1* and *Bx-omt-1*) potentially encoding three different regulatory/tailoring proteins. To determine whether *Bx-vap-P* is expressed, we designed a specific primer pair to recognize *Bx-vap-P* by RT-PCR. Subsequently, the full-length cDNA sequences of *vap-3*, *abc-1*, *rbp-1* and *omt-1* were obtained by 3'-RACE and 5'-RACE using the first-strand cDNA generated from mixed-stage *B. xylophilus* and the gene-specific primers with the BD SMART RACE cDNA amplification kit (Clontech) according to the manufacturer's instructions. The amplified PCR products were cloned into a pGEM-T easy vector for sequencing and delineation of the full-length cDNAs. All oligonucleotide primers (Table 1) used in this study were synthesized by Invitrogen Biotechnology Co. Ltd. (Shanghai, China).

2.5. Sequence analyses, alignment, and phylogenetics

Sequence similarity comparisons were carried out by BLASTP in NCBI (Altschul et al., 1997). Signal peptide predictions were made using SignalP (Nielsen et al., 1997). Protein theoretical isoelectric points and molecular weights were predicted using ProtParam (Gill and von Hippel, 1989), and motifs were searched with Motifscan (Falquet et al., 2002). Sequence alignments were performed with the DNAMAN software (version 6.0; Lynnon Biosoft, USA) using the default settings and manually modified. Phylogenetic and molecular evolutionary analyses were conducted using the neighbor-joining method with bootstrap of 1000 times and poisson correction in MEGA software (version 3.1) (Kumar et al., 2004).

2.6. In situ hybridization

For in situ hybridization, the DNA fragment used as the probe was amplified from the full-length cDNA clones of *Bx-vap-1*, -2 and -3 with the specific primer pairs VAP1s2/VAP1a2, VAP2s2/VAP2a2 and VAP3s2/VAP3a2. The digoxigenin (DIG)-labeled sense

Table 1
Primers used in this research.

Primer name	Sequence (5' → 3')
BACTs	TCTCTTCCAGCCTTCTTCTTG
BACTa	GGCCGGATTCTGCTGACT
VAP1r1	TGGCGTCCCCAGAATAACGA
VAP1r2	CTCATTCCACCACGTTGAGGA
VAP2r1	GGCGTCCCCGAATAATCC
VAP2r2	CATCCACCACATTGAGCG
VAP1s1	GATCATCAATCAAATGTTCCGTC
VAP1a1	AAGCACTGAATTTATTTGCAACAGT
VAP2s1	GGCTTGAAAAATGGTTCGAG
VAP2a1	GAAGCGCTGGATTATTTTAAATATA
VAP3f1	CTGCGCCTATCATCACCCCT
VAP3f2	GGAACCCGCATCTCGAATACTC
VAP3r1	TCGACACGTGTCAGATAGTTGTTGG
VAP3r2	GCCAGATTCTGGCCGTAAGGAG
VAP1s2	GACGAATGGGCTCAGGCC
VAP1a2	CATTGAATCAGGCGCTACACAG
VAP2s2	GCGAATCTTGCCCAATGGAGAG
VAP2a2	TCAGTCAAGATCCTTGGCTGCT
VAP3s2	CAGAGGCCAGGTTAAGCGACGATG
VAP3a2	GAGCAGGATCCCCGGTCTGA
ABC1f1	TTGTGGAGCCGAGGATTGTG
ABC1f2	CGAGAATTCAGGAGGCAAGGG
ABC1r1	AAGAATACCTCAGTTTCGGCGG
ABC1r2	CATTCTGTAAGCGGAGTAGC
OMT1f1	AAGCGGCATTGGGAGGATAAC
OMT1f2	TGGACATGGTGGATGAACGG
OMT1r1	GAATATGCAACTTCTCAAAGCCTCC
OMT1r2	CGGATGCAGTTTTCGCCC
RBP1f1	AGAAGGTTGTGCACTGTCCC
RBP1f2	GCCGAAGCTGATCAGTTGAAGG
RBP1r1	ATAGGGACAGGTGCACAACCTTC
RBP1r2	TCTTCATCAAGGCTTCTGGCAC

and anti-sense cDNA probes were synthesized from PCR products with the PCR DIG Probe Synthesis Kit (Roche Applied Science) by asymmetric PCR (Wang et al., 2001). In situ hybridization was performed with mixed-stage *B. xylophilus* as described previously (De Boer et al., 1998), except hybridization and detection were performed with the DIG High Prime DNA Labeling and Detection Starter Kit I (Roche Applied Science).

2.7. Developmental stage expression analysis

Transcripts of *vap-1*, *-2* and *-3* were analyzed at different developmental stages of *B. xylophilus* by reverse transcription (RT)-PCR with the same set of gene-specific primers used for in situ hybridizations. Approximately 20,000 differentiated eggs (first-stage and unhatched second-stage juvenile), 2000 hatched second-stage juveniles, 500 adult females and 500 males were manually separated from a batch of freshly collected nematodes cultured on fungi. Poly(A)⁺ RNA (mRNA) was isolated using the Dynabeads mRNA[®] DIRECT[™] Micro kit (Invitrogen, Carlsbad, CA, USA), and first-strand cDNA was synthesized with SuperScript III Reverse Transcriptase (Invitrogen) and oligo(dT) primers (Sangon, Shanghai, China) in a volume of 20 µl at 50 °C for 60 min and at 70 °C for 15 min. Amplification reactions were performed in a volume of 25 µl containing 1 µl cDNA or genomic DNA as template, 0.4 µM of each primer, 0.625 U of Taq polymerase (TaKaRa, Dalian, China), 200 µM of each dNTP and 1× PCR buffer (TaKaRa). PCR conditions were as follows: 3 min at 94 °C, 25–30 cycles at 94 °C for 30 s, 55 °C for 30 s, 72 °C for 30 s, and a final elongation of 10 min at 72 °C. Additionally, the constitutively expressed β-actin gene of *B. xylophilus* (GenBank accession number EU100952) used as controls was amplified from each sample with primers BACTs and BACTa. PCR products were separated on 1.5% agarose gels and stained with ethidium bromide.

3. Results

3.1. Identification and cloning of the *B. xylophilus* venom allergen-like protein gene cluster

The *B. xylophilus* gland-region-specific cDNA library produced from subtractive hybridization yielded two different clones, B056 and B083, with significant similarity to venom allergen-like proteins from *H. glycines* (GenBank accession Nos. AY033601 and AF013289). Based on these sequences, reverse gene-specific primers were designed for 5'-RACE PCR that permitted the successful cloning of two full-length cDNAs, designated as *Bx-vap-1* and *Bx-vap-2* (GenBank accession No. GQ272489 and GQ272490, respectively), encoding two different venom allergen-like proteins.

Southern blots confirmed the nematode origin of *Bx-vap-1* and *-2* (Fig. 1). The *Bx-vap-1* and *-2* cDNAs were DIG-labeled and used as probes for a DNA gel blot containing *EcoR* I- and *Bam*HI-digested DNA from *B. xylophilus*. As a negative control, the blot contained genomic DNA from the fungus (*B. cinerea*) on which the nematodes were reared. The *vap-1* probe resulted in a single hybridizable band from the *EcoR* I and *Bam*HI digestions of *B. xylophilus* DNA (lanes N1 and N2 in Fig. 1). When the *Bx-vap-2* probe was used, the same-size single band that appeared in hybridization with the *Bx-vap-1* probe was also detected in each lane (data not shown), which suggests that *Bx-vap-1* and *-2* were located close to each other in the genome of *B. xylophilus*.

The amplification of the complete sequence (2.1 kb) of *Bx-vap-1* and *-2* along with the intergenic region confirmed that they were closely linked and tandemly arranged on the same chromosome of *B. xylophilus* (Fig. 2). Additionally, the 10-kb *Bam*HI restriction fragments and the 11-kb *EcoR* I restriction fragments (Fig. 1, lanes N1 and N2) were used as templates to amplify regions flanking the *Bx-vap-1* and *-2* genes by LD-IPCR. This yielded a 3.4-kb region downstream of *vap-1* which, when sequenced, contained two other sequences similar to the venom allergen-like protein gene (Fig. 2). The first one, *Bx-vap-P* (GenBank accession No. GU451045), is oriented in the same direction as *Bx-vap-1* and *2*, while the second one, *Bx-vap-3* (GenBank accession No. GQ272491), is oriented in the opposite direction. The distance between *Bx-vap-2* and

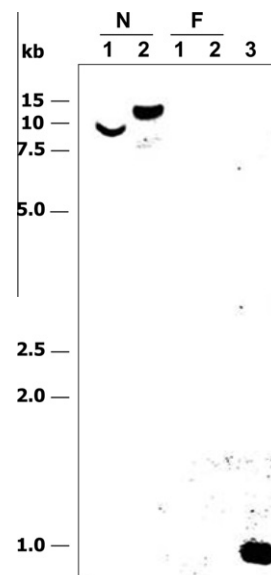


Fig. 1. Southern blot analysis of *Bx-vap-1*. Blot containing genomic DNAs from *Bursaphelenchus xylophilus* (N) and fungus *Botrytis cinerea* (F) digested with *Bam*HI (1) or *EcoR* I, (2) and PCR products of *Bx-vap-1* (3) were hybridized with a *Bx-vap-1* cDNA probe. DNA size markers are indicated in kilobases.

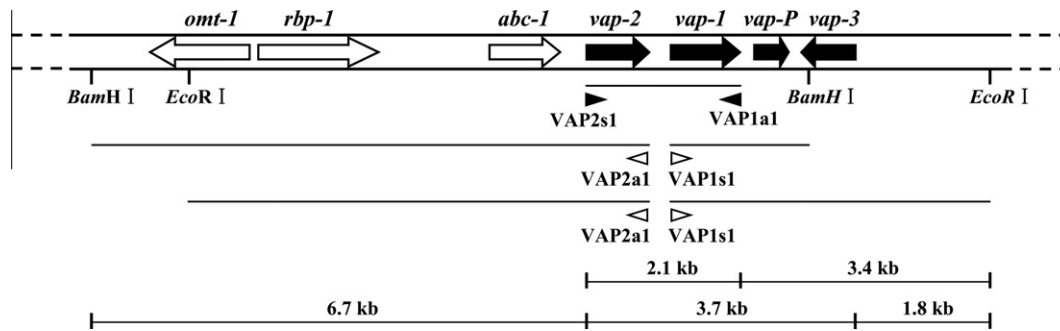


Fig. 2. The *vap* gene cluster and flanking regions in *Bursaphelenchus xylophilus*. The arrows indicate the position, orientation and relative lengths of each gene. The numbers representing the lengths (kb) of the *vap* gene cluster and its flanking regions were determined by LD-IPCR. The restriction sites (*Bam*H I and *Eco*R I) used for LD-IPCR are marked on the chromosome, and the positions of the primers used for LD-IPCR, and to define the junction of *Bx-vap-1* and -2, are indicated by open and closed arrowheads.

Bx-vap-1 is 278 bp; between *Bx-vap-1* and *Bx-vap-P*, 240 bp; and between *Bx-vap-P* and *Bx-vap-3*, 112 bp. In addition, a 6.7-kb region was sequenced upstream of *Bx-vap-2* and three entire sequences were found (Fig. 2). Two of them, *Bx-abc-1* (GenBank accession No. GU451048) and *Bx-rbp-1* (GenBank accession No. GU451050), which encode proteins with some similarity to ATP-binding cassette (ABC) transporter and RNA-binding protein, respectively, are transcribed in the same direction as *Bx-vap-1* and -2, while another sequence, *Bx-omt-1* (GenBank accession No. GU451049), which encodes a protein with some similarity to O-methyltransferase, is transcribed in the opposite direction as *Bx-vap-1* and -2. Based on the sequence data, gene-specific primers were designed for 3'-RACE and 5'-RACE PCR that permitted the successful cloning of four full-length cDNAs encoding BX-VAP-3, BX-ABC-1, BX-RBP-1 and BX-OMT-1.

3.2. Cloning and sequence analysis of *Bx-vap-1/2/3* cDNAs

The putative full-length *Bx-vap-1* cDNA contained 722 nucleotides (excluding the poly (dA) tail) with a putative open reading frame (ORF) of 615 bp. The cDNA contained 14 bp of 5'-untranslated region (5'-UTR) and 93 bp of 3'-UTR, which contained a polyadenylation signal (AATAAA). The ORF encoded a deduced protein of 204 amino acids with a theoretical molecular mass of 22434.86 Da and a pI of 4.93, predicted by ProtParam (Gill and von Hippel, 1989). A putative signal peptide predicted by the SignalP program (Nielsen et al., 1997) terminated immediately upstream of a protease cleavage site between amino acids Ala17 and Asp18. A motif search (Falquet et al., 2002) identified a possible N-glycosylation site at amino acid Asn150. The *Bx-vap-1* cDNA did not contain a nematode spliced leader sequence at the 5'-end (Ray et al., 1994).

The putative full-length *Bx-vap-2* cDNA contained 711 nucleotides (excluding the poly (dA) tail) with a putative ORF of 621 bp. The cDNA contained 20 bp of 5'-UTR and 70 bp of 3'-UTR, which contained a polyadenylation signal (AATAAA). The ORF encoded a deduced protein of 206 amino acids with a theoretical molecular mass of 22418.83 Da and a pI of 4.68. A putative signal peptide terminated immediately upstream of a protease cleavage site between amino acids Ala17 and Thr18. A motif search identified three possible N-glycosylation sites at amino acids Asn48, Asn105 and Asn150. A nematode spliced leader sequence was not present at the 5'-end of the *Bx-vap-2* cDNA.

The putative full-length cDNA designated as *Bx-vap-3* contained 691 nucleotides (excluding the poly (dA) tail) with a putative 612-bp ORF. The cDNA contained 10 bp of 5'-UTR and 69 bp of 3'-UTR containing a polyadenylation signal (AATAAA). The ORF encoded a deduced protein of 203 amino acids with a theoretical molecular mass of 22446.27 Da and a pI of 4.34. A putative signal peptide

terminated immediately upstream of a protease cleavage site between amino acids Ala18 and Arg19. A motif search identified three possible N-glycosylation sites at amino acids Asn10, Asn160 and Asn196. There was no nematode trans-spliced leader sequence at the 5'-end of the *Bx-vap-3* cDNA.

3.3. Cloning and sequence analysis of *Bx-rbp-1/abc-1/omt-1* cDNAs

The putative full-length *Bx-rbp-1* cDNA contained 1085 nucleotides (excluding the poly (dA) tail) with a putative open reading frame (ORF) of 951 bp. The cDNA contained 57 bp of 5'-untranslated region (5'-UTR) and 77 bp of 3'-UTR. The ORF encoded a 316-residue protein containing an RNA recognition motif (RRM), which showed 60% identity (76% similarity) to an RNA-binding motif protein from *Brugia malayi*.

The putative full-length *Bx-abc-1* cDNA contained 787 nucleotides (excluding the poly (dA) tail) with a putative open reading frame (ORF) of 582 bp. The cDNA contained 57 bp of 5'-untranslated region (5'-UTR) and 148 bp of 3'-UTR. The ORF encoded a 193-residue protein containing a 27-residue putative signal peptide, which showed 28% identity (56% similarity) to an ABC transporter-related protein from *Streptococcus suis*.

The putative full-length *Bx-omt-1* cDNA contained 939 nucleotides (excluding the poly (dA) tail) with a putative open reading frame (ORF) of 747 bp. The cDNA contained 83 bp of 5'-untranslated region (5'-UTR) and 109 bp of 3'-UTR. The ORF encoded a 248-residue protein, which showed 52% identity (73% similarity) to a hydroxyindole-O-methyltransferase from *Caenorhabditis brenneri*.

3.4. Genomic structure of the *Bx-vap-1/2/3/P* genes

The lengths of the genomic DNA stretches containing *Bx-vap-1*, *Bx-vap-2*, *Bx-vap-3* were 988, 893 and 767 bp, respectively (Fig. 3). The exon/intron boundaries of the genomic sequences were determined by aligning the genomic sequences with the corresponding cDNA sequences. Both *Bx-vap-1* and *Bx-vap-2* contained one shorter intron 1 (76 bp and 74 bp) and one longer intron 2 (190 bp and 108 bp), respectively, while *Bx-vap-3* contained only one short intron 1 (76 bp). All sequences of introns followed the "GT-AG rule" for cis-splicing (Blumenthal and Steward, 1997).

A comparison of the nucleotide sequence of *Bx-vap-P* and other *vap* genes indicated that the length of the genomic DNA stretch of *Bx-vap-P* was 523 bp, which began at a matching position corresponding to position 67 of the 225 nucleotide-pair exon 2 in the *Bx-vap-1* or -2 gene, and ended at AATAAA corresponding to the polyadenylation signal (Fig. 3). *Bx-vap-P* contained one short sequence which corresponded to intron 2 of *Bx-vap-1* or -2, and the intron-like sequence used the conserved dinucleotide GT at

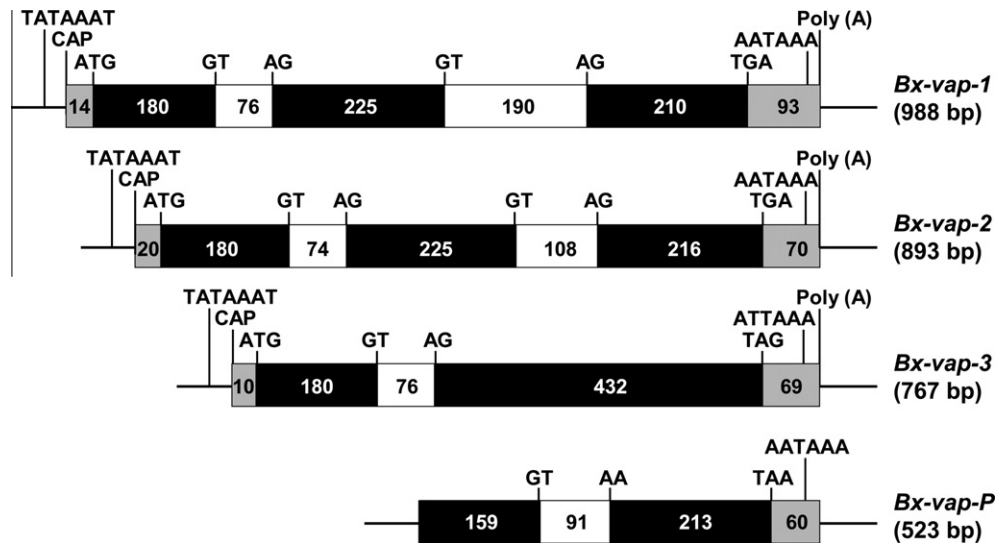


Fig. 3. Structural features of venom allergen-like protein genes *Bx-vap-1*, *Bx-vap-2*, *Bx-vap-3* and *Bx-vap-P*. 5'/3' UTRs are marked with grey boxes, exons (dark boxes), introns (open boxes). Conserved ATA boxes in the promoter region, dinucleotides at the splice junctions, and the AATAAA polyadenylation signal are marked on the top of each gene.

the 5' splice site and the mutative dinucleotide AA at the 3' splice site instead of AG. In the putative truncated reading frame, an in-frame TAG termination codon occurred at the extreme 3-end of its first exon sequence. Based on the sequence data, it was speculated that *Bx-vap-P* is a non-processed pseudogene in *B. xylophilus*. In addition, the RT-PCR experiment further indicated that *Bx-vap-P* is an untranscribed pseudogene, as transcription was not detected in mixed stages of *B. xylophilus* (data not shown).

3.5. Sequence comparison to other nematode VAP proteins

The degree of similarity between the proteins encoded by three complete genes, *Bx-vap-1*, *Bx-vap-2* and *Bx-vap-3*, the pseudogene *Bx-vap-P*, and other venom allergen-like proteins from *H. glycines* and *M. incognita* was determined. Comparison of the four BX-VAPs indicated that they are approximately 44–81% identical and 61–87% similar to each other, and the highest level of identity and similarity was found between BX-VAP-1 and BX-VAP-2. When compared with other venom allergen-like protein family members from selected nematode species such as *H. glycines* and *M. incognita*, BX-VAP-1, BX-VAP-2, BX-VAP-3 and BX-VAP-P were 25–42% identical and 38–53% similar, and the highest level of identity and similarity for BX-VAP-1, BX-VAP-2 and BX-VAP-3 was with the venom allergen-like protein from *H. glycines* (HG-VAP-2, GenBank accession No. AY033601) (Gao et al., 2001), while BX-VAP-P was more similar to the venom allergen-like protein from *M. incognita* (MI-MSP-1, GenBank accession No. AF013289) (Ding et al., 2000).

The *Bx-vap* gene sequences were then used to search the EST database in GenBank. One EST from *B. xylophilus* (GenBank accession No. CJ986258) is the *Bx-vap-1* cDNA, and another EST from *Bursaphelenchus mucronatus* (GenBank accession No. CJ976257) is similar to the *Bx-vap* gene sequences. Based on this, the full-length cDNA of a venom allergen-like gene from *B. mucronatus* was obtained by RACE-PCR. The *B. mucronatus* cDNA, designated *Bm-vap-1* (GenBank accession No. GU451046), contains a 621-bp open reading frame and encodes a 206 amino acid protein exhibiting 50–91% identity and 66–94% similarity with BX-VAP-1, -2, -3 and -P. Moreover, in the present study, a cDNA that could encode a similar protein was found amongst ESTs generated from *Bursaphelenchus doui*. This cDNA, designated *Bd-vap-1* (GenBank accession No. GU451047), is full-length and contains an open reading frame for a 202 amino acid polypeptide exhibiting 50–61% identity and 64–76% similarity with BX-VAP-1, -2, -3 and -P.

Multiple sequence alignment of the venom allergen-like proteins from three different *Bursaphelenchus* species showed significant sequence conservation (Fig. 4). Each venom allergen-like protein contained the eight conserved cysteine residues and the sequence HFT(Q/X) (X: the nonsense mutation in BX-VAP-P, CAG → TAG), which is a conserved variation of the HVTQ sequence found in most members of the venom allergen-like proteins (Hawdon et al., 1996), but it is different from the conserved variation, HWVTQ, a hallmark of venom allergen-like proteins described from the phylogenetically distant plant-parasitic nematodes, *H. glycines* (Gao et al., 2001) and *M. incognita* (Ding et al., 2000; Wang et al., 2007).

3.6. Phylogenetic analysis

A neighbor-joining phylogenetic tree based on the amino acid sequences of VAPs from three *Bursaphelenchus* species and other representative nematodes is shown in Fig. 5. Robustness was assessed by bootstrap analyses using 1000 replicates; Clades with more than 50% support are denoted with bootstrap values on the branches. The *Bursaphelenchus* VAPs cluster together and form a group that was distinct from the VAP homologues of other nematodes, including *H. glycines*, *M. incognita*, *Caenorhabditis elegans* and some animal-parasitic nematodes.

3.7. Localization of *Bx-vap-1*, *Bx-vap-2* and *Bx-vap-3* mRNA

In situ mRNA hybridization showed that digoxigenin-labeled anti-sense probes generated from *Bx-vap-1*, *Bx-vap-2* and *Bx-vap-3* specifically hybridized with transcripts in the esophageal gland cells of *B. xylophilus*, respectively (Fig. 6A–C), but it was not possible to determine which of the gland cells was the source of the signal in this nematode. No hybridization was observed in *B. xylophilus* with the control sense cDNA probes of *Bx-vap-1*, *Bx-vap-2* and *Bx-vap-3* (data not shown). Hybridization signals were detected in esophageal gland cells of female, male, and propagative larvae of *B. xylophilus*.

3.8. Developmental expression

The developmental expression patterns of *Bx-vap-1*, *Bx-vap-2* and *Bx-vap-3* were investigated by PCR amplification of stage-specific cDNAs using the gene-specific primer pairs VAP1s2/VAP1a2 for *Bx-vap-1*, VAP2s2/VAP2a2 for *Bx-vap-2* and VAP3s2/VAP3a2

BX-VAP-1	MFGLLVLTVFATPTTADRFNSQKQAILDAHNRLRRRLARGEQAQKNGAMA	52
BX-VAP-2	MVRVFLVLTALFQGIAATKFSQKQAILDAHNSRRRLANGEAEAKNGTMA	52
BX-VAP-3	MFRVLLTTSVLASIVAEARLSDDEKQILFVDNHNRLRRRLANGEITNSKNGEMP	52
BM-VAP-1	MSRVVLTALFVIGQIAATKFSQKQAILDAHNSRRRLANGEQAQNGTMA	52
BD-VAP-1	MFRILLASALLVGPIVATKYTTSEKQALVDAHNRLRRRLAKGEQAQKNGIME	52
BX-VAP-P		0
BX-VAP-1	KGANIREITWDADLESKADEWQAACNFEHPTNVGYGQNLAQMPHVGADSAA	104
BX-VAP-2	KGANIREITWDADLEAKADEWQAACNFEHPTNGGYGQNLAQMPHVGDSAV	104
BX-VAP-3	QGANINEVITWDDLEASAVEWAEACAYHHPSDAPYQGNLAQMPHTNNY..L	102
BM-VAP-1	KGANIREITWDADLESKADEWQAACNFEHPSNGGYGQNLAQMPHVSADSAA	104
BD-VAP-1	EGANIREISWNGLETQANNWAKACNFEHPDESEYQGNLAQMPFVSRASTL	104
BX-VAP-P	CVWGHNPDLKTQGNLAQMPHNSPDRAI	28
BX-VAP-1	TGSPQTWNEINRYSGDAKFEWSTRTGHFTQMAWAETNKIGCAVQNCINGDT	156
BX-VAP-2	NGSAQMWWDEIKDYSGDAKFEWSTSTGHFTQMAWAETDKIGCAIQNCINGDT	156
BX-VAP-3	TR.VEKWYNBSSKYSGNPHLEYSSGIGHFTQVWADTNKIGCAVQDCSNGLS	153
BM-VAP-1	TGSAQMWWDEIKDYSGDAKFEWSTSTGHFTQMAWAETDKIGCAIQNCINGDT	156
BD-VAP-1	TR.PDKWWSEITKYSGNPQLEWGSSTGHFTQMAWAETDKVGCALQNCINGAD	155
BX-VAP-P	SRSTDSSWREIKGYS DG.....TSFGHFTXMAWAETDRIGCALKNCSNGKA	74
BX-VAP-1	QGFEWKDWTFTVCONYMNPGNVMEADLYITGEPCKSC..QGSCSDSLCSA.	204
BX-VAP-2	QGFGEWKDWTFTVCDYMNPGNMGGADLYITGEPCKSCSSQGSQSDSLCSA.	206
BX-VAP-3	QGLGDDN.WTFTVCONYDPAGNVGDGDIYQTGDPSCSQEEEEEGNDSLCSAE	203
BM-VAP-1	QGFGEWKDWTFTVCDYMNPGNYDGEDMYIVGEPCKSCSNKGSQSDSLCSA.	206
BD-VAP-1	QGFYSYDN.WTFTVCONYLHGVNVIDENIYETGSPSCSQCS..VGQNDGLQSD.	202
BX-VAP-P	QKLDYD.DWTFTAQNYMPLGNVASYDVIIPGKPCSKCGSKGVCKNGLVA.	123

Fig. 4. Comparison of amino acid sequences of the venom allergen-like proteins from *Bursaphelenchus xylophilus* (BX), *Bursaphelenchus mucronatus* (BM) and *Bursaphelenchus doui* (BD). Gaps are introduced for obtaining optimal alignment. Amino acids are numbered to the right of each line; identical residues are in black, highly conserved are in gray; the conserved motif of venom allergen-like proteins is underlined, invariant cysteine residues are highlighted by asterisks; the nonsense mutation (x) in BX-VAP-P is indicated with the triangle; the conserved exon boundary is indicated by the vertical line above the alignment, the exon boundary having intergenic differences is indicated by the double vertical line; and signal peptide cleavage sites (between amino residues 17–18 for all but BD-VAP-1 and amino acid residues 18–19 for BD-VAP-1) are pointed out by arrows. BM-VAP-1 was deduced from the EST from *B. mucronatus* (CJ976257); BD-VAP-1 was deduced from the full-length cDNA from *B. doui*; BX-VAP-P was deduced from the predicted coding nucleotide sequence of *Bx-vap-P*.

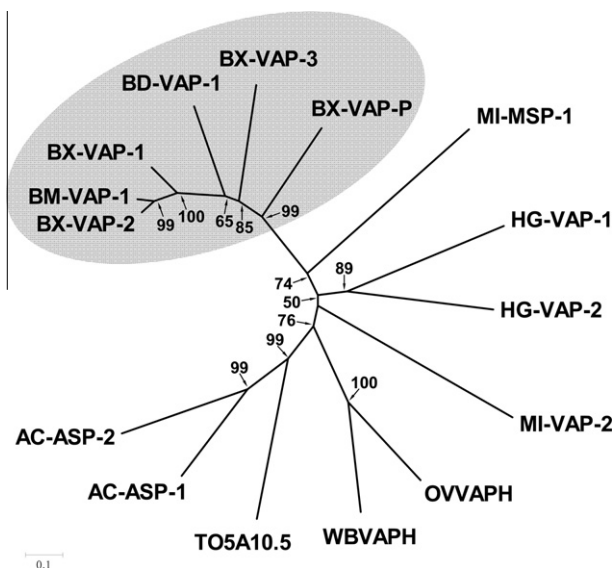


Fig. 5. Unrooted phylogenetic tree of venom allergen-like proteins is generated based on the protein sequences using the neighbor-joining method. MI-MSP-1 (AF013289) and MI-VAP-2 (EF370395) from *Meloidogyne incognita*; HG-VAP-1 (AF374388) and HG-VAP-2 (AY033601) from *Heterodera glycines*; AC-ASP-1 (Q16937) and AC-ASP-2 (AF08728) from *Ancylostoma caninum*; OVVAPH (AF042087) from *Onchocerca volvulus*; WBVAPH (AF109794) from *Wuchereria bancrofti*; T05A10.5 (NP509802) from *Caenorhabditis elegans*.

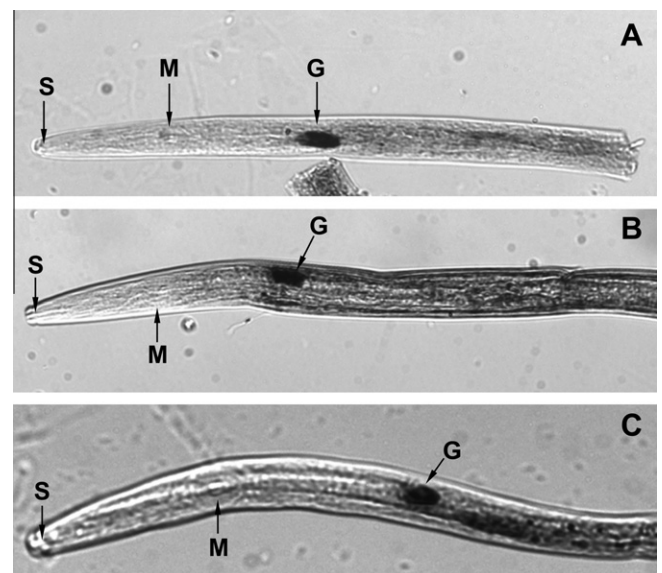


Fig. 6. Localization of *Bx-vap-1*, *Bx-vap-2* and *Bx-vap-3* transcripts in the esophageal gland cells of *Bursaphelenchus xylophilus* revealed by in situ hybridization. Nematode sections were hybridized with anti-sense *Bx-vap-1* (A), *Bx-vap-2* (B) or *Bx-vap-3* (C) digoxigenin-labeled cDNA probes. (G) Esophageal glands; (S) stylet; (M) metacarpus.

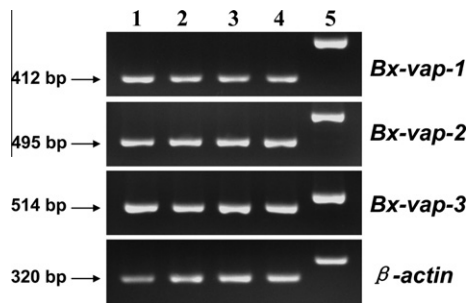


Fig. 7. Developmental expression pattern of *Bx-vap-1*, *Bx-vap-2* and *Bx-vap-3* by semi-quantitative RT-PCR. Lanes 1–5 indicate the size of *Bx-vap-1*, *Bx-vap-2* and *Bx-vap-3* amplified from cDNA templates from differentiated eggs (first-stage and unhatched second-stage juveniles) (1), hatched second-stage juveniles (2), adult females (3), adult males (4) and genomic DNA (5). As a positive control, all cDNA templates were amplified with the primers of β -actin from *Bursaphelenchus xylophilus*.

for *Bx-vap-3*, respectively. The transcription of *Bx-vap-1*, *Bx-vap-2* and *Bx-vap-3* was detectable in differentiated eggs (first-stage and unhatched second-stage juveniles), hatched second-stage juveniles, adult females and males of *B. xylophilus* (Fig. 7). Since the three pairs of primers spanned the introns of genomic DNA and allowed differentiation of PCR products from genomic DNA and cDNA, the amplification of the PCR product from cDNA templates in all the reactions indicated no genomic DNA contamination.

4. Discussion

In this work, we characterized three complete cDNAs from *B. xylophilus*, which encoded proteins showing similarity to venom allergen-like protein, and identified a corresponding putative untranscribed pseudogene. The four genomic sequences of putative venom allergen-like protein gene and pseudogene are tightly grouped and organized within a 3.7 kb cluster. The coding sequences of *Bx-vap-1*, 2 and 3 are oriented in the same direction, while that of *Bx-vap-3* is in the opposite direction. Protein sequence analysis demonstrated that BX-VAPs are a class of CRISPs. Comparison with various databases revealed similarities to known venom allergen-like proteins (VAPs) genes isolated from *H. glycines* and other nematode species. To our knowledge, the possible existence of a venom allergen-like protein gene cluster in any species has not been reported.

The presence of a signal peptide in the deduced amino acid sequence of *Bx-vap-1*, -2 and -3, and detection of corresponding mRNA expression in the esophageal glands suggest that BX-VAP-1, -2 and -3 may be synthesized in the esophageal glands and secreted from the stylet of the nematode. Because *B. xylophilus* is a serious pathogen of forest trees and secretory proteins from its esophageal glands are considered to play essential roles in pathogenicity (Kikuchi, 2008), we speculate that the BX-VAP group is involved in infection and parasitism, as in other plant-parasitic nematodes, including *Meloidogyne incognita* (Ding et al., 2000; Wang et al., 2007) and *Heterodera glycines* (Gao et al., 2001). Future research is necessary to elucidate the roles of BX-VAPs.

To investigate the evolutionary relationships between the *Bursaphelenchus* VAPs and the representative nematodes VAPs from the database, a neighbour-joining tree of the VAP gene family in nematodes was constructed based on deduced amino sequence. These results suggest that the *Bursaphelenchus* VAPs represent a novel subfamily of CRISP proteins. Furthermore, BX-VAP-2 and -3 clearly exhibit higher similarity to their presumed orthologues in *B. mucronatus* and *B. doui*, suggesting that the genes encoding

BX-VAP-1/2/3/P duplicated prior to the divergence of these species and that a corresponding gene cluster may also exist on a chromosome of *B. mucronatus* or *B. doui*. In the future, determination of the potential venom allergen-like protein gene clusters and flanking regions from these nonpathogenic *Bursaphelenchus* species, *B. mucronatus* and *B. doui*, and comparison to the known venom allergen-like protein gene cluster and flanking regions of *B. xylophilus*, will improve our ability to understand the molecular basis for this discord between *Bursaphelenchus* species and venom allergen-like protein evolution.

The previous studies indicated that transcripts of four venom allergen-like protein genes from *H. glycines* and *M. incognita* were strongly detected in both preparasitic and parasitic second-stage juveniles, weakly in late parasitic stages, not in adult females (Gao et al., 2001; Ding et al., 2000; Wang et al., 2007). This expression pattern corresponds to the migratory/infective stages during sedentary endoparasitic nematode development. In this study, transcripts of *Bx-vap-1*, 2 and 3 were detectable in differentiated egg (first-stage and unhatched second-stage juvenile), hatched second-stage juvenile, adult female and male by RT-PCR. Dispersal third- and fourth-stage juveniles were not included in this experiment since we are currently unable to reliably obtain them. Nevertheless, our RT-PCR results suggested that migratory endoparasitic nematode of *B. xylophilus* may have a different pattern of gene expression for venom allergen-like protein genes from sedentary endoparasitic nematodes. It is worth to investigate whether expression pattern of *vap* gene(s) in other migratory endoparasitic nematode such as *Radopholus* and *Pratylenchus* also display similar pattern with *B. xylophilus*, and whether this gene expression pattern is relative with its attributes of migratory endoparasitic nematode which is infective, motile and feeding throughout its life cycle.

In our research, we also identified three entire sequences (designated *Bx-rbp-1*, *Bx-abc-1*, and *Bx-omt-1*) in the same flanking region of the *Bx-vap* gene cluster of *B. xylophilus*. Their predicted cDNA translation products showed homology to RNA-binding motif protein, ABC transporter-related protein and hydroxyindole-O-methyltransferase, respectively. It is worth asking whether these putative modification, transport and regulatory protein genes flanking the *Bx-vap* cluster are presented in other nematode such as cyst and root-knot nematode, and whether they are required for the functional expression of venom allergen-like proteins. Further analysis should help us to understand the biological functions of these interesting venom allergen-like proteins.

Acknowledgment

This research was supported by the Natural Science Foundations of China (Grant No. 30670278). We are grateful to Dr. Chitwood, Nematode Laboratory, ARS, USDA, for reviewing and revising our manuscript.

References

- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zheng, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25, 3389–3402.
- 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, New York, pp. 117–146.
- Cheng, H.R., Lin, M., Li, W., Fang, Z., 1983. The occurrence of a pine wilting disease caused by a nematode found in Nanjing. *Forest Pest and Disease* 4, 1–5.
- De Boer, J.M., Yan, Y., Smart, G., Davis, E.L., Baum, T.J., 1998. In situ hybridization to messenger RNA in *Heterodera glycines*. *Journal of Nematology* 30, 309–312.
- Ding, X., Shields, J., Allen, R., Hussey, R.S., 2000. Molecular cloning and characterization of a venom allergen AG5-like cDNA from *Meloidogyne incognita*. *International Journal for Parasitology* 30, 77–81.
- Falquet, L., Pagni, M., Bucher, P., Hulo, N., Sigrist, C.J., Hofmann, K., Bairoch, A., 2002. The Prosite database, its status in 2002. *Nucleic Acids Research* 30, 235–238.

- Gao, B., Allen, R., Maier, T., Davis, E.L., Baum, T.J., Hussey, R.S., 2001. Molecular characterization and expression of two venom allergen-like protein genes in *Heterodera glycines*. *International Journal for Parasitology* 31, 1617–1625.
- Gill, S.C., von Hippel, P.H., 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry* 182, 319–326.
- Haegeman, A., Jzob, J., Vanholme, B., Kyndt, T., Mitreva, M., Gheysen, G., 2009. Expressed sequence tags of the peanut pod nematode *Ditylenchus africanus*: the first transcriptome analysis of an Anguinid nematode. *Molecular and Biochemical Parasitology* 167, 32–40.
- Hawdon, J.M., Jones, B.F., Hoffman, D.R., Hotez, P.J., 1996. Cloning and characterization of *Ancylostoma*-secreted protein. *Journal of Biological Chemistry* 271, 6672–6678.
- Hu, W.P., Tay, S.K., Zhao, Y., 2006. Endometriosis-specific genes identified by real-time reverse transcription-polymerase chain reaction expression profiling of endometriosis versus autologous uterine endometrium. *The Journal of Clinical Endocrinology and Metabolism* 91, 228–238.
- Hussey, R.S., 1989. Disease-inducing secretions of plant-parasitic nematodes. *Annual Review of Phytopathology* 27, 123–141.
- Kang, J.S., Lee, H., Moon, I.S., Lee, Y., Koh, Y.H., Je, Y.H., Lim, K., Lee, S.H., 2009. Construction and characterization of subtractive stage-specific expressed sequence tag (EST) libraries of the pinewood nematode *Bursaphelenchus xylophilus*. *Genomics* 94, 70–77.
- Kikuchi, T., Aikawa, T., Kosaka, H., Pritchard, L., Ogura, N., Jones, J.T., 2007. Expressed sequence tag (EST) analysis of the pine wood nematode *Bursaphelenchus xylophilus* and *B. mucronatus*. *Molecular and Biochemical Parasitology* 155, 9–17.
- Kikuchi, T., Jones, J.T., Aikawa, T., Kosaka, H., Ogura, N., 2004. A family of GHF45 cellulases from the pine wood nematode *Bursaphelenchus xylophilus*. *FEBS Letters* 572, 201–205.
- Kikuchi, T., Shibuya, H., Aikawa, T., Jones, J.T., 2006. Cloning and characterization of pectate lyases secreted by the pine wood nematode *Bursaphelenchus xylophilus*. *Molecular Plant–Microbe Interactions* 19, 280–287.
- Kikuchi, T., Shibuya, H., Jones, J.T., 2005. Molecular and biochemical characterization of an endo- β -1,3-glucanase from the pine wood nematode *Bursaphelenchus xylophilus* acquired by horizontal gene transfer from bacteria. *Biochemical Journal* 389, 117–125.
- Kikuchi, T., 2008. Parasitism genes of the pine wood nematode. In: Zhao, B.G., Futai, K., Sutherland, J.R., Takeuchi, Y. (Eds.), *Pine Wilt Disease*. Springer, Japan., pp. 67–80.
- Kikuchi, T., Li, H.M., Karim, N., Kennedy, M.W., Moens, M., Jones, J.T., 2009. Identification of putative expansin-like genes from the pine wood nematode, *Bursaphelenchus xylophilus*, and evolution of the expansin gene family within the nematode. *Nematology* 11, 355–364.
- Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics* 5, 150–163.
- Lambert, K.N., Allen, K.D., Sussex, L.M., 1999. Cloning and characterization of an esophageal-gland-specific chorismate mutase from the phytoparasitic nematode *Meloidogyne javanica*. *Molecular Plant–Microbe Interactions* 12, 328–336.
- Mamiya, Y., 1988. History of pine wilt disease in Japan. *Journal of Nematology* 20, 219–226.
- Mota, M., Braasch, H., Bravo, M.A., Penas, A.C., Burgermeister, W., Metge, K., Sousa, E., 1999. First report of *Bursaphelenchus xylophilus* in Portugal and in Europe. *Nematology* 1, 727–734.
- 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.
- Ray, C., Abbott, A.G., Hussey, R.S., 1994. Trans-splicing of a *Meloidogyne incognita* mRNA encoding a putative esophageal gland protein. *Molecular and Biochemical Parasitology* 68, 93–101.
- Rutherford, T.A., Mamiya, Y., Webster, J.M., 1990. Nematode-induced pine wilt disease: factors influencing its occurrence and distribution. *Forest Science* 36, 145–155.
- Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. *Molecular Cloning: A Laboratory Manual*, second ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- 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*. *Molecular and Biochemical Parasitology* 88, 203–213.
- 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*. *Molecular Plant–Microbe Interactions* 14, 536–544.
- Wang, X., Li, H.M., Hu, Y.J., Fu, P., Xu, J.H., 2007. Molecular cloning and analysis of a new venom allergen-like protein gene from the root-knot nematode *Meloidogyne incognita*. *Experimental Parasitology* 117, 133–140.
- Yi, C., Byun, B., Park, J., Yang, S., Chang, K., 1989. First finding of the pine wood nematode, *Bursaphelenchus xylophilus* (Steiner et Buhrer) Nickle and its insect vector in Korea. *Research Reports of the Forestry Research Institute Seoul* 38, 141–149.