Mining the secretome of root-knot nematodes for cell wall modifying proteins

Erwin H. A. Roze

Promotor:
Prof. dr. ir. J. Bakker
Hoogleraar in de Nematologie
i.h.b. fysiologie en moleculaire ecologie van nematoden
Wageningen Universiteit

Copromotor:
Dr. ir. G. Smant
Universitair docent
Laboratorium voor Nematologie
Wageningen Universiteit

Promotiecommissie:
Prof. dr. J.M. Vlak (Wageningen Universiteit)
Dr. R.C.H.J. van Ham (PRI, Wageningen Universiteit en Researchcentrum)
Dr. ir. F.P.M. Govers (Wageningen Universiteit)
Dr. M.-N. Rosso (INRA, Frankrijk)

Dit onderzoek is uitgevoerd binnen de onderzoeksschool 'Experimental Plant Sciences'

Mining the secretome of root-knot nematodes for cell wall modifying proteins

Erwin H. A. Roze

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. dr. M. J. Kropff,
in het openbaar te verdedigen
op vrijdag 15 februari 2008
des namiddags te half twee in de Aula.

Roze, E.H.A. (2008).

Mining the secretome of root-knot nematodes for cell wall modifying proteins.

PhD thesis Wageningen University, Wageningen, The Netherlands
with summaries in English and Dutch

ISBN: 978-90-8504-858-9

Contents

1

General introduction

Nematodes live life to the max

Nematodes, a diverse group of roundworms, are found in practically every soil sample often with high species diversity and a high abundance. The number of nematodes in the soil is generally somewhere between one to ten million/m^2 (Sohlenius, 1980). Lambshead (1993) estimated that there may be as many as one hundred million nematode species in the deep sea alone, but the number of described nematode species found in freshwater and marine sediments and soils is currently about 20,000 (Boucher and Lambshead, 1995); (Lambshead, 1993). Some soil nematode species feed on bacteria or fungi, while others are omnivorous, predaceous, or parasites of plant roots (Yeates *et al.*, 1993). On top of their broad spectrum of feeding habits, nematodes also live in a wide range of ecological niches including the most extreme environments on earth (Bird and Koltai, 2000; Wall and Virginia, 1999). Examples of such extreme environments are hot and cold deserts, which are characterised by extreme temperatures, limited water availability, and high soil pH and salinity (Freckman and Virginia, 1989). Nematode communities are for example found in the Antarctic soil ecosystem, representing cold deserts, at Machu Picchu station on King George Island, and in the McMurdo Dry Valleys of Antarctica (Mouratov *et al.*, 2001; Treonis and Wall, 2005). At the other extreme end, plant-feeding nematodes are found, even to a depth of 11-12 meters, in the Chihuahuan Desert in New Mexico, USA (Freckman and Virginia, 1989; Wall and Virginia, 1999).

Soil nematodes range in size from 0.3 to 3 mm and are active in water-filled pores and the water films surrounding soil particles. To survive adverse conditions such as drought, some species can enter into an inactive state of anhydrobiosis ('life without water') (McSorley, 2003; Treonis and Wall, 2005). That this inactive state allows long-term survival is supported by observations that nematodes could be revived after years in anhydrobiosis. One of the species capable of surviving harsh circumstances by anhydrobiosis is *Filenchus polyhypnus,* which was revived from dried plant material after a period of 39 years (McSorley, 2003). Plant parasitic nematodes, like root-knot (*Meloidogyne* spp.) and cyst (*Globodera* and *Heterodera* spp.) nematodes, also employ strategies to survive during cold seasons and in the absence of a host. These strategies include developmental dormancy and diapause, changing sex ratios, and the formation of a protective cyst that covers the eggs in the case of cyst nematode species (Gaur *et al.*, 2000; McSorley, 2003; Papadopoulou and Triantaphyllou, 1982; Perry, 1989).

Meloidogyne chitwoodi and *Meloidogyne incognita*

The root-knot nematodes mentioned above are sedentary endoparasites of the economically most damaging family of plant parasitic nematodes, the *Heteroderidae*

(Williamson and Gleason, 2003). They are obligate, biotrophic pathogens that have evolved highly sophisticated parasitic relationships with their hosts. The *Meloidogyne* species, including *M. incognita* and *M. chitwoodi*, are the most important of the plant parasitic nematodes since they infect almost all cultivated plants (Trudgill and Blok, 2001).

At present, more than ninety different root-knot nematode species have been described in the literature (Karssen and Den Nijs, 2004). The Columbia root-knot nematode *Meloidogyne chitwoodi* was first described in 1980 as an endoparasite of potato in the Pacific Northwest of the USA (Golden *et al.*, 1980). Official reports on the prevalence of *M. chitwoodi* suggest a patchy distribution limited to specific areas in the Americas, Africa, and Europe, where it occurs in regions with cooler, temperate climates (Karssen and Den Nijs, 2004). *M. chitwoodi* larvae hatch from eggs as soon as the soil temperature rises above 5-10°C. This tolerance for lower temperatures allows for three to four generations per year in Europe (www.eppo.org). *Meloidogyne incognita* (Kofoid and White, 1919) is distributed mainly in warmer temperate and tropical regions, where soil temperatures rarely fall below 3°C (Trudgill and Blok, 2001).

Based on their host plants, *Meloidogyne* species can be grouped into three groups comprising species mainly parasitizing i) dicotyledons, ii) monocotyledons, or iii) both mono- and dicotyledons (Karssen, 1999). Both *M. incognita* and *M. chitwoodi* belong to the group thriving on both mono- and dicotyledons. *Meloidogyne chitwoodi* and *M. incognita* also share a polyphagous behaviour, i.e. have a wide host range, since they parasitize thousands of different plant species (Trudgill, 1997; Trudgill and Blok, 2001; Williamson and Hussey, 1996). Among their host plants are many economically important crop species (like tomato, potato, sugar beet, maize, and wheat), flowering and ornamental plants, and common weed species (Santo *et al.*, 1980); (Den Nijs *et al.*, 2004; O'bannon, 1982). Because of their wide host ranges, short generation times, and high reproductive rates these root-knot nematode species are difficult to control (Trudgill and Blok, 2001).

Life cycle of *Meloidogyne*

The life cycle of root-knot nematodes includes a period in the soil, outside the host plant, and a parasitic period inside the roots of a host. Adult female root-knot nematodes release hundreds of eggs into a gelatinous matrix on the surface of the root. Following a first moult inside the egg, motile, second-stage juveniles (J2) hatch in the soil and penetrate a plant root behind the tip in the elongation zone. Within the root, the roundworm migrates intercellularly in a stealthy way through the cortex and induces the formation of large multinucleate cells called 'giant cells' (Bird, 1996; Bird and Koltai, 2000; Wyss *et al.*, 1992). These giant cells share features with transfer

cells and provide the sedentary nematode with a constant supply of nutrients (Jones and Dropkin, 1976). Secretory proteins synthesized by the nematode in the dorsal and two subventral oesophageal gland cells play an important role during the migration through plant roots and subsequent induction and maintenance of the giant cells (Davis *et al.*, 2000; Hussey *et al.*, 1994; Hussey *et al.*, 2002; Hussey, 1989). These proteins are secreted through the oral stylet, a hollow, protrusible feeding spear, which is also used to penetrate the plant cell wall and to withdraw nutrients from the host cell cytoplasm. Once the feeding site is established, the second-stage juveniles undergo significant changes, including swelling of the body and a switch in gland activity from subventral to dorsal dominance (Bird, 1983; Davis *et al.*, 2000; Endo, 1987; Hussey and Mims, 1990). A J2 feeds from the giant cells for 10 to 12 days, then ceases feeding and moults three times over the next two days to become an adult. Adult females swell to a pear shape and become sedentary. The cells surrounding the developing juvenile and the giant cells typically proliferate and enlarge to form a gall, hence the common name root-knot nematodes.

The secretome and the parasitome of nematodes

Their stealthy penetration of host plants and the profound cellular changes root-knot nematodes induce inside host tissues give evidence of a complex molecular interaction between nematode and host involving many parasite effector molecules. The products of parasitism genes in nematodes coding for these effector molecules have to be secreted by the nematode to reach their molecular targets at the nematode-plant interface. Therefore, the identification of members of the secretome, i.e. the repertoire of secreted proteins in an organism, has been the main focus of plant-parasitic nematode research for decades as it may lead to the subset of proteins that are at the molecular basis underlying this interaction (Hussey *et al.*, 2002; Vanholme *et al.*, 2004). This subset of proteins (below referred to as 'the parasitome') can either be secreted into the plant extracellular space (the apoplast) or into the host plant cell (the symplast) (e.g. (Huang *et al.*, 2006; Wang *et al.*, 1999)). Cell wall modifying proteins, for example, constitute apoplastically secreted nematode proteins that act on the plant cell wall, thereby facilitating the invasion and intercellular migration through the plant root. These proteins will be discussed in more detail later on in this introduction. Proteins secreted into the plant cell cytoplasm induce, either directly or indirectly, changes in gene expression and cell development in the host (Gheysen and Fenoll, 2002; Williamson and Hussey, 1996). Besides the elaborate modifications of host cells, the parasitic lifestyle of nematodes inside a host also likely involves suppression or evasion of host defence mechanisms conditioned by molecules released by the parasite. For example, antidefence proteins are produced in the hypodermis and coat the surface of the nematode, these include

antioxidant proteins such as superoxide dismutase, thioredoxin peroxidase, and a lipoxygenase inhibiting protein, which may block jasmonic acid-dependent defence responses (Prior *et al.*, 2001; Robertson *et al.*, 1999; Robertson *et al.*, 2000).

Hussey *et al.* (2002) confined the definition of the nematode parasitome to secretions produced in the oesophageal secretory glands by defining it as 'the complete profile of parasitism gene products secreted through the nematode stylet during the parasitic cycle' (Hussey *et al.*, 2002). Although the oesophageal secretory glands are the principal sources of secretions involved in plant parasitism, other nematode structures are also involved in the production of secretions and therefore a broader definition of parasitome was chosen: 'the products of parasitism genes secreted by a range of secretory organs, including the oesophageal, amphidial, and rectal glands, the intestine, and the hypodermis'. The implication of this definition is that in order to identify members of the parasitome within the secretome, it is necessary to systematically localise their primary site of expression in the nematode.

Identification of members of the parasitome using expressed sequence tags

Expressed sequence tags (ESTs) are unedited sequences obtained from either end of a cDNA molecule, which are generated by single-pass sequencing of randomly selected cDNA clones from a library (McCarter *et al.*, 2000). The generation of ESTs is a relatively simple, rapid and inexpensive procedure that has become a widely embraced approach for large-scale gene discovery in eukaryotic plant pathogens, such as nematodes, oomycetes, and fungi (Dautova *et al.*, 2001; Kamoun *et al.*, 1999; Popeijus *et al.*, 2000a; Qutob *et al.*, 2000; Skinner *et al.*, 2001; Torto *et al.*, 2003).

The Parasitic Nematode Sequencing Project is underway to produce over 300,000 ESTs from thirty different parasitic nematode species, of which nearly 117,000 ESTs (July 2006) originate from whole nematode libraries of plant parasitic nematodes (http://www.nematode.net/; (Wylie *et al.*, 2004)). Mining of EST datasets has already resulted in the identification of a suite of transcripts encoding parasitism-related proteins in plant parasitic nematodes, e.g. pectate lyase (Popeijus *et al.*, 2000b), β-1,4-endoglucanase, β-1,4-endoxylanase (Dautova *et al.*, 2001), polygalacturonase (Jaubert *et al.*, 2002), glutathione peroxidase (Jones *et al.*, 2004), and chorismate mutase (Jones *et al.*, 2003). EST analysis has not only given insight in the identity of putative phylum- and family-specific genes, but also revealed novel horizontally transferred gene candidates in *Meloidogyne* (McCarter *et al.*, 2003).

As is the case for any approach, there are strengths and weaknesses associated with the use of ESTs as research tools (McCarter *et al.*, 2000). Expressed sequence tags reflect the abundance of mRNAs present in the starting material that is used to create the cDNA library. Because these libraries are for example prepared from a specific

life stage of an organism, sequences can be selected for further study based on their abundance in the EST dataset. Following this criterion, it is possible to select sequences even without the finding of any similarity to known sequences in databases. A limitation, however, is that it is difficult to point out putative parasitism-related sequences among the ESTs that are neither abundant nor have any similarity to known sequences. In addition, the fact that ESTs reflect the abundance of mRNAs also means that rare transcripts are likely to be missing from EST datasets (McCarter *et al.*, 2000).

Because ESTs are unedited, single-pass sequences, a weakness is that they can contain nucleotide sequencing errors. Secondly, in most cases they represent fragments of transcripts. Both characteristics can make the annotation of sequences difficult and, therefore, their putative function can remain obscure. The problems of low quality and partial nature of individual ESTs can be addressed to some extent by constructing contigs, i.e. groups of ESTs assembled together based on sequence similarity (McCarter *et al.*, 2003) (see Chapter 2). Contig building also allows to reduce and to estimate the redundancy within a library. Reduction of redundancy results in a decrease of the size of the dataset that can be used for further analyses. The estimation of redundancy provides a measure of complexity of the cDNA library, where more complex means less redundant. In addition, assembly of ESTs into contigs enables the identification of abundant transcripts (McCarter *et al.*, 2003).

Assembled ESTs can be further analysed to identify members of the secretome and subsequently the parasitome by combining a series of algorithms and *in situ* hybridisation. BLAST searches can be performed to search for similarity with known sequences in order to assign putative functions. Based on evident similarities to secretory proteins in a database, interesting candidates can be pointed out. Most secreted proteins in eukaryotes, including plant parasitic nematodes, are exported through the classical secretory pathway via short, N-terminal, leader peptide sequences known as signal peptides. This means that most members of the secretome can be identified by screening the sequences for the presence of such signal peptides. The SignalP program, that was developed using machine learning methods, can assign signal peptide prediction scores and cleavage sites to amino acid sequences with a high degree of accuracy (Menne *et al.*, 2000; Nielsen *et al.*, 1999; Schneider and Fechner, 2004). With the use of SignalP it is possible to identify proteins containing putative signal peptides from ESTs irrespective of the presence of a matching sequence in the database (see Chapter 2). Several EST projects have used such predictions to identify putative secreted proteins in plant pathogens, for example in the plant pathogenic oomycete *Phytophthora infestans* ((Torto *et al.*, 2003), reviewed in (Kamoun, 2006)), animal parasitic nematode *Nippostrongylus brasiliensis* (Harcus *et al.*, 2004), and the plant parasitic nematode *Heterodera schachtii* (Vanholme *et al.*, 2006).

The set representing members of the secretome thus identified can be further narrowed down by comparative EST analysis to assess if they are nematode specific, specific for plant-parasites, or family specific. However, the final and most important criterion to point out candidate parasitome members involves *in situ* hybridisations on whole mount nematodes to see if the corresponding genes are expressed in organs associated with secretion of molecules to the host-parasite interface. The members of the parasitome selected in this way require further validation by functional assays such as *in planta* expression and biochemical analysis.

Plant cell wall

This research was focused on the identification of members of the parasitome of two species (*M. incognita* and *M. chitwoodi*) that feed on both monocotyledons and dicotyledons, which differ substantially in the constitution of their cell walls. We hypothesized that *M. chitwoodi* and *M. incognita* have to match this diversity in plant cell walls with their repertoire of cell wall modifying proteins to be able to penetrate thousands of different plant species in a stealthy fashion. The hypothesis of the matching diversity was tested in this thesis, but for the sake of clarity I will first elaborate on the constitution of plant cell walls followed by a review of cell wall modifying proteins in nematodes which were known at the outset of the project.

The plant cell wall serves many functions in a plant, on one hand, it enables intercellular adhesion and is mechanically rigid in order to provide structural support, while on the other hand, it remains sufficiently flexible to allow for cell elongation until the cell ceases to grow. Overall, the plant cell wall dictates the size and shape of the plant cell it surrounds and ultimately also the shape of the plant itself. Another function of the plant cell wall is to protect the plant against external attacks undertaken by parasites or pathogens that (try to) invade the plant.

When root-knot nematodes enter the plant root and migrate intercellularly through the root system, they are faced with the physical barrier formed by the plant cell wall. The root tissues through which the *Meloidogyne* species migrate mostly include differentiating and growing cells, and therefore, they will not have to breakdown secondary plant cell walls, which are very resistant to enzymatic degradation. In the cell wall between growing cells, three layers can be distinguished: a primary cell wall, the middle lamella and again a primary cell wall. The basic structure of the primary cell wall in all land plants appears to be very similar, but the arrangements and molecular composition of the cell wall vary depending on the plant species, tissue type, cell type and stage of differentiation (Carpita and McCann, 2000). The most abundant, but often overlooked, component of the plant cell wall is water, which makes up about two thirds of the wall mass in growing tissues (Cosgrove, 1997a). The main structural framework of the growing

plant cell wall consists of many different polysaccharides, which are held together primarily by non-covalent bonds. This polysaccharide network is formed by cellulose, i.e. β-1,4-linked D-glucose polymers, organized in inextensible microfibrils, which are tethered by hemicelluloses. The latter is a general term for a diverse group of cross-linking glycans and xylans that can hydrogen-bond to cellulose microfibrils (Carpita and McCann, 2000). In turn, the cellulose-hemicellulose network is embedded in a pectin matrix, which extends into the middle lamella (Carpita and McCann, 2000; Carpita and Gibeaut, 1993). The middle lamella is very rich in pectins which 'glue' the cells together. Pectins constitute a heterogeneous group of polysaccharides with a backbone consisting of partially methylesterified galacturonic acid (Cosgrove, 1997a). Different regions are recognized in pectin, the so-called smooth and hairy regions. The smooth region is composed of homogalacturonan, a linear polymer of α-1,4-linked D-galacturonic acid. This region is interrupted by rhamnogalacturonan I and, with lesser abundance, rhamnogalacturonan II (hairy region). The latter two are highly branched and thus complex carbohydrates containing diverse sugars in varying linkages.

There are different models explaining how these variant polysaccharide networks may interact (Cosgrove, 1999). In one model of the plant cell wall, hemicelluloses, such as xyloglucan (XyG), coat the surface of cellulose and directly link the microfibrils together (*e.g.* in (Carpita and Gibeaut, 1993; Cosgrove, 1997b); (Nishitani, 1997)). An alternative model envisions cellulose microfibrils that are coated with hemicelluloses and these are in turn embedded in additional layers of matrix polymers (Talbott and Ray, 1992). The latter model implicates that there is no direct cross-linking of microfibrils in primary cell walls of plants.

In addition to polysaccharides, the plant cell wall contains a relatively small amount of aromatic substances and structural proteins. These proteins may also form networks in the wall and they are classified in four major groups according to their predominant amino acid composition: (i) hydroxy-proline-rich glycoproteins (*e.g.* extensins), (ii) proline-rich proteins, (iii) glycine-rich proteins and (iv) proteoglycans (Showalter, 1993). Many of these proteins are highly glycosylated; proteoglycans can even be more than 95% carbohydrate. On a dry weight basis, the relative amounts of the polysaccharides and proteins in the wall of enlarging plant cells is approximately 30% cellulose, 30% hemicellulose and 35% pectin, with only 1 to 5% structural protein (Cosgrove, 1997a). However, substantial deviations from these figures may be found, particularly in the grasses (Carpita, 1996), where, for example, the walls of growing maize coleoptiles consist of approximately 55% hemicellulose, 25% cellulose, and only 10% pectin.

Next to the differences in the relative amounts of the main polysaccharides, grass primary cell walls differ in more respects from those of all other flowering plants

(Carpita, 1996). Cell walls of most dicots and the non-commelinoid line of monocots contain a so-called 'Type I' cell wall (Carpita and Gibeaut, 1993). In this type of walls, the main hemicellulose is XyG, a branched polymer consisting of a backbone of β-1,4-linked D-glucose chains with short side chains containing xylose, galactose and, often, a terminal fucose (Fry, 1989; McNeil *et al.*, 1984). In the 'Type II' walls, found in the commelinoid line of monocots, the principal polymers that interlock the microfibrils are glucuronoarabinoxylans (GAXs) instead of XyGs (Carpita, 1996). In general, 'Type II' walls are rich in xylans, pectin-poor and have a very low amount of structural protein compared to dicots and other monocots. The grass (order Poales) 'Type II' wall contains a third major cross-linking glycan, called 'mixed-linkage' (1→3),(1→4) β-D-glucans (Carpita and McCann, 2000; Carpita, 1996). Another major feature of the cell walls in the order Poales and its relatives, is the enrichment of aromatic compounds in non-lignified walls (Carpita, 1996).

Cell wall modifying proteins secreted by plant parasitic nematodes

Secretory proteins synthesized in the two subventral oesophageal gland cells of plant-parasitic nematodes play an important role during migration through host plant roots. The combined effect of these gland secretions and the physical damage inflicted by the fierce protrusions of the oral stylet cause the cell walls to break down. The secretory proteins from the subventral oesophageal gland cells include a range of cell wall degrading enzymes, cellulose binding proteins and β-expansins.

The first genes coding for cell wall degrading enzymes (cellulases; EC 3.2.1.4), were identified in cyst nematodes of the genera *Globodera* and *Heterodera* (Smant *et al.*, 1998; Yan *et al.*, 1998). Since then, cellulases have been identified from several other cyst nematode species, the root-knot nematode *Meloidogyne incognita*, the root-lesion nematode *Pratylenchus penetrans*, and the pine wood nematode *Bursaphelenchus xylophilus* (De Meutter *et al.*, 1998; Gao *et al.*, 2002; Goellner *et al.*, 2000; Kikuchi *et al.*, 2004; Rosso *et al.*, 1999; Uehara *et al.*, 2001; Yan *et al.*, 2001). These cellulases, with the exception of those identified from *B. xylophilus*, belong to the glycosyl hydrolase family 5 and comprise two domain architectures (i.e. a catalytic domain with and without a cellulose binding domain). Beta-1,4-endoglucanases degrade polysaccharides possessing β-1,4-glucan backbones such as cellulose and the hemicellulose xyloglucan. Another enzyme acting on the hemicellulose fraction of plant cell walls is endo-1,4-β-xylanase (EC 3.2.1.8).

Since *Meloidogyne* species migrate intercellularly, they go through the middle lamella, which is rich in pectic polysaccharides. For pectin degradation, the combined action of several enzymes is needed. One group of pectin degrading enzymes consists of hydrolases and lyases (depolymerases), which cleave the backbone of pectin (Tamaru and Doi, 2001). Pectate lyases (EC 4.2.2.2) belong to this

group and are known to play a critical role in pectin degradation by catalyzing the random cleavage of internal α-1,4-linkages of pectate by β-elimination (Barras *et al.*, 1994). In addition, a polygalacturonase (EC 2.1.15), that cleaves α-1,4-glycosidic linkages in pectate by hydrolysis, was identified from *M. incognita* (Barras *et al.*, 1994; Jaubert *et al.*, 2002).

Next to these cell wall degrading enzymes, two sequences encoding cellulose-binding proteins were isolated from *H. glycines* (HG-CBP-1) and *M. incognita* (MI-CBP-1) (Ding *et al.*, 1998; Gao *et al.*, 2004a). The cellulose-binding protein HG-CBP-1 only contains a cellulose-binding module (CBM), whereas MI-CBP-1 consists of a CBM linked to an unknown domain. Both proteins lacked cellulase activity, but bound to cellulose and may be indirectly involved in cell wall degradation.

A novel class of nematode secreted cell wall modifying proteins is formed by the β-expansin identified from *G. rostochiensis* (Kudla *et al.*, 2005; Qin *et al.*, 2004). Expansins lack hydrolytic activity and are proposed to weaken non-covalent interactions between cellulose and hemicellulose polymers resulting in a rapid induction of plant cell wall extension (McQueen-Mason and Cosgrove, 1994). Through this action, expansins are believed to make the plant cell wall polysaccharides more accessible to enzymatic attack (Cosgrove, 2000a) and therefore facilitate the invasion of plants by cell wall degrading enzyme secreting plant pathogens.

Outline of the thesis

This thesis deals with the identification and characterization of members of the parasitome of the root-knot nematode species *M. chitwoodi* and *M. incognita*. Within this subset of proteins the focus was on the group of cell wall modifying proteins.

The objective in **Chapter 2** was to identify members of the secretome and subsequently the parasitome of *M. chitwoodi*. We present an efficient pipeline of bio-informatic algorithms which was used to mine the publicly available *M. chitwoodi* ESTs for secretory proteins. The most abundant transcripts in the secretome were screened with *in situ* hybridization microscopy to test whether the corresponding genes were expressed in organs associated with secretion of molecules to the host-parasite interface. With this approach we obtained the first comprehensive overview of the secretome of *M. chitwoodi* and identified the first candidate parasitism genes from this nematode species.

To study the genes expressed in the penetration phase of the nematode-plant interaction, the focus was on cell wall degrading enzymes in the invasive J2 stage of *M. chitwoodi*. In the study described in **Chapter 3**, ESTs of *M. chitwoodi* were screened for similarity with cell wall modifying protein encoding sequences, resulting in the identification of an elaborate repertoire of enzymes that targets essentially all major

cell wall components. It includes β-1,4-endoglucanases, a β-1,4-endoxylanase, pectate lyases, polygalacturonases and a cellulose binding protein. Characteristics of each (group of) cell wall modifying proteins from *M. chitwoodi* are described and discussed.

In **Chapter 4** the question is addressed whether the secretion of expansins, in combination with cell wall modifying enzymes, is unique to cyst nematodes or that similar combinations of cell wall modifying proteins are also secreted by other plant-parasitic nematodes and even by other unrelated plant pathogens. The expansin-like sequences that were identified from *M. chitwoodi* are described. In addition, the results of a screening of a number of other root-knot nematode species, a root-lesion nematode, a plant pathogenic oomycete and fungi for expansin-like sequences are shown. We also discuss the proposed nomenclature of the expansin superfamily and the value of our findings regarding the definition and key amino acid signatures of expansins.

In **Chapter 5** the characterization of an endo-1,4-β-xylanase, an enzyme that catalyzes the hydrolysis of substituted xylan polymers, from the root-knot nematode *M. incognita* is described. It concerns the first finding of a functional endo-1,4-β-xylanase gene from an animal.

In **Chapter 6** the results of the thesis are summarised and discussed, followed by a layman's version of the summary in Dutch (**Chapter 7**).

2

Mining the secretome of the root-knot nematode *Meloidogyne chitwoodi* for candidate parasitism genes

Erwin Roze, Bram Hanse, Makedonka Mitreva,
Bartel Vanholme, Jaap Bakker and Geert Smant

Molecular Plant Pathology (2008) 9 (1), p. 1-10

SUMMARY

Parasite proteins secreted at the interface of nematode and host are believed to play an essential role in parasitism. Here, we present an efficient pipeline of bio-informatic algorithms and laboratory experiments to identify candidate parasitism genes within nematode secretomes, i.e. the repertoire of secreted proteins in an organism. We performed our approach on 12,218 ESTs originating from three life stages of the plant parasitic nematode *Meloidogyne chitwoodi* – a molecularly unexplored root-knot nematode species. The ESTs from *M. chitwoodi* were assembled into 5,880 contigs and open reading frames translated from the consensus sequences were searched for features of putative signal peptides for protein secretion and trans-membrane regions, resulting in the identification of 398 secretome members. The products of parasitism genes are secreted by a range of organs, including the oesophageal, amphidial, and rectal glands, the intestine, and the hypodermis. To localise the site of expression in *M. chitwoodi*, we subjected the most abundant secretome members to *in situ* hybridization microscopy. We found hybridization of one tag in the dorsal oesophageal gland, seven in the two subventral oesophageal glands, two in the intestine, and one tag hybridised to the tail tip in the proximity of the phasmids. Four sequences showed similarity to putative parasitism genes from other nematode species, whereas seven represented pioneering sequences. Our approach presents an efficient method to identify candidate parasitism genes, which does not require sophisticated cDNA isolation and selection protocols, and can therefore be used as a powerful starting point for the molecular investigation of parasites.

INTRODUCTION

The Columbia root-knot nematode *Meloidogyne chitwoodi* was first described in 1980 as an obligate endoparasite of potato in the Pacific Northwest of the USA (Golden *et al.*, 1980). Further investigations revealed that *M. chitwoodi* has a wide host range, including many arable crops, flowering and ornamental plants, trees and common weed species. The polyphagous nature of this nematode makes it a difficult parasite to control with traditional strategies, like crop rotation, which are rarely effective on *M. chitwoodi* and require extremely good weed control (Den Nijs *et al.*, 2004; O'bannon, 1982; Santo *et al.*, 1980). Attempts to manage *M. chitwoodi* infestations using environment unfriendly nematicides have not been adequate either (Ingham *et al.*, 2000; Pinkerton *et al.*, 1986; Santo and Wilson, 1990). Official reports on the prevalence of *M. chitwoodi* suggest a patchy distribution limited to specific areas in the Americas, Africa, and Europe. For its potential to become established all over continents and its high damage probability, *M. chitwoodi* acquired a quarantine status in Europe (Den Nijs *et al.*, 2004).

The parasitic cycle of *M. chitwoodi* starts with the infective second stage juveniles hatching from eggs that are deposited by the females in egg masses on root surfaces. These pre-parasitic juveniles (ppJ2-s) penetrate the plant root and migrate intercellularly through the cortex (Wyss *et al.*, 1992). To go around the impervious barrier formed by the Casparian strip, root-knot nematodes migrate down towards the root tip, make a U-turn and enter the vascular cylinder. In the differentiation zone they establish a permanent feeding site built from host cells transformed into giant cells by the induction of nuclear division without cytokinesis (Williamson and Gleason, 2003). Feeding nematodes take up nutrients from the cytoplasm of giant cells through their stylet (a hollow, protrusible buccal spear). After three consecutive moults at the start of feeding, the nematodes develop into either vermiform, motile adult males or pear-shaped sedentary females to complete the parasitic cycle within two months.

The parasitic lifestyle of nematodes inside a host involves elaborate modifications of host cells and suppression or evasion of host defence mechanisms conditioned by molecules released by the parasite. The identification of members of the secretome, i.e. the repertoire of secreted proteins in an organism, of a parasite will aid to unravel these molecular interactions at the interface of host and parasite. Secretory proteins produced in the oesophageal, amphidial, and rectal glands and in the hypodermis and intestine, are believed to play crucial roles in plant parasitism (Davis *et al.*, 2000). The Parasitic Nematode Sequencing Project is underway to produce over 300,000 expressed sequence tags (ESTs) from thirty different parasitic nematode species

(http://www.nematode.net; (Wylie *et al.*, 2004). The generation of ESTs, by single-pass sequencing of randomly selected cDNA clones, forms an invaluable resource for gene discovery (McCarter *et al.*, 2000; McCarter *et al.*, 2003). Our objective was to mine this resource for candidate parasitism genes by identifying and analyzing members of the secretome of a parasite.

Although several members of the secretome have been identified from the distantly related root-knot nematode *Meloidogyne incognita*, *M. chitwoodi* was virtually unexplored at the molecular level (reviewed in Davis *et al.*, 2004; Davis *et al.*, 2000; Hussey *et al.*, 2002; Vanholme *et al.*, 2004). Up to July 2006, the Parasitic Nematode Sequencing Project has generated nearly 117,000 ESTs from whole nematode libraries of plant parasitic nematodes of which 12,218 originate from three distinct life stages of *M. chitwoodi*. In this study, we clustered these *M. chitwoodi* ESTs to identify and to analyze members of its secretome by combining a sequence of algorithms and *in situ* hybridization microscopy. The algorithms aimed to identify N-terminal leader peptide sequences for protein secretion and corresponding cleavage sites, membrane anchors, trans-membrane regions, and other sorting signal motifs. *In situ* hybridization microscopy was subsequently done with the most abundant transcripts in the secretome to test if the corresponding genes are expressed in organs associated with secretion of molecules to the host-parasite interface. We were able to identify eleven candidate parasitism genes from *M. chitwoodi*, which can be used in further research to unravel their exact role in the parasitic cycle and to identify new targets for nematode control. Our strategy provides an efficient starting point for the molecular identification of parasitism genes without sophisticated cDNA isolation and selection methods and signal peptide trapping procedures.

MATERIALS AND METHODS

Contig assembly

ESTs from *M. chitwoodi* were downloaded from the dbEST division of GenBank on 10 September 2005. They originate from 3 different life stages (Table 1). These ESTs were assembled into contigs using the programs Phred, Phrap and Consed (McCarter *et al.*, 2003; Mitreva *et al.*, 2004b). Sequence trace files and the contig fasta files of NemaGene *Meloidogyne chitwoodi* are available at http://www.nematode.net.

Protein sorting algorithms

Contig consensus sequences were used as input for the Perl-based program SPIT (Vanholme *et al.*, 2006). The SPIT program generates polypeptide sequences in the six

Table 1 Characteristics of the *M. chitwoodi* cDNA libraries (I) and the EST and contig statistics (II)

(I) cDNA library characteristics				(II) EST and contig statistics		
Life stage	Directionality	PCR	#EST[b]		EST	Contig
egg	directional	SMART PCR	4,965	number	12,218	5,880
egg	non-directional	SL1 - oligo(dT) PCR	1,766	mean length (nt)[c]	428	445
ppJ2[a]	directional	SMART PCR	4,067	minimum length (nt)	49	51
adult female	directional	SL1-SMART oligo(dT) PCR	1,420	maximum length (nt)	730	1,746
		total	12,218	total length (nt)	5,217,184	2,615,370

[a]Pre-parasitic second stage juvenile (ppJ2)
[b]Number of expressed sequence tags (ESTs) downloaded from the dbEST division of GenBank on 10 September 2005
[c]Length in nucleotides (nt) of the EST and the contig consensus sequence

different reading frames. Ambiguities in the consensus sequence were translated into 'X' in the amino acid sequence. The longest of the six generated amino acid sequences, preferentially starting with a methionine, was retained in an output file. The minimum length of the translated polypeptide to be retained in the output file was set at 50 amino acids. Names are assigned to the polypeptide sequences in the SPIT-output file including the main characteristics of the amino acid sequence: (F)orward or (R)everse orientation, name of contig, length of EST and amino acid sequence, start (M)ethionine or (O)ther. These sequences were fed into SignalP 3.0 (http://www.cbs.dtu.dk/services/SignalP/; Bendtsen *et al.*, 2004) to predict the presence and cleavage site of signal peptides. A signal peptide was considered to be present when a sequence scored above the cut-off value for both NN and HMM. Putative signal peptide containing sequences were screened with trans-membrane hidden Markov models (TMHMM) v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) to exclude proteins with predicted transmembrane regions (Krogh *et al.*, 2001). Basic local alignment search tool (BLAST) analyses were performed on the *nr* (built October 2005) and est_others (non-mouse and non-human; built April 2006) databases at NCBI to identify sequence similarity (Altschul *et al.*, 1990). A WoLF PSORT analysis (http://wolfpsort.seq.cbrc.jp/) was performed to predict the subcellular localization of putatively secreted proteins based on their amino acid sequence minus the predicted signal peptide for secretion (Horton *et al.*, 2006).

Nematodes

Meloidogyne chitwoodi was propagated on tomato (*Lycopersicum esculentum*) cultivar Moneymaker in a greenhouse. Infected tomato roots were harvested and thoroughly but carefully rinsed with tap water to obtain clean roots with intact egg masses. These roots were cut into pieces and put in a mist chamber to hatch the nematodes at 19°C (Oostenbrink, 1960). Hatched nematodes were put on a double cotton wool

filter (Hygia Favorit II) to select living nematodes, which were collected after three to four hours. To check the species purity, we performed two PCRs with JMV primers on DNA obtained from the *M. chitwoodi* sample by phenol/chloroform extraction and ethanol precipitation (Wishart *et al.*, 2002). The JMV1 primer was used in combination with both JMV2 and JMV hapla to distinguish the three temperate species *M. chitwoodi*, *M. fallax* and *M. hapl*a. A separate PCR was performed with primers JMV1 and JMV tropical to detect the presence of the tropical species *M. arenaria*, *M. incognita*, *M. javanica* and *M. mayaguensis*.

In situ hybridization microscopy

Sense and anti-sense probes were generated from contigs encoding members of the secretome including the most abundant ESTs. The probes were obtained by linear amplification and labelled using digoxigenin-dUTP in the transcription reaction (Roche Diagnostics, Mannheim, Germany). Clones and primers used for amplification are listed in Table 2. Nematodes were fixed overnight in 2% (w/v) paraformaldehyde at 4°C, followed by 4 hours at room temperature and cut into two to five pieces. *In situ* hybridization procedure was performed as described (De Boer *et al.*, 1998). Overnight hybridization was done at 50°C. Nematodes were labelled with alkaline phosphatase-conjugated sheep anti-digoxigenin Fab fragments (Roche Diagnostics). Alkaline phosphatase activity was detected overnight at 4°C with nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate as substrates. The nematodes were examined using differential interference contrast microscopy (Leica, Deerfield).

Table 2 Oligonucleotide primers used to prepare probes from an EST member of a contig for *in situ* hybridization

Contig	EST	Forward primer (5' - 3')	Reverse primer (5' - 3')
MC01229	ri68b01	TTGCCCCAAACCCAAAAC	ACCGCCACCAGCCGAATG
MC01222	rj19b04	GAGTCCAGAGTATGCTAAAATAAC	TTTGCCTGATAACAATAACCATC
MC01220	ri62c03	CTGCTACACCCGCTTCAAC	AACAGAGACAGTCGTATCATCAC
MC01214	rj20g01	TTGGTTTTGGATGTTATTATTGCC	TTGTGGTCAAATCACCGATG
MC01209	ri72e01	ACTTATACATGCGATAACACTTGG	TCTTGGCTGTGGATGCTTAC
MC01206	ri43d12	GACGGCTCCCAACTTCAC	GGTAGGTTCTTCATCATCATCATC
MC01202	ri43d01	GTGAACCAACTAATGAAGAATGTG	GGTGAGTTTGTAAAGCATATACG
MC01203	ri82e04	AAGGATAATGGTGAAGTAGC	TTGAGGTCAAATAGTTGGTC
MC01194	ri66h08	TGGTCATCATCATCATACTAAAAG	AAGTTGTTGTTGAAGAAGTAGAAG
MC01180	ri38g03	TTCTATTTTGTCAAGGCAG	AACTTTGTCCTTTGCATGTC

RESULTS

Contig assembly

In total, 12,218 ESTs originating from three developmental stages of *M. chitwoodi* were used in this study (Table 1). Genes that were highly expressed at the time of sampling are represented by a higher number of ESTs. To reduce this redundancy in the dataset, all ESTs were assembled in contigs based on sequence similarity. ESTs within a contig appear to derive from (nearly) identical transcripts (McCarter *et al.*, 2003; Mitreva *et al.*, 2004b) and therefore contig-assembly is indicative of the level of gene expression. Consensus sequences thus created can have a higher quality and length compared to individual ESTs. Contig-assembly of the *M. chitwoodi* ESTs resulted in a slight increase of the mean sequence length and a 50% reduction of the total sequence length that has to be analyzed (Table 1). The assembly of 12,218 ESTs into contigs resulted in 4,720 singletons, i.e. contigs containing one EST, and 1,160 multi-member contigs (Fig. 1). The redundancy of the *M. chitwoodi* EST dataset, estimated by the number of reads that were assembled into multi-member contigs relative to the total number of reads, is approximately 61%. The percentage of the proteome of *M. chitwoodi* that is encoded by the current set of ESTs is calculated to be 25.2%. This percentage is based on the number of predicted proteins of *Caenorhabditis elegans*, which is estimated to be 23,375 (Wormpep database, release wormpep176).

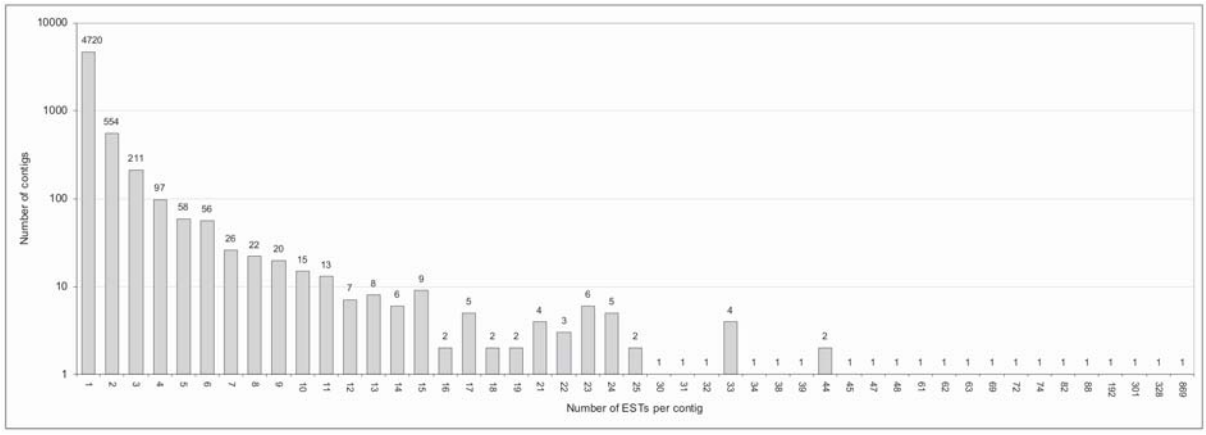


Fig. 1 Distribution of contigs by number of assembled ESTs per contig.

The secretome of *M. chitwoodi*

We carried out an *in silico* analysis of contig consensus sequences using a sequence of search algorithms, including SPIT, SignalP and TMHMM, in order to identify transcripts coding for secreted proteins (Fig. 2). The 1,160 contig consensus sequences

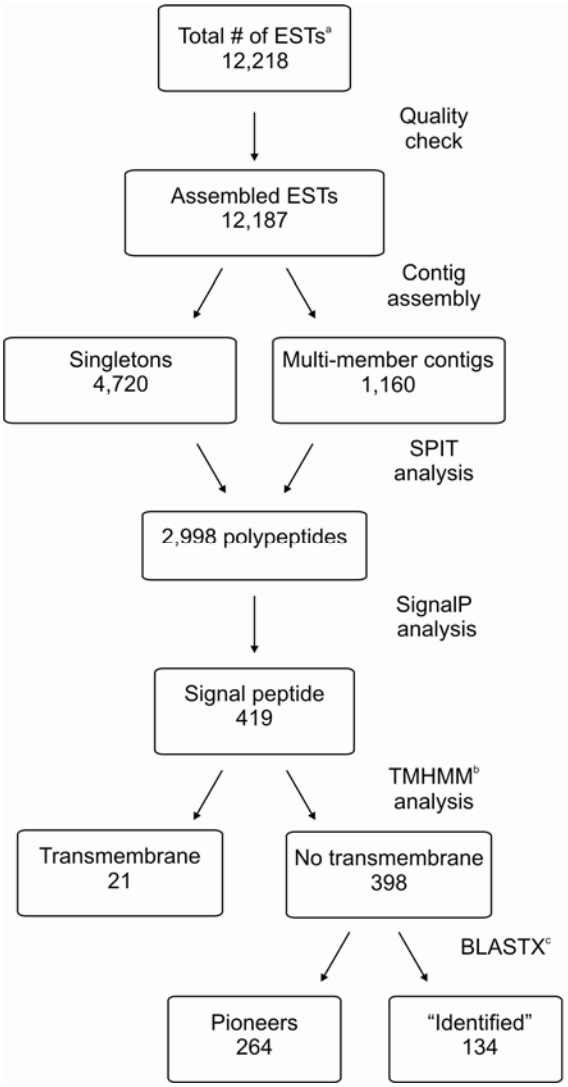


Fig. 2 Flowchart of the *in silico* analyses to identify members of the *Meloidogyne chitwoodi* secretome. The SPIT analysis was performed with high stringent selection criteria.
[a] Total number of ESTs sequenced by the Parasitic Nematode Sequencing Project
[b] TMHMM = trans-membrane hidden Markov models
[c] BLASTX performed on the non-redundant (*nr*) database at NCBI

and 4,720 singletons from the assembly were analyzed with the SPIT-algorithm resulting in 2,998 polypeptide sequences of at least 50 amino acids. These amino acid sequences were fed into the SignalP 3.0 neural networks (NN) and hidden Markov models (HMM) algorithms to search for signal peptides for protein secretion and cleavage sites. A total of 419 proteins was identified with D-score probability of 0.432 or higher in the NN algorithm, and with signal peptide probability of 0.402 or higher in the HMM algorithm. Of these, 21 sequences were excluded from further analysis since they also include a trans-membrane domain in the mature protein as predicted by TMHMM. The resulting dataset of 398 putatively secreted proteins from *M. chitwoodi* was used as search query in BLASTX on the non-redundant (*nr*) database to find significantly matching sequences. For 264 proteins, i.e. 66%, no significant

similarity was found to other known proteins in the *nr* database (E-value cut-off 0.001). The remaining 134 proteins do have a significant similarity to known sequences (E-value < 0.001). Among the latter sequences are secretory proteins previously identified in the oesophageal glands, amphids, and hypodermis of other plant parasitic nematode species, e.g. cell wall modifying enzymes, expansins, venom allergen homologs, SXP/RAL-2 protein, the putative avirulence protein MAP-1, SEC-2 protein and cuticle collagens.

Analysis of highly expressed members of the secretome

Sequence similarity

Among the top ten most abundantly represented transcripts in the dataset of 12,218 ESTs are four likely members of the secretome identified by our *in silico* analysis: MC01229, MC01226, MC01224, and MC01222 (Table 3). Two other sequences in this top ten, MC01225 and MC01221, are also likely to encode secretory proteins based on their similarity with known proteins. However, sequences MC01225 and MC01221 were not retrieved by our algorithms. The matching sequence of MC01225, SEC-2 (also named Gp-FAR-1), is a retinol- and fatty acid-binding protein from *Globodera pallida* (Prior *et al.*, 2001). A closer look at the MC01225 sequence and a pair-wise alignment with SEC-2 generated in BLASTX shows that part of the 5' un-translated region of MC01225 is included in the open reading frame due to a frame shift. As a result, the signal peptide is not recognized with sufficient probability in SignalP. In the case of MC01221, which matches the MAP-1 gene from *M. incognita* (Semblat *et al.*, 2001), the contig is incomplete at its 5'-end and the encoded open reading frame lacks amino acids at the amino terminus.

Although most of the abundantly expressed members of the secretome do not have evident homologs in the *nr* database (Table 4), MC01222 shows significant similarity to the oesophageal gland secretory protein MSP37 from *M. incognita* (Huang *et al.*, 2004a). Similarly, a conserved domain search with sequence MC01209 showed similarity with C-type lectin or carbohydrate-recognition domain (CD00037 and SMART00034; Marchler-Bauer and Bryant, 2004). In addition, a BLASTX analysis using MC01209 showed similarity to several C-type lectins from snake venoms. Among the latter lectins are mucrocetin and trimecetin from *Protobothrops mucrosquamatus* (synonym *Trimeresurus mucrosquamatus*) involved in platelet agglutination (Genbank accession numbers AAQ93687 and AAW69916; Chiou *et al.*, 1996; Huang *et al.*, 2004b). Weak similarity was also found with the N-terminal C-type lectin domain of SCN1018, which is expressed in the hypodermis of *Heterodera glycines* females (Genbank accession number AY043250; De Boer *et al.*, 2002b). Also for sequence MC01185, a match was found with a C-type lectin domain, and a PSI-

Table 3 Top 10 most abundantly represented transcripts in the *M. chitwoodi* EST dataset. Underlined contigs encode putatively secreted proteins identified by our *in silico* analysis. BLASTX hit was considered significant when E-value ≤ 0.001

	Contig	#EST[a]	nt[b]	BLASTX top hit	E-value
1	MC01230	869	1746	unnamed protein product XP_453836 (*Kluyveromyces lactis*)	2e-17
2	MC01229	328	572	-	-
3	MC01228	301	555	-	-
4	MC01227	192	848	-	-
5	MC01226	88	354	-	-
6	MC01225	82	664	SEC-2 protein Y09293 (*G. pallida*)	2e-30
7	MC01224	74	466	-	-
8	MC01223	72	674	unnamed protein product XP_453836 (*K. lactis*)	4e-18
9	MC01222	69	720	putative oesophageal gland cell secretory protein 37 AY422830 (*M. incognita*)	5e-17
10	MC01221	63	690	putative avirulence protein CAC27774 (*M. incognita*)	3e-30

[a]Number of ESTs assembled in a contig
[b]Length of contig consensus sequence in nucleotides (nt)

BLAST hit upon a C-type lectin (clec-11) from *C. elegans* (Genbank accession number NP_493162; unpublished). Sequences MC01203 and MC01180 show similarity to β-1,4-endoglucanases and pectate lyases respectively, which are cell wall degrading enzymes also identified in other plant parasitic nematode species. These enzymes were found to be expressed in the subventral oesophageal glands and seem to facilitate the migration through the root system (Popeijus *et al.*, 2000b; Smant *et al.*, 1998). Sequence MC01184 shows similarity to venom allergen-like proteins expressed in the subventral oesophageal glands of *H. glycines*, *Globodera rostochiensis*, and *M. incognita* (Ding *et al.*, 2000; Gao *et al.*, 2001).

M. chitwoodi specific sequences

Thirteen of the 20 most abundant members of the secretome do not have significant similarity to other known proteins in the *nr* database. BLASTN and TBLASTX searches on the est_others database (E-value cut-off 0.001) were carried out utilizing these twenty sequences to investigate if they are specific for *M. chitwoodi*. For MC01229, MC01226, MC01214, MC01210, MC01206, MC01201, and MC01194, similarity was only found with *M. chitwoodi* ESTs indicating that they are specific for this species, with a possibility to shift to any of the other two groups (described below) as more sequence data become available. For sequences MC01224, MC01222, MC01220, MC01209, MC01203, MC01202, MC01198, MC01196, MC01185, and MC01177, similarity was found with ESTs from other root-knot nematode species, including *M. hapla*, *M. incognita*, *M. arenaria*, and *M. javanica*. For sequences MC01207, MC01184, and MC01180, the BLAST searches identified similarity to transcripts from cyst nematode species *H. glycines* and *H. schachtii*, in addition to ESTs from root-knot nematode species.

Table 4 Top 20 most abundantly represented putative secretome members of *M. chitwoodi*. The signal peptide (SP) cleavage site was predicted by the SignalP 3.0 software. BLASTX hit was considered significant when E-value ≤ 0.001. Contigs with an asterisk encode mature secretory proteins with a nuclear localization signal predicted by WoLF PSORT

	Contig	#EST[a] (stage)	nt[b]	aa[c]	start[d]	SP cleavage site	BLASTX top hit	E-value	ISH labelling[e]
1	MC01229*	328 (ppJ2)	572	76	M	20-21	-	-	DG
2	MC01226	88 (female)	354	67	M	18-19	-	-	n.d.
3	MC01224	74 (female)	466	105	M	26-27	-	-	n.d.
4	MC01222	67 (ppJ2) 2 (egg)	719	178	M	18-19	putative oesophageal gland cell secretory protein 37 AY422830 (*M. incognita*)	4e-25	SvG
5	MC01220	62 (ppJ2)	912	275	M	21-22	-	-	intestine
6	MC01214*	44 (ppJ2)	483	97	M	21-22	-	-	SvG
8	MC01210	33 (female)	386	66	M	19-20	-	-	n.d.
7	MC01209	32 (ppJ2) 1(egg)	578	179	M	18-19	mucrocetin beta chain AAQ93687 (*P. mucrosquamatus*)	0.001	SvG
9	MC01207	33 (egg)	618	189	M	21-22	cysteine protease related protein 4 AAA92327 (*C. elegans*)	5e-15	n.d.
10	MC01206*	32 (ppJ2)	642	205	M	21-22	-	-	SvG
12	MC01203	25 (ppJ2)	682	211	M	21-22	beta-1,4-endoglucanase 1 precursor CAC12958 (*H. schachtii*)	3e-10	SvG
11	MC01202	24 (ppJ2) 1 (egg)	789	146	M	22-23	-	-	tail tip region
14	MC01201	24 (female)	343	67	M	18-19	-	-	n.d.
13	MC01198	24 (ppJ2)	649	185	M	21-22	-	-	intestine
15	MC01196	23 (egg)	878	290	M	34-35	-	-	n.d.
16	MC01194	23 (ppJ2)	554	95	M	22-23	-	-	SvG
17	MC01185	21 (ppJ2)	581	167	M	23-24	-	-	n.d.
18	MC01184	21 (ppJ2)	568	171	M	21-22	vap-1 AAK60209 (*H. glycines*)	2e-19	n.d.
19	MC01180	15 (ppJ2) 3 (egg)	566	178	M	18-19	pectate lyase 1 AAS88579 (*M. incognita*)	2e-42	SvG
20	MC01177	9 (egg) 8 (female)	568	131	O	18-19	hypothetical protein C16C10.11 CAA86749 (*C. elegans*)	7e-18	n.d.

[a]Number of ESTs assembled in a contig
[b]Length of contig consensus sequence in nucleotides (nt)
[c]Length of the polypeptide sequence in amino acids (aa) generated by the SPIT program and retained in the output file
[d]First amino acid of the polypeptide sequence which is either a methionine (M) or other (O) amino acid
[e]Labelling pattern observed with *in situ* hybridization microscopy (ISH). DG = dorsal oesophageal gland, SvG = subventral oesophageal glands, n.d. = not determined

Identification and putative nuclear targeting of candidate parasitism genes

Parasitism genes are expressed in organs and tissue capable of releasing molecules to the interface of host and parasite, i.e. oesophageal, amphidial, and rectal glands, intestine and hypodermis. *In situ* hybridization microscopy was performed to localize the site of expression of the most abundantly represented transcripts from the secretome of *M. chitwoodi*. These analyses were only performed on transcripts from the *M. chitwoodi* ppJ2 stage and not on transcripts exclusively represented in either the adult female or the egg stage. Transcripts exclusively represented in the egg stage are not expected to play a direct role in the nematode-plant interaction and are thus considered less important for our understanding of parasitism. Similarly, transcripts specific for adult females are more likely to be involved in reproduction than in feeding. Eleven putative members of the secretome were tested on whole mount nematode sections (Table 4). One of these specifically and intensely labelled the dorsal oesophageal gland and seven labelled the subventral oesophageal glands. Two probes hybridised to the intestine and one to the tail tip in the region where the phasmids and rectal glands are located (Fig. 3). In case of three oesophageal gland specific parasitism genes, MC01229, MC01214, and MC01206, the predicted subcellular localization in host cells following cleavage of the leader peptide for protein secretion was nuclear (Table 4).

DISCUSSION

In a straightforward *in silico* approach we generated the first comprehensive overview of the secretome from the plant-parasite *M. chitwoodi*. From a dataset of 12,218 ESTs we identified 398 members of the nematode's secretome (Fig. 2). Two-third of these 398 sequences has no significant similarity to other known proteins in the *nr* database, which complicates the deduction of possible functions of these members of the secretome and their involvement in the host-parasite interaction. Genes encoding secretory proteins in the oesophageal, amphidial and rectal glands, intestine, and hypodermis are considered candidate parasitism genes (Davis *et al.*, 2000; Hussey and Mims, 1990). *In situ* hybridization microscopy with a selection of the most abundant members of the secretome resulted in the identification of seven subventral oesophageal gland specific genes, one dorsal gland specific gene, two intestine specific genes, and one gene specific for the tail tip region.

The finding that at least eight of the twenty most abundantly represented members of the secretome are specifically expressed in the single-celled oesophageal

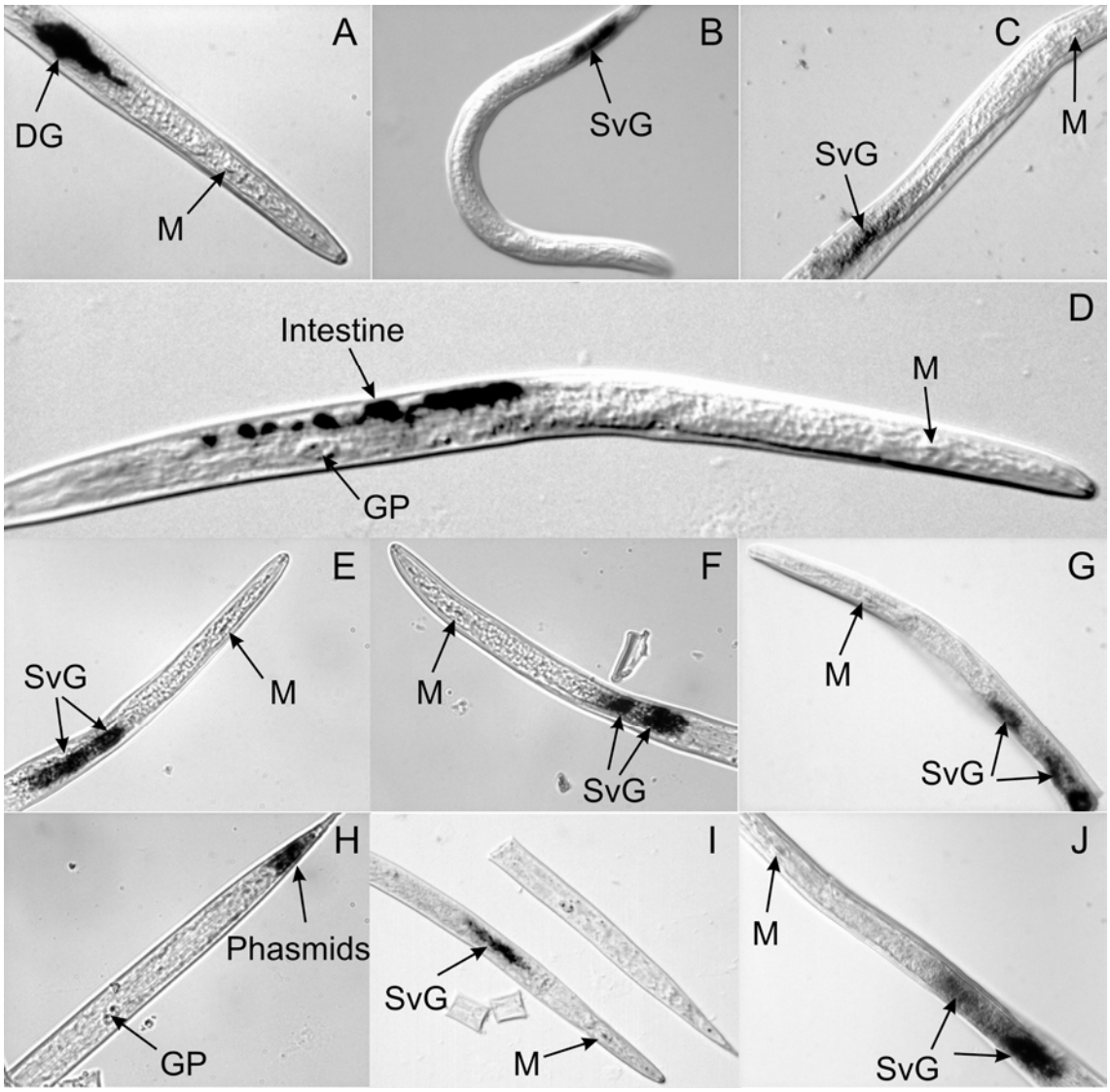


Fig. 3 *In situ* hybridizations of the most abundantly expressed members of the secretome in pre-parasitic second-stage juveniles of *Meloidogyne chitwoodi*. Magnification between brackets. Sections of the nematode incubated with antisense probe designed on contigs: **A** MC01229 (400x), **B** MC01222 (400x), **C** MC01214 (400x), **D** MC01220 (160x), **E** MC01209 (400x), **F** MC01206 (400x), **G** MC01203 (400x), **H** MC01202 (400x), **I** MC01194 (160x), **J** MC01180 (400x). DG = Dorsal Gland, M = Metacorpus, SvG = Subventral Glands, GP = Genital Primordium.

glands is striking. This relative abundance of oesophageal gland genes may be explained by the fact that these glands are extraordinarily large cells that constitute a major portion of the nematode's oesophageal region (Hussey and Mims, 1990). Furthermore, transcription in each of these gland cells, which are packed with secretory granules, is dedicated to the expression of members of the secretome (Hussey and Mims, 1990). Thus, even in whole nematode cDNA libraries the chance of generating tags from oesophageal gland specific genes is relatively high.

Two of the subventral oesophageal gland specific genes, MC01203 and MC01180, show similarity to β-1,4-endoglucanases and pectate lyases respectively. These cell

wall degrading enzymes were previously identified in other plant parasitic nematodes and facilitate their migration through plant roots (Huang *et al.*, 2005; Popeijus *et al.*, 2000b; Rosso *et al.*, 1999; Smant *et al.*, 1998; Wang *et al.*, 1999). Another subventral oesophageal glands specific gene, MC01222, has significant similarity with putative oesophageal gland cell secretory protein MSP37 from *M. incognita* (Huang *et al.*, 2004a). MSP37 is a pioneering sequence, and was shown to be specific for the dorsal oesophageal gland cell in parasitic J2-s and later stages. No expression was observed in pre-parasitic J2-s of *M. incognita*. Interestingly, our sequence sharing similarity with this protein was expressed in the subventral oesophageal gland cells of *M. chitwoodi* pre-parasitic J2-s. The subventral oesophageal glands specific MC01209 shows similarity with several C-type lectins from snake venoms and weak similarity with the N-terminal C-type lectin domain of SCN1018. The latter gene is reported to be up-regulated in parasitic life stages and expressed in the hypodermis of *H. glycines* females (De Boer *et al.*, 2002b). De Boer *et al*. (2002b) suggested a role in antibacterial defence for this gene in *H. glycines*.

No significant similarity was found for the two intestine and the tail tip specific sequences (Table 4). In addition, the remaining three subventral and the dorsal oesophageal gland specific genes do not have significant similarity with sequences in the *nr* database and the *est_others* database. These oesophageal gland genes are thus considered to be pioneers and specific for *M. chitwoodi*. Other studies on nematode parasitism genes have shown similar percentages (65-75%, E-value cut-off 0.005) of pioneers in their samples (Gao *et al.*, 2003; Huang *et al.*, 2003). The vast genetic diversity at loci involved in parasitism may reflect an evolutionary activity resulting from forces that aim to improve the capacity to modify host cells, while at the same time avoid recognition by specific host defence systems (Harcus *et al.*, 2004).

We investigated possible targets of the pioneer genes with subcellular localization predictions of the mature proteins (with the signal peptide cleaved off). We found that the subcellular target in host cells of the mature products of the pioneer oesophageal gland specific secretome members, except for MC01194, is predicted to be nuclear. This suggests that, once injected into the host cells, these proteins are directed to the plant cell nucleus. A C-terminal green fluorescent protein (GFP) fusion of Hs-UBI1, a parasitism protein from the plant-parasitic nematode *H. schachtii*, showed a nucleolar localization of this parasite protein in plant cells (Tytgat *et al.*, 2004). Nuclear targeting was also observed with three parasitism genes from *H. glycines* suggesting that the nucleus may be an important target for secreted parasitism gene products containing a nuclear localization signal (Davis *et al.*, 2004).

We have used the ESTs generated in the Parasitic Nematode Sequencing Project (www.nematode.net) from whole nematode cDNA libraries to identify members of the secretome. To date, a range of pre-selection methods have been successfully

deployed to enrich libraries for genes encoding extracellular proteins and for specific tissues. These pre-selections were based on tissue specific hybridizations (e.g. suppressive subtractive hybridization; Huang *et al.*, 2004a), micro aspiration of mRNA from specific regions in the parasite (Gao *et al.*, 2003; Huang *et al.*, 2003; Wang *et al.*, 2001), signal sequence trapping (Wang *et al.*, 2001) and differential expression patterns (e.g. cDNA-AFLP and micro arrays; De Boer *et al.*, 2002b; Qin *et al.*, 2000). Because these pre-selection methods require significant technological effort and skills that need to be adapted for each parasite, we chose to use the whole nematode libraries to start the molecular exploration of *M. chitwoodi*. Secondly, we did not want to rely on assumptions on the importance of specific organs in parasitism. The consequence of using whole nematode cDNA libraries, without pre-selection, is that our strategy is most effective with larger number of ESTs, which are now being made available by the Parasitic Nematode Sequencing Project.

The primary criterion in our selection strategy, i.e. the presence of a signal peptide for protein secretion, relies on predictions of protein sorting motifs in amino acid sequences. This implicates a strong dependency on i) confidence levels of the prediction algorithms and ii) the proportion of full-length cDNAs in the libraries. The performance of the prediction software is subject of ongoing improvements by combining different prediction algorithms and output scores, and validating reference datasets (Bendtsen *et al.*, 2004). Primers designed on the 5'-end *trans*-spliced leader (SL) sequences have been used to enrich cDNA libraries for full-length cDNA molecules by PCR (Mitreva *et al.*, 2004a). The majority of the most recent ESTs sequenced by the Parasitic Nematode Sequencing Project are generated from cDNA libraries enriched for full-length cDNA inserts using the SMART™ technology (Zhu *et al.*, 2001). The *M. chitwoodi* ESTs analyzed in our study are sequenced from cDNA libraries enriched for full-length sequences using both approaches (Table 1).

An example of a likely false negative in the output of our sequence of algorithms due to incomplete 5'-end cDNA sequence in the library is MC01221. The gene shows similarity to the C-terminal part, i.e. amino acids 305 to 436, of the amphid-specific putative avirulence protein MAP-1 from *M. incognita* (Semblat *et al.*, 2001). The open reading frame of MAP-1 starts with a classical leader peptide for protein secretion. Because of its presence only in strains of *M. incognita* that are avirulent on the host plants harbouring the *Mi-1* resistance gene, this nematode gene is believed to be involved in parasite recognition by the defence systems of the host (Semblat *et al.*, 2001). Sequence MC01221 also shows similarity to the C-terminal expansin catalytic domain of EXPB1 and EXPB2 from the potato cyst nematode *G. rostochiensis*; EXPB1 is secreted by infective juveniles of *G. rostochiensis* (Kudla *et al.*, 2005; Qin *et al.*, 2004).

As a starting point for our analysis, we used consensus sequences of contigs assembled from ESTs. A contig gives a good indication of the abundance of a

transcript at the time of sampling because the ESTs within a contig are derived from (nearly) identical transcripts (McCarter *et al.*, 2003). To go from transcript level to gene level, contigs are frequently being grouped in clusters. The interpretation of a cluster, however, is not straightforward. Contigs within a cluster may represent different alleles, transcript splice isoforms, or closely related gene family members (McCarter *et al.*, 2003). It is thus not clear whether a cluster represents a single gene, a gene family, or a mixture of genes with homologous domains and, therefore, we did not use the clusters in our analysis.

By combining *in silico* analyses with *in situ* hybridizations, we identified the first candidate parasitism genes from *M. chitwoodi*. Our strategy to identify secreted proteins and the products of parasitism genes within the full repertoire of secreted proteins forms the basis for the functional analysis of parasite proteins involved in the nematode-host interaction. We believe that this methodology for initial molecular exploration of parasites provides an unbiased view on the nature of the proteins involved in parasitism. Furthermore, as the approach does not require technologically advanced pre-selection procedures, it can be widely applied to animal- and plant-parasites, and beyond.

ACKNOWLEDGEMENTS

We would like to thank G. Karssen from the Plant Protection Service and Laboratory of Nematology, Wageningen (NL), for supplying us with *M. chitwoodi* nematodes which we used in *in situ* hybridizations, J. Wishart from the Scottish Crop Research Institute for the JMV primers and David Messina from the Washington University for technical assistance. The work was supported by European Union grant DREAM (QLRT-1999-1462) and in part by NIH-NIAID research grant AI 46593.

3

The repertoire of cell wall modifying proteins from the root-knot nematode *Meloidogyne chitwoodi*

Erwin Roze, Biljana Stojanova, Hein Overmars, Makedonka Mitreva,
Aska Goverse, Jaap Bakker and Geert Smant

Summary

Meloidogyne chitwoodi is a root-knot nematode species with a wide range of host plants, including both monocotyledons and dicotyledons. The polymer constitution of the cell walls in this wide range of host plants varies extensively. *M. chitwoodi* penetrates and migrates through the host plant roots to finally establish a permanent feeding site inside the root system. The migration phase takes entirely place in the apoplast and we therefore hypothesized that *M. chitwoodi* has to match the polymer diversity in plant cell walls with a repertoire of cell wall modifying proteins to be able to successfully invade its hosts. To study the genes expressed in the penetration phase of the nematode-plant interaction we have undertaken an expressed sequence tag (EST) project with the invasive second-stage juveniles (J2-s) of *M. chitwoodi*. In total, 4,067 ESTs were produced and screened for cell wall modifying protein encoding sequences. This resulted in the identification of an elaborate repertoire of enzymes that targets essentially all major cell wall components. It includes β-1,4-endoglucanases, a β-1,4-endoxylanase, pectate lyases, polygalacturonases and a cellulose binding protein. The subventral oesophageal gland specific expression and presence of predicted secretion signal peptides suggest that these cell wall modifying proteins are secreted from the nematode during invasion of the plant root. Characteristics of each (group of) cell wall modifying proteins identified from *M. chitwoodi* are described and discussed, including their possible roles during the intercellular migration phase and their impact on the host plant range of *M. chitwoodi*.

Introduction

The plant cell wall is a highly dynamic and organized composite of diverse polysaccharides, proteins and aromatic substances, of which the polysaccharides are the main components (Cosgrove, 1997a). The plant cell wall polymer arrangement and composition differs between plant species, tissues and cells (Carpita and McCann, 2000). The scaffolding structure in the cell wall is formed by a network of cellulose microfibrils, which are interlocked by cross-linking glycans, i.e. hemicelluloses. Two major cross-linking glycans are the xyloglucans and glucuronoarabinoxylans. Xyloglucans cross-link polymers in the walls of all dicots and the non-commelinoid monocots, whereas glucuronoarabinoxylans are the major cross-linking glycans in the commelinoid line of monocots. Cereals and grasses, commelinoid species of the order Poales, contain 'mixed-linkage'(1→3), (1→4) β-D-glucans which is the third type of major cross-linking glycan (Carpita and Gibeaut, 1993). The cellulose/hemicellulose network is embedded in a matrix of heterogeneous pectic polysaccharides. Pectins are a diverse group of polysaccharides rich in galacturonic acid with a variable level of methyl esterification (Tamaru and Doi, 2001).

The functions of plant cell walls are as diverse as their composition. Plant cell walls are e.g. responsible for tensile strength, flexibility, porosity, intercellular adhesion, cell shape, and defence against invading organisms (Cosgrove, 1997a). The latter may be achieved by increased deposition of structural polymers like lignin and callose or oligosaccharides cleaved from pectic polysaccharides can elicit defence responses (Bowles, 1990; Esquerré-Tugayé *et al.*, 2000).

Root-knot (*Meloidogyne* spp.) and cyst (*Heterodera* and *Globodera* spp.) nematodes are sedentary endoparasites of plant roots. In contrast to cyst nematodes, root-knot nematodes have a very broad host plant range including thousands of different plant species (Trudgill and Blok, 2001). Based on their host plants in the field, *Meloidogyne* species can be grouped into three groups; species mainly parasitizing i) dicotyledons, ii) monocotyledons, or iii) both mono- and dicotyledons (Karssen, 1999). *M. chitwoodi* belongs to the group parasitizing on mono- and dicotyledons, which includes the most successful species of the *Meloidogyne* genus. Among the host plants of *M. chitwoodi* are many economically important crop species like tomato, potato, sugar beet, maize and wheat (Den Nijs *et al.*, 2004; O'bannon, 1982; Santo *et al.*, 1980). For its polyphagous nature, high damage probability and potential to establish widely within Europe, *M. chitwoodi* has a quarantine organism status (Den Nijs *et al.*, 2004).

Both cyst and root-knot nematodes establish their feeding site, called syncytium and giant cell respectively, in the vascular cylinder of the host plant root (Williamson

and Gleason, 2003). A remarkable difference between cyst and root-knot nematodes is their mode of migration in the plant. Cyst nematodes migrate intracellularly through the cortex in a more or less straight line, leaving a track of necrotic cells. Root-knot nematodes on the other hand migrate intercellularly down the cortex towards the root tip. In the root tip they make a U-turn, enter the base of the vascular cylinder and migrate up the root (Wyss *et al.*, 1992). Amazingly, this extensive intercellular migration does not inflict any detectable damage to the cells along the migratory track (Hansen *et al.*, 1996).

During plant penetration and migration through the plant root system, cyst and root-knot nematodes are confronted with the plant cell walls. To breach this barrier, the infective second-stage juveniles (J2-s) of these nematodes secrete a mixture of cell wall degrading enzymes from their stylet, a hollow protrusible feeding spear, combined with stylet thrusts (Davis *et al.*, 2000; Davis *et al.*, 2004). It has been discovered that plant-parasitic nematodes use cellulases (De Meutter *et al.*, 1998; Gao *et al.*, 2002; Goellner *et al.*, 2000; Kikuchi *et al.*, 2004; Rosso *et al.*, 1999; Smant *et al.*, 1998; Uehara *et al.*, 2001; Yan *et al.*, 2001), xylanase (Mitreva-Dautova *et al.*, 2006), pectate lyases (De Boer *et al.*, 2002a; Doyle and Lambert, 2002; Huang *et al.*, 2005; Kikuchi *et al.*, 2006; Popeijus *et al.*, 2000b), and exo-polygalacturonase (Jaubert *et al.*, 2002) to facilitate the invasion of the host plant root system.

The polymer constitution of the cell walls in the wide range of host plants of *Meloidogyne chitwoodi* varies extensively. Combined with the fact that the migration through the roots takes entirely place in the apoplast this indicates that great flexibility is required in the substrate specificity of the enzyme complex in the stylet secretions of this nematode species. To study the genes expressed in the initial phases of nematode-plant interaction we have undertaken an expressed sequence tag (EST) project with the invasive J2 stage of *M. chitwoodi.* We searched the ESTs for cell wall modifying protein (CWMP) encoding sequences, which resulted in the identification of β-1,4-endoglucanases, a β-1,4-endoxylanase, pectate lyases, polygalacturonases and a cellulose binding protein. The latter represents a novel putative activity in nematode stylet secretions. Together with the expansin-like proteins described in Chapter 4, these CWMPs form the most elaborate repertoire found in a single plant-parasitic nematode species so far. All CWMPs were found to be expressed in the subventral secretory glands and together they target all major cell wall components. Our results also show that random sequencing of a cDNA library is a powerful method to identify genes involved in parasitism in a relatively unexplored nematode species.

Materials and methods

M. chitwoodi pre-parasitic second stage juveniles (ppJ2-s) cDNA-library construction

Eggs of *M. chitwoodi* isolate Co (Zijlstra, 2000) were separated from egg masses by NaOCl treatment and cleaned over a sucrose gradient (McClure *et al.*, 1973). The sucrose was removed by washing the eggs with sterile tap water on a 38 micron sieve. The eggs were subsequently layed on a 25 micron sieve, covered with sterile tap water for hatching at 19.5°C in the dark. Hatched ppJ2-s were collected in sterile low adhesion tubes (BIOzym, Landgraaf, The Netherlands) and stored at -80°C until further use. To isolate RNA, the nematodes were mixed with TRIzol reagent (Gibco BRL, Life Technologies), frozen in liquid nitrogen and homogenised using a motor-driven pestle (Heidolph Instruments, Schwabach, Germany). Total RNA was isolated following the TRIzol protocol (Gibco BRL, Life Technologies) and dissolved in di-ethyl pyrocarbonate (DEPC)-treated water. The quality of the RNA was checked by agarose gel electrophoresis. Purification of mRNA from total RNA was performed using GenoPrep mRNA beads according to the manufacturers' instructions (GenoVision, Philadelphia, USA).

We performed two PCR amplifications with nematode species-specific JMV primers on DNA recovered from the TRIzol interphase during the RNA isolation to check the species purity of the sample (Wishart *et al.*, 2002). The JMV1 primer was used in combination with both primers JMV2 and JMV-hapla to discriminate the three temperate root-knot nematode species *M. chitwoodi*, *M. fallax* and *M. hapl*a (Adam *et al.*, 2007). A second PCR was done with the primers JMV1 and JMV-tropical to test for the presence of the tropical species *M. arenaria*, *M. incognita*, *M. javanica* and *M. mayaguensis*.

Messenger RNA was used to synthesise cDNA following the long-distance (LD) PCR protocol of the Creator™ SMART™ cDNA Library Construction kit (Clontech, Palo Alto, CA, USA). Twenty-two cycles were run to amplify the cDNA by LD PCR using primers including *Sfi I* restriction sites. Following *Sfi* I digestion and size fractionation, the SMART cDNA was directionally ligated to the pDNR-LIB vector and used to transform electrocompetent TOP10 *E. coli* (Invitrogen, Breda, The Netherlands).

M. chitwoodi ppJ2-s EST sequencing and screening for cell wall modifying proteins

In total, 4,067 5'-end expressed sequence tags (ESTs) were produced by sequencing random clones from the *M. chitwoodi* ppJ2-s full length enriched, directionally cloned

cDNA library. Clone handling and sequencing were performed as described (Hillier *et al.*, 1996; McCarter *et al.*, 2003). Prior to submission to the Expressed Sequence Tags database at the National Center for Biotechnology Information (dbEST; http://www.ncbi.nlm.nih.gov/projects/dbEST), the ESTs were processed to assess quality, to clean the sequences and to identify similarities with the basic local alignment search tool (BLAST) (Hillier *et al.*, 1996). Sequence trace files and clone request information are available at http://www.nematode.net.

The complete *M. chitwoodi* ppJ2-s EST dataset was screened for tags with similarity to proteins involved in cell wall modification by annotation search and using Mi-ENG-1 (AAK21882), MI-CBP-1 (AAC05133), Mi-XYL-1 (AAF37276), MI-PEL-1 (AAQ09004), and MI-PG-1 (AAM28240) as queries in a TBLASTN search. The ESTs encoding putative CWMPs were assembled into contigs based on sequence similarity.

Generation of full-length sequences

Full-length sequences encoding CWMPs, including the polyA tail, were obtained by either additional 3′-end sequencing of the original library clones or corresponding full-length sequences were amplified and sequenced from the original cDNA library. The 5′-end of *Mc-xyl-1* was amplified with the library vector specific primer pDNR-LIB-for (5′-CATTATACGAAGTTATCAGTC-3′) in combination with primer Mc-xyl-1-rev (5′-AACCATAAGCATGTCTAGC-3′), followed by the complete ORF with primers Xyl-CD418879ORF-for (5′-AGTTTCTAACCAATTTAATATA-3′) and Xyl-CD418879ORF-rev (5′-CAACAACAAATTTTACAATATTC-3′). Both the 5′-end and ORF of *Mc-xyl-1* were cloned in pCR4-TOPO and transformed into TOP10 *E. coli* for sequencing (Invitrogen, Breda, The Netherlands). Sequencing of clones was done with the vector specific primers M13for (5′-TGTAAAACGACGGCCAGT-3′) and M13rev (5′-CAGGAAACAGCTATGACC-3′) and insert specific primers, which are listed in Table 1.

Analysis of full-length sequences

BLAST analyses were performed on the non-redundant (*nr*) and est_others (non-mouse and non-human) databases at NCBI to identify sequence similarity (Altschul *et al.*, 1990). The protein sequences encoded by the full-length sequences were fed into SignalP 3.0 to predict the presence and cleavage site of signal peptides for secretion (http://www.cbs.dtu.dk/services/SignalP/; Bendtsen *et al.*, 2004). A ScanProsite search was performed to screen mature protein sequences for N-glycosylation consensus sites (http://www.expasy.org/tools/scanprosite/?PS00001). Conserved domain searches

Table 1 Insert specific oligonucleotide primers used for sequencing of cell wall modifying protein (CWMP) encoding sequences from *M. chitwoodi*

CWMP	Primer name	Primer sequence (5' - 3')
Mc-eng-1	Mc-eng-1-for	TCATAATGCAACAAGTCAAG
Mc-eng-2	Mc-eng-2-for	TCACACAAATAATGCTCCAG
Mc-xyl-1	Mc-xyl-1-rev	AACCATAAGCATGTCTAGC
Mc-pel-1	Mc-pel-1-for	TTCTATTTTGTCAAGGCAG
	Mc-pel-1-rev	AACTTTGTCCTTTGCATGTC
Mc-pg-1	Mc-pg-1-for1	TCAAGTAGCAAGTGGAGGTTC
	Mc-pg-1-for2	AGTGTTCGTTCAGTTGATTC
	Mc-pg-1-rev1	CGATTGTTAATTACTGAAATG
	Mc-pg-1-rev2	CAACACTAACATCCATAACAG

were performed with the mature protein sequences at NCBI and InterProScan (Marchler-Bauer *et al.*, 2007) (http://www.ebi.ac.uk/InterProScan/).

Identification of *Meloidogyne* β-1,4-endoglucanases

The clustered *Meloidogyne arenaria*, *M. hapla*, *M. javanica* and *M. paranaensis* ESTs (Nematode Net, NemaGene Clustering January 2007) were screened for β-1,4-endoglucanase encoding sequences. A NemaGene Cluster TBLASTN was performed against the EST contigs of these *Meloidogyne* species on www.nematode.net with Mc-ENG-2 minus signal peptide as query. The contig consensus sequences producing High-scoring Segment Pairs (HSPs) were selected for phylogenetic analysis.

Phylogenetic analysis

The protein sequences showing significant similarity to β-1,4-endoglucanase and pectate lyase genes in *M. chitwoodi* were aligned using the ClustalW algorithm in BioEdit version 6.0.7 (Isis Pharmaceuticals, CA, USA). The putative signal peptide sequences as predicted by SignalP-3.0 were removed and the initial alignment was further optimized manually. In the case of the cellulases the final alignment included 340 amino acid sites, while for the pectate lyases we could identify 180 homologous positions in the alignment. Four Bayesian runs (MrBayes, version 3.1) were performed with random trees as starting point using the BLOSUM substitution model with each alignment. Each run spanned 1 million generations, sampling took place every 100th generation, and the first 75,000 generations were discarded as burn-in. In the resulting unrooted phylogenetic tree, nodes with a posterior probability lower than 0.95 are considered to be unresolved.

Nematodes

Meloidogyne chitwoodi was propagated on tomato (*Lycopersicum esculentum*) cultivar Moneymaker in a greenhouse. Infected tomato roots were harvested and thoroughly

Table 2 Oligonucleotide primers used to prepare probes for *in situ* hybridization from an EST member of a contig encoding a cell wall modifying protein (CWMP)

CWMP	EST	Forward primer (5' - 3')	Reverse primer (5' - 3')
Mc-eng-3	ri38h05	GTTCAAATGGACAAGTTGC	AGTATATGCAGGTTCATTCC
Mc-cbp-1	ri64a10	AAATGAATTCTTTATTTTACG	AAACTGAACCAGAACCTCC
Mc-xyl-1	ri78d08	AGATGCATTTGCCAGAGTTG	AACTGAGCCATAACATAAGC
Mc-pel-1	ri38g03	TTCTATTTTGTCAAGGCAG	AACTTTGTCCTTTGCATGTC
Mc-pg-1	ri63g10	GTTAAAGCAGAGGAAGGTG	CCATTAGAATCAACTGAACG

but carefully rinsed with tap water to obtain clean roots with intact egg masses. These roots were cut into pieces and put in a mist chamber to hatch the nematodes at 19°C (Oostenbrink, 1960). Hatched nematodes were put on a double cotton wool filter (Hygia Favorit II) to select living nematodes, which were collected after three to four hours.

In situ hybridization microscopy

Sense and anti-sense probes were generated from clones encoding CWMPs. The probes were obtained by linear amplification and labelled using digoxigenin-dUTP in the transcription reaction (Roche Diagnostics, Mannheim, Germany). Clones and primers used for amplification are listed in Table 2. Nematodes were fixed overnight in 2% (w/v) paraformaldehyde at 4°C, followed by 4 hours at room temperature and cut into two to five pieces. *In situ* hybridization procedure was performed as described (De Boer *et al.*, 1998). Overnight hybridization was done at 50°C. Nematodes were labelled with alkaline phosphatase-conjugated sheep anti-digoxigenin Fab fragments (Roche Diagnostics). Alkaline phosphatase activity was detected overnight at 4°C with nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate as substrates. The nematodes were examined using differential interference contrast microscopy (Leica, Deerfield).

Results

Sequence characteristics of the cell wall modifying proteins from *M. chitwoodi*

β-1,4-endoglucanases

Screening of the *M. chitwoodi* ppJ2-s EST dataset for sequences with similarity to β-1,4-endoglucanases (cellulases; EC 3.2.1.4) resulted in the identification of 10 ESTs that clustered into four contigs, which were named *Mc-eng-1* to *Mc-eng-4* (Table 3). The full-length sequences obtained for *Mc-eng-1*, *Mc-eng-2*, and *Mc-eng-3*, including

Table 3 Overview of cell wall modifying proteins (CWMP) identified from *M. chitwoodi*. Listed are the number of expressed sequence tags (ESTs) representing each CWMP in the EST dataset generated from the *M. chitwoodi* ppJ2-s cDNA library and the sequence characteristics of each CWMP. The figures of the deduced protein sequences include the predicted signal peptide for secretion. Similarity searches were performed using BLASTP on the non-redundant database at NCBI using the protein sequences minus signal peptide for secretion as search queries

CWMP	ESTs	Length (nt)	Protein (aa)	SP cleavage site	MW (kDa)	pI	Similarity	E-value
Mc-eng-1	5	1435	308	Ala_{19}/Val_{20}	34.8	5.51	β-1,4-endoglucanase Pp-eng-2 dbj BAB68523.1 (*Pratylenchus penetrans*)	$3e^{-80}$
Mc-eng-2	2	1345	405	Ala_{17}/Ala_{18}	44.8	8.63	β-1,4-endoglucanase Mi-eng-1 gb AAD45868.1 (*M. incognita*)	e^{-137}
Mc-eng-3	2	1147	338	Ala_{22}/Val_{23}	38	7.07	β-1,4-endoglucanase Mi-eng-2 gb AAK21881.1 (*M. incognita*)	e^{-127}
Mc-eng-4	1	823	-	-	-	-	β-1,4-endoglucanase Mi-eng-1 gb AAD45868.1 (*M. incognita*)	$2e^{-22}$
Mc-cbp-1	2	793	195	Ala_{20}/Glu_{21}	20.9	4.08	β-1,4-endoglucanase Mi-eng-1 gb AAD45868.1 (*M. incognita*)	$1e^{-21}$
Mc-xyl-1	1	1385	394	Gly_{21}/Glu_{22}	44.8	9.01	Possible xylan degradation enzyme (GHF30-like domain and Ricin B-like domain) AAK76864.1 (*Clostridium acetobutylicum*)	$4e^{-74}$
Mc-pel-1	11	1040	264	Ala_{18}/Asp_{19}	29.5	8.92	Pectate lyase MJ-PEL-1 AF455757 (*M. javanica*)	$5e^{-74}$
Mc-pel-2	1	967	274	Cys_{33}/Phe_{34}	30.4	9.14	Pectate lyase MJ-PEL-1 AF455757 (*M. javanica*)	$2e^{-46}$
Mc-pel-3	2	688	-	Ala_{18}/Asp_{19}	-	-	Pectate lyase MJ-PEL-1 AF455757 (*M. javanica*)	$3e^{-66}$
Mc-pel-4	2	786	-	-	-	-	Pectate lyase MI-PEL-1 AF527788 (*M. incognita*)	$2e^{-36}$
Mc-pg-1	4	2029	624	Ala_{21}/Glu_{22}	67.7	6.9	Polygalacturonase AAM28240.1 (*M. incognita*)	$7e^{-141}$
Mc-pg-2	1	918	-	-	-	-	Polygalacturonase AAM28240.1 (*M. incognita*)	$4e^{-72}$

Abbreviations: (nt) nucleotides, (aa) amino acids, (SP) signal peptide, (MW) molecular weight, (kDa) kilodalton, (pI) isoelectric point

the polyA tail, consisted of 1435, 1345 and 1147 nt respectively. The ORFs deduced from *Mc-eng-1*, *Mc-eng-2*, and *Mc-eng-3* were all predicted to have an N-terminal signal peptide for secretion (Table 3). *In situ* hybridization localized the site of *Mc-eng-3* expression in the subventral oesophageal glands of *M. chitwoodi* ppJ2-s (Fig. 1). BLASTP searches performed on the *nr* database with the mature Mc-ENG-1, Mc-ENG-2 and Mc-ENG-3 protein sequences as queries revealed similarity to glycoside hydrolase family 5 (GHF5) cellulases from eukaryotic origins and a wide range of bacterial genera. Conserved domain searches with Mc-ENG-1, Mc-ENG-2 and Mc-ENG-3 revealed the presence of a catalytic domain consisting of approximately 260

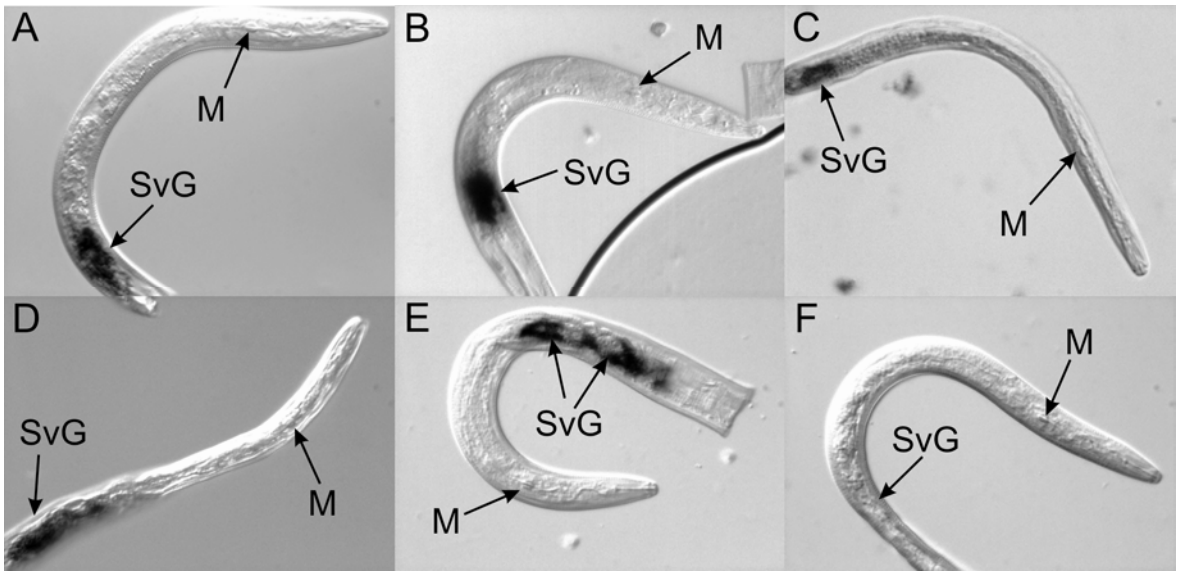


Fig. 1 Whole mount *in situ* localization of cell wall modifying protein encoding transcripts in pre-parasitic second stage juveniles of *Meloidogyne chitwoodi* using digoxigenin-labelled cDNA probes and an alkaline phosphatase-conjugated secondary antibody to digoxigenin. Sections of the nematode incubated with antisense probe designed on *Mc-eng-3* (**A**)**,** *Mc-cbp-1* (**B**)**,** *Mc-xyl-1* (**C**)**,** *Mc-pel-1* (**D**) and *Mc-pg-1* (**E**) showing a specific alkaline phosphatase staining of the subventral esophageal gland cells (SvG). Incubations of nematode sections with sense probe resulted in no specific staining (**F**). The metacorpus is indicated with M and pictures were taken under 400x magnification.

amino acids which is recognized as a GHF5 (pfam00150) cellulase with E-values ranging from $5e^{-27}$ to $8e^{-41}$ (Marchler-Bauer *et al.*, 2005) (http://www.ebi.ac.uk/InterProScan/). The eukaryotic matches were from the plant-parasitic nematode genera *Meloidogyne*, *Globodera*, *Heterodera* and *Pratylenchus*, the longicorn beetle species *Apriona germari* and *Psacothea hilaris*, and the protist *Spirotrichonympha leidyi* which is a symbiotic protist in the hindgut of the lower termite *Coptotermes formosanus* (De Boer *et al.*, 1999; De Meutter *et al.*, 1998; Inoue *et al.*, 2005; Rosso *et al.*, 1999; Smant *et al.*, 1998; Sugimura *et al.*, 2003; Uehara *et al.*, 2001). In its catalytic domain, Mc-ENG-1 shares 55% and 31% similarity at the amino acid level with Mc-ENG-2 and Mc-ENG-3 respectively, whereas Mc-ENG-2 and Mc-ENG-3 show 35% similarity. The similarity and conserved domain searches did not show the presence of a specific ancillary C-terminal domain in either of the three *M. chitwoodi* β-1,4-endoglucanases, although Mc-ENG-2 contains a C-terminal extension of approximately 80 amino acids.

The GHF5 β-1,4-endoglucanases belong to a superfamily of 8-fold β/α-barrels with identical amino acids at their active sites (Jenkins *et al.*, 1995). The nucleophile is a glutamic acid (E) located close to the carboxy-terminus of β-strand seven. In addition, they have the sequence asparagine-glutamate (N-E) approximately 100 residues N-terminal of the nucleophile, close to the carboxy-terminus of β-strand

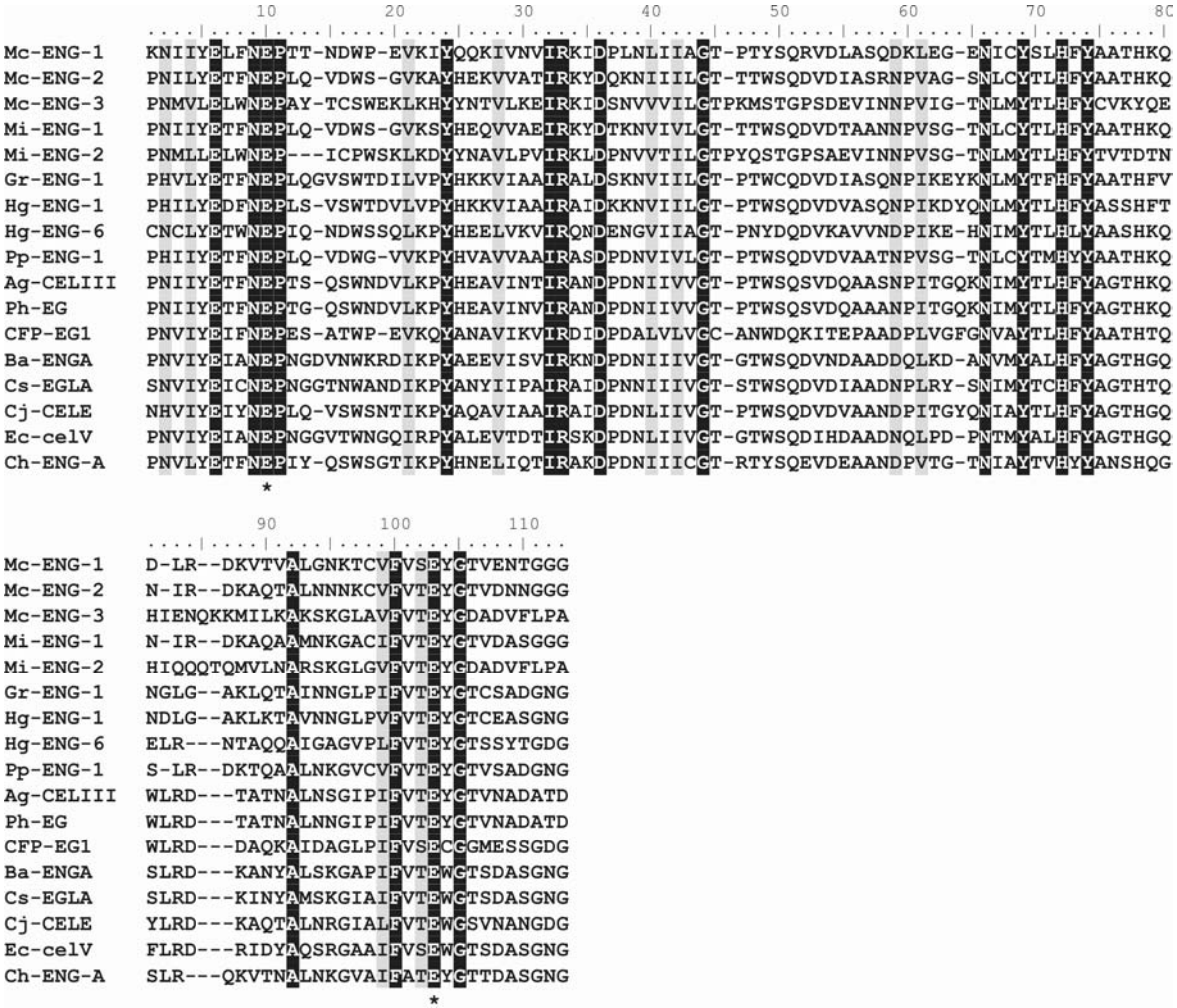


Fig. 2 Alignment of the β-1,4-endoglucanases Mc-ENG-1, Mc-ENG-2 and Mc-ENG-3 from *Meloidogyne chitwoodi* with GHF5 β-1,4-endoglucanases from other organisms. Mi-ENG-1 and Mi-ENG-2 from *M. incognita* (GenBank accession no. AAD45868 and AAK21881), Gr-ENG-1 from *Globodera rostochiensis* (AAC63988), Hg-ENG-1 and Hg-ENG-6 from *Heterodera glycines* (AAC48327 and AAO25506), Pp-ENG-1 from *Pratylenchus penetrans* (BAB68522), Ag-CELIII from *Apriona germari* (AAX18655), Ph-EG from *Psacothea hilaris* (BAB86867), CFP-EG1 from *Spirotrichonympha leidyi* (BAD90558), Ba-ENGA from *Bacillus amyloliquefaciens* (AAL99668), Cs-EGLA from *Clostridium saccharobutylicum* (AAA23230), Cj-CELE from *Cellvibrio japonicus* (CAA60493), Ec-celV from *Pectobacterium carotovorum* (CAA53592) and Ch-ENG-A from *Cytophaga hutchinsonii* (ZP_00308849). Residues identical in all β-1,4-endoglucanases are shaded black and similar residues are shaded grey. The two catalytic residues in GHF5 β-1,4-endoglucanases are indicated with an asterisk.

four, of which the glutamic acid is the acid/base (Jenkins *et al.*, 1995; Py *et al.*, 1991). An alignment of Mc-ENG-1, Mc-ENG-2 and Mc-ENG-3 with other GHF5 cellulases revealed that the conserved regions including the active site glutamic acids are also present in the cellulases from *M. chitwoodi* (Fig. 2).

We found another EST, CD418822, having *M. incognita* β-1,4-endoglucanase Mi-ENG-1 as best match in BLAST searches (Rosso *et al.*, 1999). The cDNA insert of the corresponding library clone, including the polyA tail, was 823 nt in length and

named *Mc-eng-4* (Table 3). *Mc-eng-4* showed similarity to amino acid regions 241-323 (63% identity) and 409-504 (31% identity) of Mi-ENG-1, which are part of the catalytic domain and the complete carbohydrate binding module (CBM) of Mi-ENG-1 respectively. These two domains are separated by a putative linker of approximately 50 amino acids that is almost completely constituted by asparagine, glycine, serine, lysine and threonine. Since *Mc-eng-4* is likely to encode the C-terminal part only of a β-1,4-endoglucanase, we tried to amplify the full length *Mc-eng-4* from the original cDNA library, but these attempts failed.

Phylogeny of *Meloidogyne* β-1,4-endoglucanases

The three novel root-knot nematode β-1,4-endoglucanases Mc-ENG-1, Mc-ENG-2 and Mc-ENG-3 were included in a bayesian analysis to infer the phylogeny of root-knot nematode β-1,4-endoglucanases. For *M. incognita* we retrieved β-1,4-endoglucanase sequences Mi-ENG-1 to Mi-ENG-4 (gi 13398418, 20543838, 40037070 and 40037077 respectively) and the sequences gi 13398416 and gi 5639669 from the *nr* database. The screening of clustered ESTs from *Meloidogyne* spp. resulted in the identification of eight *M. hapla* sequences with NemaGene contig IDs MH02744, MH02523, MH02829, MH01029, MH07431, MH01728, MH01769 and MH02046 (E-values ranging from $1.9e^{-125}$ to $6.6e^{-23}$), four *M. arenaria* sequences MA00503, MA03107, MA01394 and MA00787 (E-values ranging from $2e^{-92}$ to $2e^{-20}$), and three *M. javanica* sequences MJ00626, MJ04631, MJ04449 (E-values ranging from $3e^{-91}$ to $1e^{-41}$). In case of *M. paranaensis*, no sequences were identified with similarity to β-1,4-endoglucanase. In addition, *P. pratylenchus* β-1,4-endoglucanases Pp-ENG-1 and Pp-ENG-2 (gi 15777927 and 15777928) were included in the phylogenetic analysis and the β-1,4-endoglucanase Gr-ENG2 (gi 2654525) from *G. rostochiensis* was used as out-group.

An initial alignment of the partial β-1,4-endoglucanases, led to the exclusion of sequences MH02829, MH01029, MH07431, MH01728, MA03107 and MA00787 due to insufficient sequence overlap. The phylogeny generated from the alignment revealed four distinct clusters of β-1,4-endoglucanases with high support values (posterior probabilities of 1.00, Fig. 3). One cluster is formed by *M. chitwoodi* Mc-ENG-3, *M. incognita* Mi-ENG-2 and *M. hapla* cellulases MH02046 and MH01769. The second cluster is comprised of cellulases from all root-knot nematode species included in the analysis, i.e. *M. hapla*, *M. javanica*, *M. arenaria*, *M. incognita* and *M. chitwoodi*. Among the latter are Mi-ENG-1, Mi-ENG-3 and Mi-ENG-4 from *M. incognita* and Mc-ENG-2 from *M. chitwoodi*. The third cluster of root-knot nematode cellulases is formed by Mc-ENG-1 and *M. javanica* cellulase MJ04631. The *P. pratylenchus* cellulases Pp-ENG-1 and Pp-ENG-2 form a separate cluster. The four clusters collapse to polytomy at the

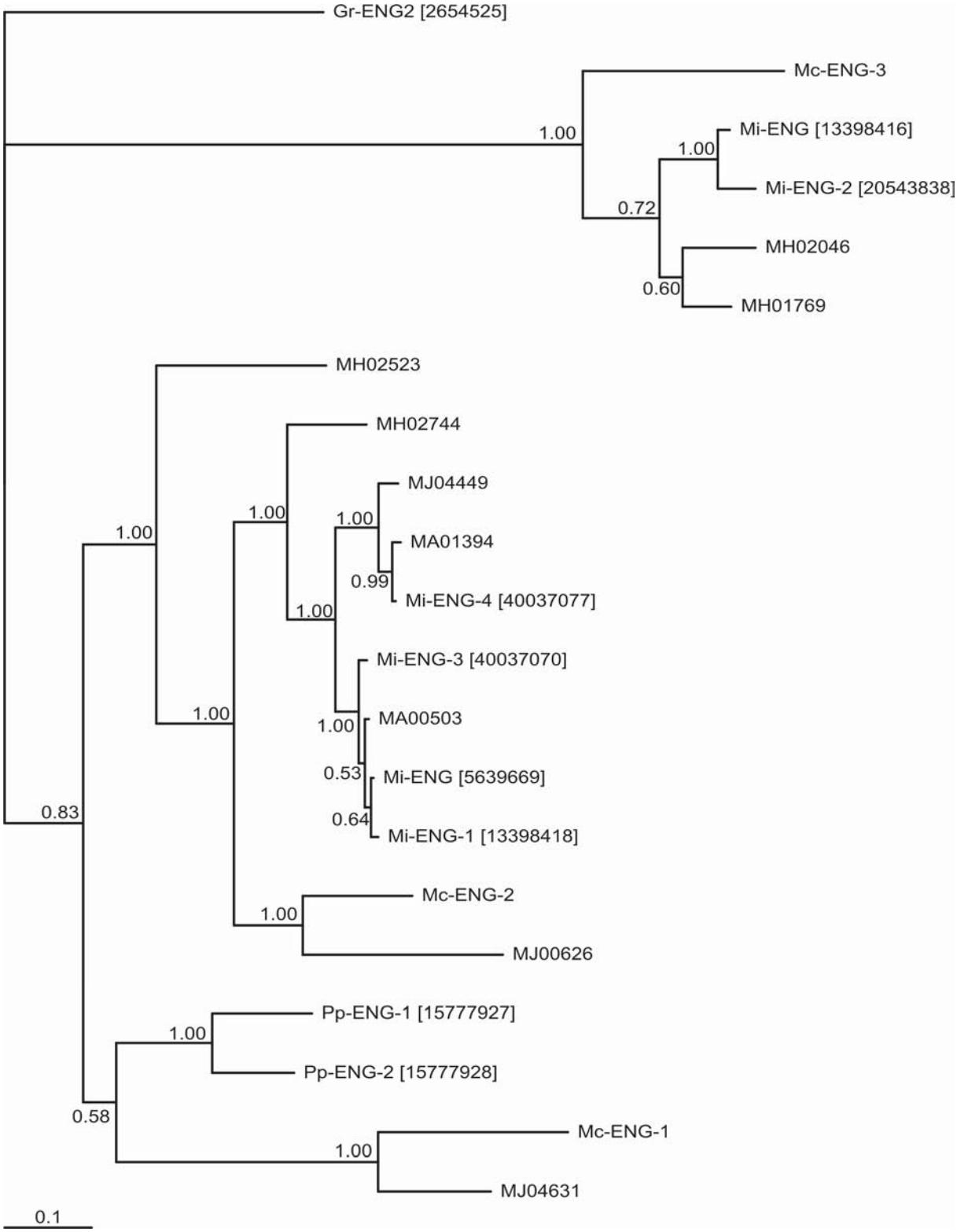


Fig. 3 Phylogenetic tree of root-knot nematode β-1,4-endoglucanases. The tree was constructed using Bayesian analysis based on a ClustalW alignment of the deduced catalytic domain of β-1,4-endoglucanases from *Meloidogyne chitwoodi* (Mc), *M. incognita* (Mi), *M. hapla* (MH), *M. javanica* (MJ) and *M. arenaria* (MA). Two β-1,4-endoglucanases from *Pratylenchus penetrans*, Pp-ENG-1 and Pp-ENG-2, were also included and β-1,4-endoglucanase Gr-ENG2 from *Globodera rostochiensis* was used as out-group. Gi numbers are included between [brackets].

basal nodes due to low support values, and therefore no robust conclusions can be drawn on the evolutionary history of these four clusters.

Cellulose binding protein

Two ESTs (CB930801 and CB931574) were found with similarity to the cellulose binding protein precursor CBP-1 of *Meloidogyne incognita* (Genbank accession AAC05133). The full-length cDNA sequence assembled from the insert sequences of the two corresponding library clones, which was named *Mc-cbp-1*, consisted of 793 nucleotides (nt), including a polyA tail (Table 3). The largest ORF in *Mc-cbp-1* encodes a protein of 195 amino acids with a calculated molecular weight of 20.9 kDa and an isoelectric point (pI) of 4.08. The protein is predicted to have a signal peptide for secretion at its N-terminus with a cleavage site between amino acid position 20 and 21. In addition, the site of *Mc-cbp-1* expression was localized to the subventral oesophageal secretory glands of *M. chitwoodi* ppJ2-s (Fig. 1). No consensus sites for N-linked glycosylation were predicted in the mature Mc-CBP-1.

A BLASTP search of the *nr* database with Mc-CBP-1 revealed that the C-terminal 98 amino acids share 34 to 51% identity with the CBMs in β-1,4-endoglucanases from root-knot and cyst nematodes, in the cellulose binding protein Mi-CBP-1 from *M. incognita*, and in the expansin Gr-EXPB1 from *G. rostochiensis* (Fig. 4). These CBMs are all classified as bacterial family 2 CBMs, which are characterized by two highly conserved cysteine residues positioned at each extremity of the domain. These cysteines have been shown to be involved in an intramolecular disulfide bond (Gilkes *et al.*, 1991a). We used the Interpro database accession IPR001919 to determine if these conserved cysteine residues are present in the CBMs of Mc-CBP-1 and of other nematode proteins (http://www.ebi.ac.uk/InterProScan/). Notably, Mc-CBP-1 contains a cysteine near both the N- and C-terminus of the CBM, whereas the CBM of the β-1,4-endoglucanases from root-knot and cyst nematodes as well as Mi-CBP-1 only have one C-terminal cysteine. The N-terminal cysteine in these latter CBMs is replaced by valine (V). Next to the two conserved cysteines, there are also four conserved tryptophans (W) in bacterial family 2 CBMs of which two are involved in cellulose binding (Gilkes *et al.*, 1991b). These latter two residues are also conserved in CBMs of plant-parasitic nematodes found to date, including the CBM of Mc-CBP-1 (Fig. 4). The third tryptophan is replaced by valine (V) in the CBMs of the plant-parasitic nematodes and the fourth is either conserved or replaced by the aromatic tyrosine (Y). Another characteristic of bacterial family 2 CBMs are the conserved asparagine (N) and glycine (G) residues, which are also present in the CBM of Mc-CBP-1 (Fig. 4)(Gilkes *et al.*, 1991b).

Fig. 4 Alignment of the carbohydrate-binding module (CBM) of Mc-CBP-1 from *M. chitwoodi* with CBMs from other plant-parasitic nematode species and bacterial species. Mi-CBP-1: a cellulose binding protein from *M. incognita* (GenBank accession no. AAC05133), β-1,4-endoglucanase Mi-ENG-1 from *M. incognita* (AAD45868), Gr-ENG-1 from *Globodera rostochiensis* (AAC63988), Gts-ENG-1 from *Globodera tabacum solanacearum* (AAD56392), Hg-ENG-1 from *Heterodera glycines* (AAC48327), Hs-ENG-1 from *Heterodera schachtii* (CAC12958), Gr-EXPB1: an expansin from *G. rostochiensis* (CAC83611), CENA: an endoglucanase A precursor from *Cellulomonas fimi* (P07984), EGA: an endoglucanase A precursor from *Pseudomonas fluorescens* (P10476) and EGD: an endoglucanase D precursor from *Clostridium cellulovorans* (P28623). Identical residues are shaded black and residues that are similar are shaded grey. Residues conserved in all CBMs are indicated with an asterisk, these include the tryptophans (W) that are putatively involved in cellulose binding.

Database searches with the N-terminal 80 amino acids of the mature Mc-CBP-1 protein on the SWISS-PROT and *nr* databases only revealed weak similarity (29% identity and E-value 0.25) with part of the fibronectin type 3 domain (Fn3, cd00063) of the ephrin type-A receptor 7 precursor from several animal species. Fn3-like domains (Fn3) are a common constituent of animal proteins and are also found in one particular class of bacterial proteins, namely the extracellular glycosyl hydrolases (Kataeva *et al.*, 2002; Little *et al.*, 1994). A further BLASTN and TBLASTN search on the est_others database was performed with Mc-CBP-1 to identify similar sequences in other plant-parasitic nematode species. No ESTs were found to match the N-terminal 80 amino acids of Mc-CBP-1.

Comparison of the N-terminal domains of Mc-CBP-1 and MI-CBP-1 (34% identity) shows that both have a region rich in serine (S), which was identified as a Q-type linker in MI-CBP-1 (Ding *et al.*, 1998; Gilkes *et al.*, 1991b). Intriguing is the fact that

approximately one out of three amino acids of the N-terminal region of both CBP proteins are charged and/or polar amino acids, i.e. aspartic acid (D), glutamic acid (E) and asparagine (N) (data not shown).

Endo-1,4-β-xylanase

One cDNA sequence of 548 nt, CD418879, was found in the *M. chitwoodi* ppJ2 EST dataset with similarity to a possible xylan degradation enzyme from *Clostridium acetobutylicum* (GenBank accession number AAK76864). Alignment of the sequence of CD418879 with the central parts of various endo-xylanases in the *nr* protein database indicated that the 5′-end of the sequence was lacking. The full-length cDNA sequence was amplified by PCR and sequenced from the *M. chitwoodi* ppJ2 cDNA library and named *Mc-xyl-1*. It consists of 1385 nt, including the polyA tail and a polyadenylation signal (AATAAA) in its 3′-untranslated region. The open reading frame encodes a protein of 394 amino acids with a molecular weight of 44.8 kDa (Table 3). The first 21 amino acids at the N-terminus code for a signal peptide for secretion and two consensus sites for N-linked glycosylation were predicted (N_{274} to T_{276} and N_{347} to S_{349}). *In situ* hybridization localized the *Mc-xyl-1* transcripts in the subventral esophageal secretory glands of *M. chitwoodi* ppJ2-s (Fig. 1).

A BLASTP search on the *nr* database with the Mc-XYL-1 protein sequence without its signal peptide showed significant similarity to (possible) xylanases of several bacterial species and Mi-XYL1 from *M. incognita*, which was the first functional animal endo-1,4-β-xylanase identified (Mitreva-Dautova *et al.*, 2006). These xylanases are all classified as GHF5 members in the CAZy database of glycoside hydrolases and therefore Mc-XYL-1 is also believed to belong to this family (http://afmb.cnrs-mrs.fr/CAZY/fam/acc_GH.html) (Table 4). An alignment of the catalytic domain of Mc-XYL-1 with the other GHF5 xylanases revealed that the two putative active site glutamic acids (E) flanking β-sheets β4 and β7 are also present in our *M. chitwoodi* xylanase (Fig. 5). The amino acid motif G-F-G-G that makes up the first β-sheet in the GHF5 endoxylanase members and that is absent in the cellulase members of this family is present in Mc-XYL-1 as $G_{15}F_{16}G_{17}A_{18}$ (numbers refer to positions in alignment of Fig. 5).

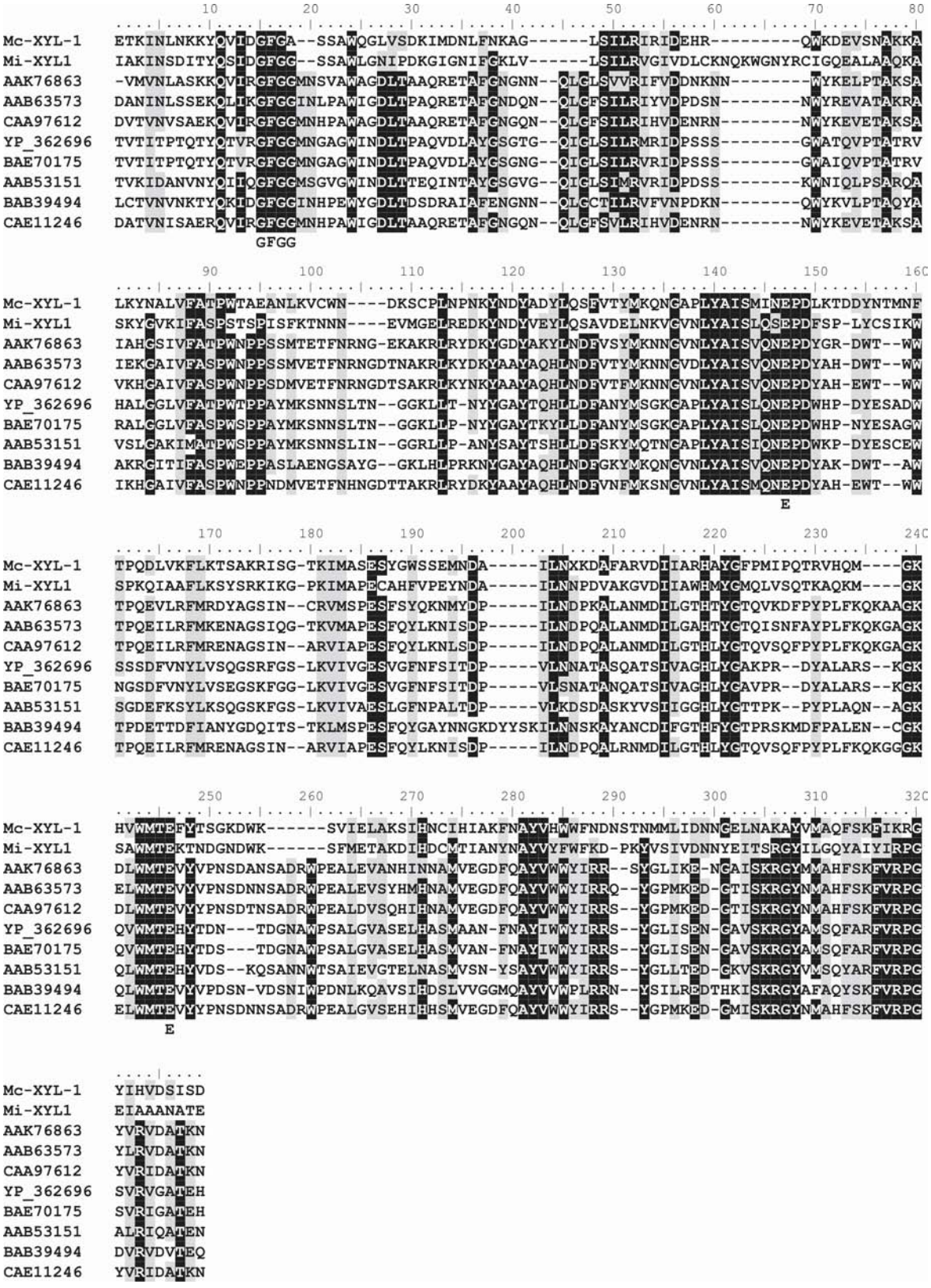


Fig. 5 Alignment of the putative catalytic domain of Mc-XYL-1 from *M. chitwoodi* with the catalytic domains of GHF5 xylanases from bacterial species and *Meloidogyne incognita*. Mi-XYL1: xylanase from *Meloidogyne incognita*, AAK76863 (GenBank accession number): possible xylan degradation enzyme from *Clostridium acetobutylicum*, AAB63573: xylanase D from *Aeromonas punctata*, CAA97612: YnfF from *Bacillus subtilis*, YP_362696: xylanase precursor from *Xanthomonas campestris pv. vesicatoria*,

Fig. 5 (continued) BAE70175: xylanase from *Xanthomonas oryzae*, AAB53151: xylanase from *Erwinia chrysanthemi*, BAB39494: xynC from *Ruminococcus albus*, CAE11246: YnfF from *Bacillus amyloliquefaciens*. Identical residues are shaded black and residues that are similar are shaded grey. Indicated are the two putative active site glutamic acids (E) and the G-F-G-G amino acid motif that makes up the first β-sheet in the catalytic domain of GHF5 endoxylanase members and that is absent in the cellulase members in this family. The latter motif is present in Mc-XYL-1 as G-F-G-A.

Table 4 Results of BLASTP search on the non-redundant database at NCBI using either (I) the putative catalytic domain (CD, amino acid region 22 to 318) or (II) the C-terminal putative xylan binding module (XBM, amino acid region 319 to 394) of Mc-XYL-1 from *M. chitwoodi* as search query

Organism	Accession number	Annotation	CAZy	%ID / %Pos with CD (I)	%ID / %Pos with XBM (II)
Clostridium acetobutylicum	AAK76864	Possible xylan degradation enzyme	GHF5	43/56	33/53
Meloidogyne incognita	AAF37276	xylanase	GHF5	39/60	-
Clostridium thermocellum	EAM45109	Clostridium cellulosome enzyme	n.d.	35/54	35/57
Aeromonas punctata	BAA13641	Endo-xylanase	GHF5	36/54	33/52
Clostridium acetobutylicum	AAK76863	Possible xylan degradation enzyme	GHF5	38/56	42/64
Aeromonas punctata	AAB63573	Xylanase D	GHF5	36/54	34/55
Bacillus subtilis	CAA97612	YnfF	GHF5	36/55	42/57
Clostridium cellulolyticum	EAV71621	Carbohydrate binding family 6	n.d.	33/53	22/52
Xanthomonas campestris	CAJ22596	Xylanase precursor	GHF5	34/49	30/44
Xanthomonas oryzae	BAE70175	xylanase	GHF5	33/49	26/42
Erwinia chrysanthemi	AAB53151	xylanase	GHF5	32/53	27/40
Ruminococcus albus	BAB39494	xynC	GHF5	33/50	36/55

Mc-XYL-1 and Mi-XYL1 share 39% identity in their catalytic domain (Table 4). Further comparison of these two root-knot nematode xylanases showed that Mc-XYL-1 has an ancillary stretch of 76 amino acids at its C-terminus. Among the bacterial xylanases that match this C-terminal domain, with levels of identity in the range of 22 - 42%, is the GHF5 xylanase XynA (accession number AAB53151) from the bacterial plant pathogen *Erwinia chrysanthemi* (Table 4 and Fig. 6) (Larson *et al.*, 2003). The matching C-terminal region of XynA is suggested to be an unclassified

xylan binding module (XBM) (Larson *et al.*, 2003). These results suggest that Mc-XYL-1 may comprise both a catalytic domain and a C-terminal XBM. Notably, the putative XBM of Mc-XYL-1 contains a high number of the aliphatic and hydrophobic amino acid isoleucine (I) and the polar amino acid serine (S). The amino acid stretch $R_{316}F_{317}V_{318}R_{319}P_{320}G_{321}A_{322}$ that forms the linker between the catalytic domain and the XBM of XynA from *E. chrysanthemi* is present as $K_{299}F_{300}I_{301}K_{302}R_{303}G_{304}Y_{305}$ in Mc-XYL-1(Larson *et al.*, 2003).

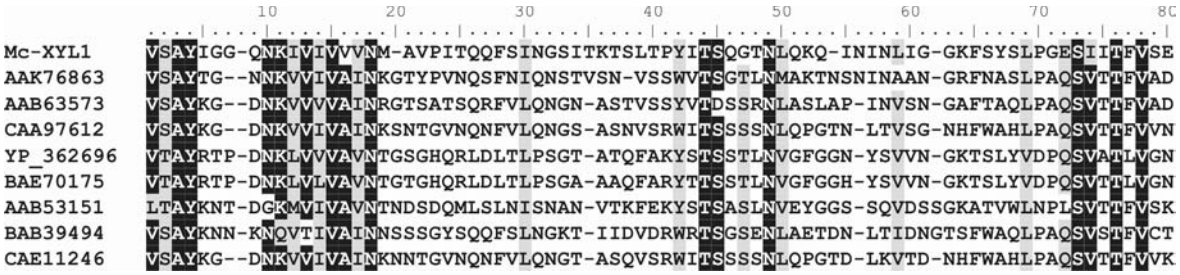


Fig. 6 Alignment of the C-terminal putative xylan binding module (XBM) of Mc-XYL-1 from *M. chitwoodi* with matching amino acid regions of GHF5 xylanases from bacterial species. AAK76863: possible xylan degradation enzyme from *Clostridium acetobutylicum*, AAB63573: xylanase D from *Aeromonas punctata*, CAA97612: YnfF from *Bacillus subtilis*, YP_362696: xylanase precursor from *Xanthomonas campestris pv. vesicatoria*, BAE70175: xylanase from *Xanthomonas oryzae*, AAB53151: xylanase from *Erwinia chrysanthemi*, BAB39494: xynC from *Ruminococcus albus*, CAE11246: YnfF from *Bacillus amyloliquefaciens*. Identical residues are shaded black and residues that are similar are shaded grey.

Pectate lyases

In total, 16 ESTs were found in our dataset with significant similarity (E-values of most similar matches smaller than e^{-35}) to pectate lyases. Assembly of these tags resulted in four contigs named *Mc-pel-1* to *Mc-pel-4* (Table 3). *Mc-pel-1* is represented by eleven ESTs and its full length sequence consisted of 1040 nt, including a polyA-tail. This sequence encodes a protein of 264 amino acids with a calculated molecular weight of 29.5 kDa (Table 3). SignalP analysis predicted a signal peptide for secretion of 18 amino acids at its N-terminus (Bendtsen *et al.*, 2004). Four consensus sites for N-linked glycosylation were predicted in the Mc-PEL-1 sequence, i.e. $N_{27}ITV_{30}$, $N_{37}DSL_{40}$, $N_{180}LTA_{183}$, and $N_{200}DTA_{203}$. A conserved domain search with Mc-PEL-1 revealed that the amino acid region 46 to 206 matched protein family pfam03211 built from pectate lyases (E-value $1e^{-28}$, Marchler-Bauer *et al.*, 2007). A BLASTP search on the *nr* database with Mc-PEL-1 revealed highest similarity with the family 3 pectate lyases (PL3) MJ-PEL-1 from *M. javanica* (GenBank accession number AF455757) and MI-PEL-1 from *M. incognita* (GenBank accession number AF527788), both with 55% identity and 71% similarity (Doyle and Lambert, 2002; Huang *et al.*, 2005; Shevchik *et al.*, 1997). In addition, Mc-PEL-1 shares 30 to 43% identity and 48 to

59% amino acid similarity with PL3 pectate lyases from a range of bacterial and fungal species, pine wood nematode species *Bursaphelenchus xylophilus* and *B. mucronatus*, and the cyst nematode species *G. rostochiensis* and *H. schachtii*. The site of *Mc-pel-1* expression in *M. chitwoodi* ppJ2-s was localized to the subventral esophageal secretory glands (Fig. 1).

The full-length sequence of *Mc-pel-2*, obtained from clone CD418834, consists of 967 nt including a polyA tail. It encodes a protein of 274 amino acids with a calculated molecular weight of 30.4 kDa (Table 3). SignalP analysis predicted a signal peptide for secretion at the N-terminus of Mc-PEL-2 with its cleavage site between C_{33} and F_{34} (Bendtsen *et al.*, 2004). Two consensus sites for N-linked glycosylation were predicted in the Mc-PEL-2 sequence, i.e. $N_{57}FTR_{60}$ and $N_{198}LTA_{201}$. A conserved domain search revealed that amino acid region 54 to 216 matches protein family pfam03211 built from pectate lyases (E-value $1e^{-21}$, Marchler-Bauer *et al.*, 2007). A pair-wise alignment showed that Mc-PEL-2 had 42% identity and 59% similarity to Mc-PEL-1. A BLASTP search on the *nr* database with Mc-PEL-2 as query showed similarity to the same PL3 pectate lyases as identified for Mc-PEL-1. At the amino acid level, Mc-PEL-2 shared 43% identity and 60% similarity to the pectate lyases MJ-PEL-1 from *M. javanica* and MI-PEL-1 from *M. incognita* (Doyle and Lambert, 2002; Huang *et al.*, 2005). The identity and similarity shared with pectate lyases from bacteria, fungi and other plant-parasitic nematode species ranged from 25 to 45% and 41 to 60% respectively.

Site-directed mutagenesis of the PL3 pectate lyase gene Pel-15 (accession number AB011839) from *Bacillus* sp. strain Ksm-P15 resulted in the identification of seven charged amino acid residues involved in the catalytic activity and/or calcium binding of the protein (E_{21}, D_{51}, E_{71}, D_{99}, K_{100}, K_{122} and R_{125}; the numbers refer to the amino acid positions in the alignment in Fig. 7)(Hatada *et al.*, 2000). Mc-PEL-1 and Mc-PEL-2 significantly matched Pel-15, both with 39% identity and 57% similarity. In Mc-PEL-1, six of the seven charged amino acids were conserved, only K_{122} is present as R_{122}. In Mc-PEL-2, four of these seven residues could be identified, i.e. E_{21}, D_{51}, E_{71} and K_{100}. The amino acids D_{99}, K_{122} and R_{125} were substituted with G_{99}, E_{122} and C_{125}, respectively. Most members of the PL3 family have two sets of closely spaced cysteines to support two disulphide bonds and these are also present in both Mc-PEL-1 and Mc-PEL-2 (C_{55}, C_{59}, C_{127} and C_{130} in Fig. 7).

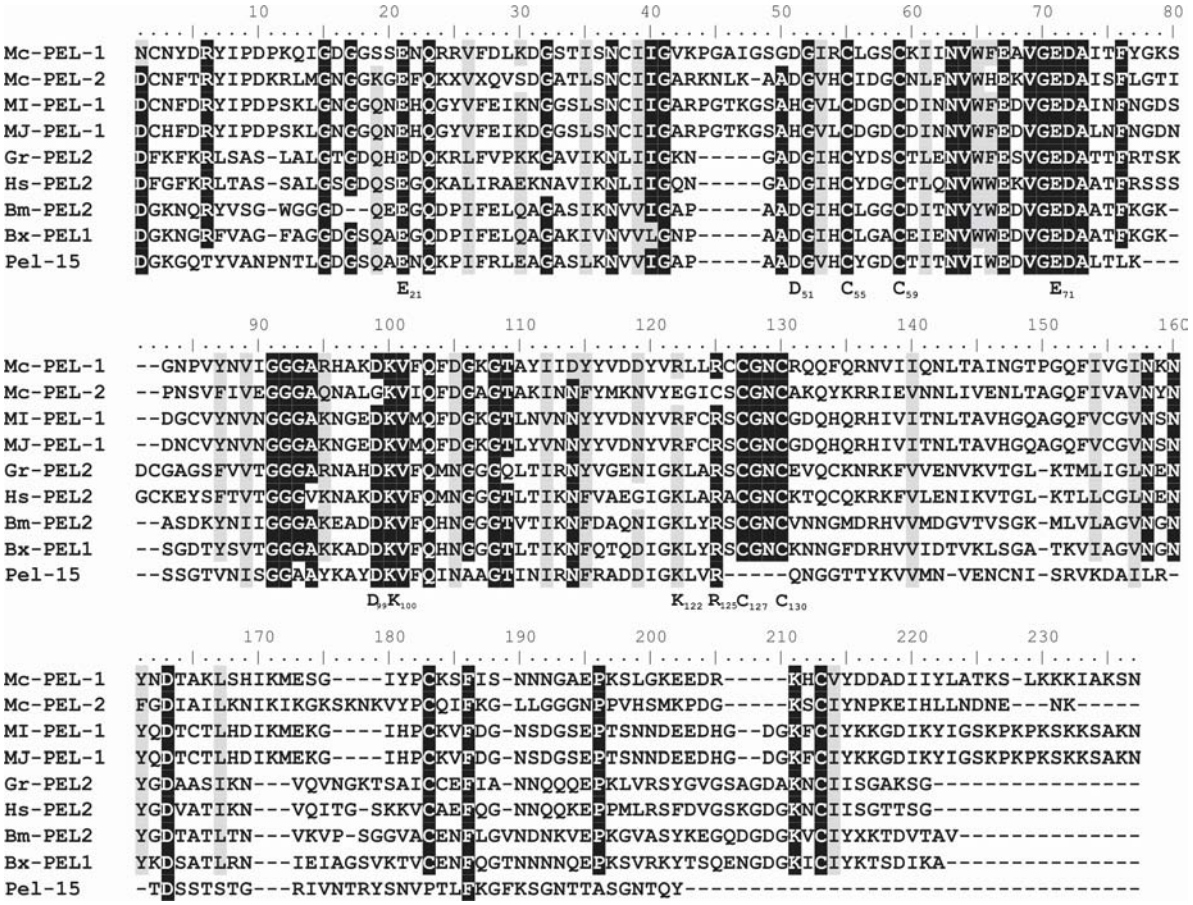


Fig. 7 Alignment of pectate lyases Mc-PEL-1 and Mc-PEL-2 from *M. chitwoodi* with family 3 (PL3) pectate lyases from other plant-parasitic nematode species and Pel-15 from *Bacillus* sp. strain Ksm-P15 (protein database accession number 1EE6_A). MI-PEL-1 from *Meloidogyne incognita* (GenBank accession number AAS88579). MJ-PEL-1 from *Meloidogyne javanica* (AAL66022), Gr-PEL2 from *Globodera rostochiensis* (AAM21970), Hs-PEL2 from *Heterodera schachtii* (ABN14272), Bm-PEL2 from *Bursaphelenchus mucronatus* (BAE48375) and Bx-PEL1 from *Bursaphelenchus xylophilus* (BAE48369). Identical residues are shaded black and residues that are similar are shaded grey. Indicated are the seven charged amino acid residues that are involved in the catalytic activity and/or calcium binding of the protein (E_{21}, D_{51}, E_{71}, D_{99}, K_{100}, K_{122} and R_{125}) (Hatada *et al.*, 2000). The two sets of closely spaced cysteines that are present in most members of the pectate lyase PL3 family, including Mc-PEL-1 and Mc-PEL-2 from *M. chitwoodi*, are marked with C_{55}, C_{59}, C_{127} and C_{130}.

The putative pectate lyases *Mc-pel-3* and *Mc-pel-4* are represented by partial cDNA sequences in the *M. chitwoodi* ppJ2 cDNA library. *Mc-pel-3*, obtained from clones CB855486 and CD420515, consisted of 688 nt (Table 3) and the deduced amino acid sequence shared 57% identity and 74% similarity with the N-terminal amino acid region 8 to 208 of MJ-PEL-1 from *M. javanica* and MI-PEL-1 from *M. incognita*. The partial Mc-PEL-3 sequence contained an N-terminal signal peptide for secretion of 18 amino acids (Bendtsen *et al.*, 2004). Pairwise alignments between Mc-PEL-3 and Mc-PEL-1, and Mc-PEL3 and Mc-PEL-2 showed that the identity at the amino acid level in the overlapping region is 85% and 48% respectively. *Mc-pel-4* consisted of 786 nt

including a polyA tail. The 201 amino acid sequence encoded by *Mc-pel-4* shows similarity to amino acid region 63 to 255 (46% identity) of *M. incognita* MI-PEL-1 and *M. javanica* MJ-PEL-1. Pairwise alignments between Mc-PEL-4 and Mc-PEL-1, and Mc-PEL-4 and Mc-PEL-2 revealed that the amino acid identity in the overlapping region was 42 and 49% respectively. In both Mc-PEL-3 and Mc-PEL-4, five of the seven charged amino acid residues involved in the catalytic activity and/or calcium binding of the pectate lyase are conserved (E_{21}, D_{51}, E_{71}, K_{100} and R_{125}). In Mc-PEL-3, both sets of closely spaced cysteines are present as is the case for most PL3 members. Notably, Mc-PEL-4 only contains the cysteines C_{127} and C_{130} (data not shown).

The phylogeny of fifty members of the PL3 family, including Mc-PEL-1 to Mc-PEL-4, reveals four monophyletic clades (named PL3-NEMs) with all currently known pectate lyases of *M. chitwoodi* clustered in clade PL3-NEM-2 (Fig. 8). PL3-NEM-2 also includes a pectate lyase from two other root-knot nematode species, i.e. *M. incognita* and *M. javanica*. While inspecting the underlying alignment of nematode pectate lyases we discovered that the PL3-NEM clades fall in two distinct categories identified by two highly conserved regions. The first region matches the canonical sequence G_{15}-G-E-x-Q_{23} (numbers refer to positions in the alignment in Fig. 7), while the second includes the residues V_{64}-W-x-E-x-V-G-E-D_{72}. These regions, which have previously been described as part of the conserved BF domain in family 3 polysaccharide lyases (Hatada *et al.*, 2000), are present in clades PL3-NEM-2, -3 and -4, but absent in PL3-NEM-1.

Polygalacturonases

Five cDNA tags were found with similarity to exo-polygalacturonase PehB (AAC46001; EC 3.2.1.82) from the Proteobacterium *Ralstonia solanacearum* with E-values ranging from $2e^{-51}$ to $5e^{-20}$ in BLAST search output (Huang and Allen, 1997). Four of these sequence tags (CB931841, CB931230, CB931913, and CB856055) clustered into one contig, which was named *Mc-pg-1*. Sequence tag CB932008 is a significantly different singleton and was therefore named *Mc-pg-2*.

The full-length sequence of *Mc-pg-1* was extracted from the original library clones and consisted of 2029 nt including a polyA tail. The longest ORF in this sequence encodes a protein of 624 aa (Table 3). The SignalP analysis (in Neural Network mode) predicted the presence of an N-terminal signal peptide for secretion with its most likely cleavage site between amino acid positions 21 (alanine) and 22 (glutamic acid), whereas the hidden Markov models (HMM) in the SignalP analysis predicted that Mc-PG-1 is not a secretory protein. *In situ* hybridization performed on *M. chitwoodi* ppJ2-s, with a 374 nt probe derived from *Mc-pg-1*, revealed a specific expression confined to the subventral esophageal secretory glands (Fig. 1).

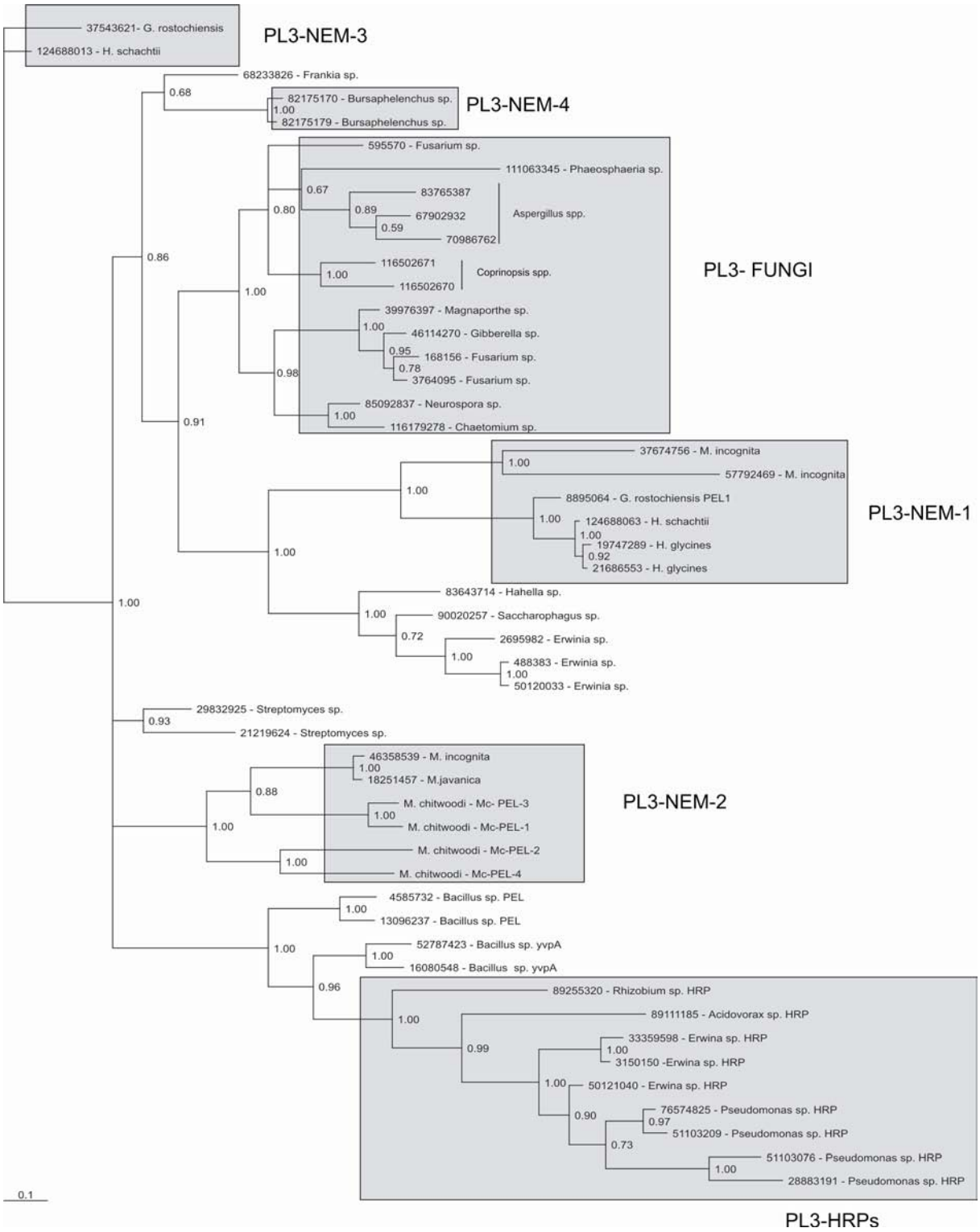


Fig. 8 Phylogenetic tree of fifty members of pectate lyase family 3 (PL3). The tree was constructed using Bayesian analysis based on a ClustalW alignment of deduced pectate lyase amino acid sequences that showed significant similarity to the pectate lyases Mc-PEL-1 to Mc-PEL-4 from root-knot nematode *M. chitwoodi*. The four clusters built from PL3 pectate lyases of nematode origin are indicated as PL3-NEM-1 to PL3-NEM-4. The PL3 pectate lyases from *M. chitwoodi*, Mc-PEL-1 to Mc-PEL-4, are all clustered in PL3-NEM-2. The gi numbers and (part of) species name are indicated for each PL3 pectate lyase.

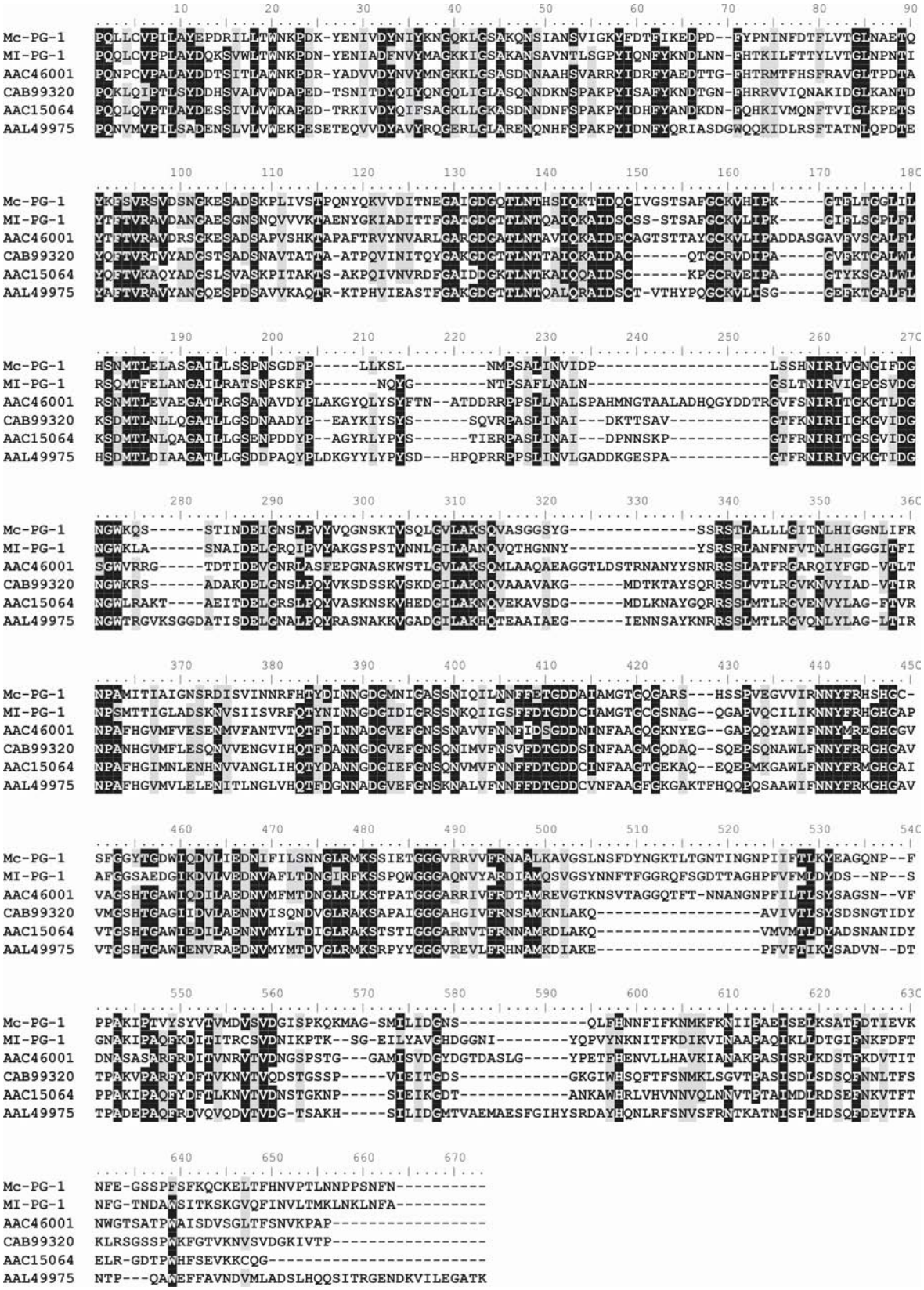


Fig. 9 Alignment of polygalacturonase Mc-PG-1 from *M. chitwoodi* with polygalacturonases from *M. incognita* and several bacterial species. MI-PG-1: polygalacturonase from *M. incognita* (accession number AAM28240), exo-polygalacturonase (EC 3.2.1.82) of *R. solanacearum* (AAC46001), *Erwinia chrysanthemi* (CAB99320), *Yersinia enterocolitica* (AAC15064) and polygalacturonase of *Klebsiella oxytoca* (AAL49975). Identical residues are shaded black and similar residues are shaded grey.

A BLASTP search on the *nr* database with the mature Mc-PG-1 showed significant similarity to the polygalacturonase MI-PG-1 from *M. incognita* (AAM28240; 47% identity / 65% similarity; Jaubert *et al.*, 2002) (see Fig. 9 for alignment). Matches were also found with exo- and endo-polygalacturonases of several bacterial species (30 to 40% identity and 44 to 55% similarity). The annotation of the bacterial endo-polygalacturonases is derived from a consensus built on twelve protein sequences (NCBI conserved domain database accession number COG5434). A conserved domain search with Mc-PG-1 also revealed similarity to the COG5434 endo-polygalacturonase domain PGU1 (E-value $1e^{-55}$). The protein sequences underlying COG5434 are all classified as polygalacturonases of GHF28, but the enzyme activity of only two of them is experimentally shown. The activity of protein sequence AAD35522 is classified as EC 3.2.1.67, whereas CAA89686 is typed as an EC 3.2.1.15 enzyme, which represent exo-polygalacturonases and polygalacturonases respectively. In addition, Mc-PG-1 is 40%, 38%, and 37% identical at the amino acid level to the exo-polygalacturonase (EC 3.2.1.82) of *R. solanacearum* (AAC46001), *Erwinia chrysanthemi* (CAB99320), and *Yersinia enterocolitica* (AAC15064), respectively (Fig. 9) (Huang and Allen, 1997; Liao *et al.*, 1999). Based on these similarity search results we believe that Mc-PG-1, like MI-PG-1, codes for an exo-polygalacturonase.

The sequence of *Mc-pg-2* from library clone CB932008 contains 918 nt including a polyA tail (Table 3). The 226 amino acid sequence encoded by *Mc-pg-2* shows similarity to the C-terminal amino acid region of 223 amino acids (62% identity) of *M. incognita* MI-PG-1 (Jaubert *et al.*, 2002). Since *Mc-pg-2* is likely to encode the C-terminal part only of a polygalacturonase, we tried to amplify the full length *Mc-pg-2* from the original cDNA library, but these attempts failed.

Discussion

β-1,4-endoglucanases

Classification of glycoside hydrolases into families is based on amino acid sequence similarities and hydrophobic cluster analysis. The current number of GHFs in this system is 108 (Bourne and Henrissat, 2001; Henrissat, 1991) (http://afmb.cnrs-mrs.fr/CAZY/fam/acc_GH.html). Fourteen GHFs contain cellulases, of which the families GHF5, GHF9 and GHF45 include animal cellulases. The endogenous β-1,4-endoglucanases identified in plant-parasitic nematodes all belong to family GHF5, except for the GHF45-cellulases from the pine wood nematode *B. xylophilus* (de Boer *et al.*, 1999; De Meutter *et al.*, 1998; Kikuchi *et al.*, 2004; Rosso *et al.*, 1999; Smant *et al.*,

1998; Uehara *et al.*, 2001). Here we report the first β-1,4-endoglucanases identified from *M. chitwoodi*, which all belong to family GHF5.

Cellulolytic enzymes have to bind to the cellulose substrate, which is aided by the carbohydrate binding module. Several nematode β-1,4-endoglucanases contain a CBM, whereas others lack such domain and thus only consist of a catalytic domain (Rosso *et al.*, 1999; Smant *et al.*, 1998; Uehara *et al.*, 2001). Also for *M. chitwoodi* we found evidence for the production of β-1,4-endoglucanases with two different domain architectures. The β-1,4-endoglucanases Mc-ENG-1, Mc-ENG-2 and Mc-ENG-3 only comprise a catalytic GHF5 domain. Mc-ENG-2 has a C-terminal extension of approximately 80 amino acids next to its catalytic domain, but for the lack of similarity of this extension to other sequences in the database the biological function for this region remains to be investigated. The similarity of Mc-ENG-4 with the C-terminal part of Mi-ENG-1, i.e. part of the catalytic domain and the complete CBM, is a strong indication that *M. chitwoodi* also produces a β-1,4-endoglucanase with both a catalytic domain and a CBM.

Until now, phylogenetic analyses on GHF5 cellulase members from plant-parasitic nematodes only included β-1,4-endoglucanases from *M. incognita* (Gao *et al.*, 2004b; Ledger *et al.*, 2006). The identification of β-1,4-endoglucanases Mc-ENG-1, MC-ENG-2 and Mc-ENG-3 from *M. chitwoodi* enabled us to improve the resolution of the phylogeny of root-knot nematode β-1,4-endoglucanases. A further improvement was achieved by including partial β-1,4-endoglucanase sequences from *M. hapla*, *M. arenaria* and *M. javanica* matching the GHF5 catalytic domain. Ledger *et al.* (2006) found one cluster of *M incognita* β-1,4-endoglucanases MI-ENG-1, MI-ENG-3 and MI-ENG-4 whereas MI-ENG-2 did not cluster with the other nematode cellulases, suggesting that this gene might represent a unique lineage of cellulases (Ledger *et al.*, 2006). We found that MI-ENG-2 does no longer stand out but rather clusters with β-1,4-endoglucanases Mc-ENG-3, MH02046 and MH01769, from *M. chitwoodi* and *M. hapla* respectively, in a single monophyletic clade (Fig. 3, cluster I). With the inclusion of more root-knot nematode cellulases, we showed that the major cellulase lineages are likely represented in all root-knot nematodes. Cluster II contains MI-ENG-1, MI-ENG-3 and MI-ENG-4 together with β-1,4-endoglucanases from *M. chitwoodi* (Mc-ENG-2), *M. hapla*, *M. arenaria* and *M. javanica* whereas cluster III only consists of two root-knot nematode β-1,4-endoglucanases, i.e. Mc-ENG-1 and MJ04631. Interestingly, we found no sequence from *M. hapla* matching this cluster despite the fact that this species is represented by the highest number of ESTs (24,331) in dbEST. The results of the phylogenetic analyses of Ledger *et al*. (2006) suggest that PP-ENG-1 and PP-ENG-2 are more related to MI-ENG-1 than to MI-ENG-2, although the exact positioning of MI-ENG-2 is unclear due to the low node resolutions. We have found no support for such a closer evolutionary relationship.

Cellulose binding protein

The putative cellulose binding protein Mc-CBP-1 from *M. chitwoodi* comprises a C-terminal domain of 98 amino acids that shares similarity with CBMs identified in cellulases, expansin and cellulose binding proteins from plant-parasitic nematodes (Ding *et al.*, 1998; Gao *et al.*, 2004a; Qin *et al.*, 2004; Rosso *et al.*, 1999; Smant *et al.*, 1998). This so-called CBM contains the conserved residues of bacterial family 2 CBMs, among which are the tryptophans whose aromatic side chains are likely involved in carbohydrate binding (Gilkes *et al.*, 1991b). CBMs have been classified into 48 different families based on the primary amino acid sequence, binding specificity, and structure of the proteins (http://afmb.cnrs-mrs.fr/CAZY/fam/acc_CBM.html) (Shoseyov *et al.*, 2006). Family 2 CBMs are modules of approximately 100 residues that are found in a large number of bacterial carbohydrate active enzymes (Tomme *et al.*, 1998). The possible functions of CBMs are i) to mediate the binding of cellulolytic enzymes to the cellulose substrate, hereby increasing the effective enzyme concentration at the surface of the substrate and ii) to facilitate the non-hydrolytic disruption of cellulose fibers (Boraston *et al.*, 2004; Din *et al.*, 1991; Shoseyov *et al.*, 2006; Tomme *et al.*, 1998). The CBM of endoglucanase A from the bacterium *Cellulomonas fimi* was shown to enhance cellulose degradation by modifying its surface (Din *et al.*, 1991). CBMs are also found in a few nonhydrolytic proteins, e.g. as part of a scaffolding subunit that organizes the catalytic subunits into a cohesive multi-enzyme complex known as cellulosome (Shoham *et al.*, 1999; Shoseyov *et al.*, 2006).

Approximately one out of three amino acids of the N-terminal region of Mc-CBP-1 (80 amino acids) is a charged and/or polar amino acid (D, E, and N). Such highly charged stretches of amino acids are found on the surface of proteins to interact with other proteins, which suggests that it could serve as a flexible adaptor linking the carbohydrate binding moiety to (an)other protein(s) with catalytic properties. This N-terminal region of Mc-CBP-1 further shares some similarity with Fn3-like domains. Fn3 domains are involved in protein-protein interactions, and either act as a linker conferring proper interaction between catalytic and substrate binding modules, bind to the cellulosome, or bind directly to polysaccharides (Campbell and Spitzfaden, 1994; Kataeva *et al.*, 2002; Little *et al.*, 1994). We therefore hypothesize that Mc-CBP-1 is involved in the hydrolysis of cell wall polymers as a docking station for other catalytic CWMPs and/or by rendering the cell wall polymers more susceptible to enzymatic attack by modifying the polymer surface.

Endo-1,4-β-xylanase

The bacterial endo-1,4-β-xylanase members of GHF5 are all multidomain proteins, including a catalytic domain of approximately 280 amino acids and a putative C-terminal CBM of approximately 90 amino acids. The first identified functional animal endo-1,4-β-xylanase, Mi-XYL1, only consists of a catalytic domain of approximately 280 amino acids. Here we present an animal endo-1,4-β-xylanase from *M. chitwoodi*, Mc-XYL-1, that comprises a catalytic domain and a putative xylan-binding module. These results suggest that nematode endoxylanases occur in two architectures, with or without a CBM, such as is the case for nematode cellulases (Smant *et al.*, 1998).

CBMs are grouped into sequence-based families, although they have also been classified into 'types' based on the topology of the binding site, which reflects the nature of the target ligand (Boraston *et al.*, 2004; McCartney *et al.*, 2006). McCartney *et al.* studied the ligand specificity of xylan binding CBMs for polysaccharides embedded within plant cell walls. Their data show that there is significant variation in the specificity of the different xylan-binding modules for plant cell walls, indicating that the structure and/or context of the xylans embedded in these composite structures is a critical specificity determinant for these protein modules. This means that different CBMs have the capacity to target associated catalytic modules to specific cell walls in diverse species (McCartney *et al.*, 2006). So far, the putative function of the C-terminal domain of Mc-XYL-1 is only deduced by sequence similarity searches. Functional (immunohistochemical) studies with this putative xylan binding module could shed light on its function and specificity for target polysaccharides in plant cell walls of different plant species.

Pectin degradation

During intercellular migration, root-knot nematodes move through the middle lamella between adjacent plant root cells, which has a high content of pectin. Pectin further forms the matrix in which the cellulose-hemicellulose network is embedded (Carpita and McCann, 2000; Carpita and Gibeaut, 1993). Pectins constitute a heterogeneous group of polysaccharides with a backbone consisting of partially methylesterified galacturonic acid (Cosgrove, 1997a; Tamaru and Doi, 2001). Different regions are recognized in pectin, the so-called smooth and hairy regions. The smooth region is composed of homogalacturonan, a linear polymer of α-1,4-linked D-galacturonic acid. This region is interrupted by rhamnogalacturonan I and, with lesser abundance, rhamnogalacturonan II (hairy region). The rhamnogalacturonans are highly branched and thus complex carbohydrates containing diverse sugars in varying linkages.

Pectic substances in the plant cell wall are important targets for degradation by plant pathogenic bacteria and fungi. The pathogenicity of these organisms results in

part from the secretion of a variety of pectinolytic enzymes and hence, these enzymes are important virulence factors (Herron *et al.*, 2000; Huang and Allen, 1997; Hugouvieux-Cotte-Pattat *et al.*, 1996; Ten Have *et al.*, 2002). In the case of root-knot nematodes, pectate lyases and polygalacturonases, both pectinolytic enzymes, are likely to assist the nematode in its intercellular migratory phase (Doyle and Lambert, 2002; Huang *et al.*, 2005; Jaubert *et al.*, 2002). In this study, we identified four novel pectate lyases (*Mc-pel-1* to *Mc-pel-4*) and two polygalacturonases (*Mc-pg-1* and *Mc-pg-2*) from *M. chitwoodi*.

Pectate lyases

The *M. chitwoodi* pectate lyases *Mc-pel-1* to *Mc-pel-4* distinctly differ at the amino acid level, but they all contain (most of) the charged residues and cysteines characteristic for PL3 pectate lyases. Therefore, we regard them as members of this pectate lyase family, which also includes all the pectate lyases identified from other plant-parasitic nematode species so far. Our finding of a family of at least four pectate lyases in *M. chitwoodi* indicates that the occurrence of small multigene families of pectate lyases in root-knot nematodes is not restricted to tropical *Meloidogyne* species, but is a more widespread phenomenon among root-knot nematode species (Doyle and Lambert, 2002; Huang *et al.*, 2005).

Pectate lyases (EC 4.2.2.2) are known to play a critical role in pectin degradation by catalyzing the random cleavage of internal α-1,4-linkages of pectate, i.e. non-methylated or low methylesterified homogalacturonan, by β-elimination (Barras *et al.*, 1994; Herron *et al.*, 2000). Herron and associates (2000) stated that a pathogen that secretes a battery of pectinolytic enzymes would be expected to have a broad host range because each of these enzymes recognizes a different sequence of methylated and non-methylated oligogalacturonate units (Herron *et al.*, 2000). The finding that *M. chitwoodi*, which has a wide host range, produces at least four different pectate lyases, in addition to two polygalacturonases, supports this statement. Structural studies on the pectate lyases from *M. chitwoodi* may reveal further differences in the substrate-binding groove and hence indicate if the optimal *in vivo* substrate of each enzyme differs.

Polygalacturonases

The SignalP analysis of Mc-PG-1 in Neural Network mode predicted the presence of an N-terminal signal peptide for secretion, whereas in its HMM mode the software predicted that Mc-PG-1 was not a secretory protein. We observed a similar outcome when we tested polygalacturonase MI-PG-1 from *M. incognita*. *Mi-pg-1* was shown to be expressed in the subventral esophageal glands of pre-parasitic juveniles of *M. incognita*. We too found that transcripts of the *Mc-pg-1* gene are specifically localized

in the subventral esophageal secretory glands of infective juveniles. In theory, non-secretory polygalacturonases could have an internal role in the gland cells, however, this seems unlikely for their substrate specificity is limited to pectic substances which do not occur in the cytoplasm of nematode cells. It is therefore more likely that Mc-PG-1 is secreted through the stylet by *M. chitwoodi* and that it plays a role in facilitating nematode invasion and intercellular migration through the host plant root.

Pectin degrading enzymes can be divided in exo-acting and endo-acting enzymes. In general, exo-acting enzymes release monomeric or dimeric glycosyl moieties, whereas endo-acting enzymes cleave a polysaccharide randomly along the chain. The latter results in a very rapid decrease in average chain length, and therefore, an endo-polygalacturonase will have more impact on the integrity of a pectin polymer than an exo-polygalacturonase (Ten Have *et al.*, 2002). The activity of endo-polygalacturonases on pectic substances releases oligogalacturonides, which may serve as elicitors of defence responses by plants (Cervone *et al.*, 1989; Favaron *et al.*, 1988). Biotrophic pathogens, like *Meloidogyne* species, must avoid host defence responses and therefore it seems likely that they will profit from a cell wall degrading enzyme repertoire that does not release oligogalacturonides. Based on the similarity search results, Mc-PG-1 is marked as a putative exo-polygalacturonase, like MI-PG-1 (Jaubert *et al.*, 2002), although biochemical evidence for this classification is lacking. With respect to the different impact of exo-polygalacturonases and endo-polygalacturonases on cell wall integrity and the role of endo-polygalacturonases in plant defence responses, it is interesting to investigate the enzymatic properties of Mc-PG-1 in order to establish to which class of polygalacturonases it belongs.

Conclusion

M. chitwoodi migrates in between cells inside plant roots and as a result of that behaviour, the nematode is constantly faced with structurally and biochemically complex cell walls during its invasion. Furthermore, *M. chitwoodi* has a very broad host plant range, comprising both monocots (including grasses) and dicots, which implicates that the nematode has to be equipped to handle a great diversity in complexity of cell wall components. For several types of CWMPs it has now been shown that they are required for infectivity of nematodes. The subventral esophageal gland specific expression and presence of predicted secretion signal peptides suggest that CWMPs are secreted from the nematode during invasion of the plant root. Post-transcriptional gene silencing by soaking nematode juveniles in double-stranded RNA (RNA interference) was successfully applied to assess the importance of β-1,4-endoglucanases and pectate lyases for infection of plant roots by cyst nematodes

(Chen *et al.*, 2005; Vanholme *et al.*, 2007). At the start of this research we hypothesized that *M. chitwoodi* deploys a wide repertoire of CWMPs including many different activities to match the diversity in the cell walls during invasion. The results of our current study support this hypothesis. The CWMPs identified from *M. chitwoodi* target essentially all major plant cell wall polymers, i.e. cellulose, hemicellulose and pectin. Together with the expansin-like proteins described in Chapter 4, these proteins constitute the most elaborate repertoire of CWMPs found in a single plant-parasitic nematode species so far. In addition, we identified small families within individual classes of CWMPs, e.g. in case of the β-1,4-endoglucanases and pectate lyases, which adds further to the cell wall degrading capabilities of the nematodes. Each of these family members might recognize different sequences of polysaccharide units and therefore complement each other, enabling the nematode to breakdown even more different types of plant cell walls. RNA interference experiments targeting *M. chitwoodi* genes encoding classes of CWMPs and specific members within these classes, followed by infection tests on monocots and dicots, may reveal the impact of the individual CWMPs on the host range of *M. chitwoodi*.

4

Identification and characterization of expansin-like proteins from the root-knot nematode *Meloidogyne chitwoodi*

Erwin Roze, Hein Overmars, Makedonka Mitreva, Jaap Bakker and Geert Smant

Submitted for publication

Summary

Expansins are defined by their typical non-enzymatic pH-dependent activity causing polymer creep in plant cell walls. Recently, the first active expansin protein was found outside the group of land plants in cyst nematodes. In this paper we report the identification of expansin-like proteins from a root-knot nematode species, *Meloidogyne chitwoodi*, showing that the production of these proteins is not a unique evolutionary innovation in cyst nematodes. We named these proteins 'expansin-like' since we were not able to produce active recombinant protein *in planta*. Among the assembled ESTs from *M. chitwoodi* we found four expansin-like sequences, named *Mc-EXP1* to *Mc-EXP4*, including different domain architectures. Both Mc-EXP1 and Mc-EXP2 are multi-domain proteins; they consist of an N-terminal signal peptide for secretion, followed by either a CBM (Mc-EXP1) or a Lysin-motif domain (LysM; Mc-EXP2) both linked to a C-terminal expansin-like domain. On the other hand, Mc-EXP3 and Mc-EXP4 only contain an expansin-like domain with an N-terminal signal peptide for secretion. Probes designed on *Mc-EXP1* localised the gene transcription in the subventral oesophageal secretory glands of ppJ2-s of *M. chitwoodi*. The same localisation was found for the Mc-EXP1 and Mc-EXP2 proteins in immunofluorescence microscopy experiments. These data strongly suggest that Mc-EXP1 and Mc-EXP2 are secreted by the nematode. In addition to the expansin-like sequences from *M. chitwoodi*, we found evidence for the presence of expansin-like sequences in five other root-knot nematode species, a root-lesion nematode species, and plant pathogenic oomycete and fungal species. It remains to be shown whether the expansin-like proteins identified in this study are functional expansins, but our findings imply that expansins are likely to be involved in many different plant-pathogen interactions. Next to the sequence characteristics we discuss the proposed nomenclature of the expansin superfamily and the value of our findings regarding the definition and key amino acid signatures of expansins.

Introduction

The plant cell wall is a complex composite mainly made of polysaccharides. The scaffold is formed by a network of cellulose microfibrils which are tethered by hemicelluloses. This cellulose-hemicellulose network is embedded in a matrix of heterogeneous pectic polysaccharides and a small proportion of structural proteins (Carpita and Gibeaut, 1993; Cosgrove, 1997a). Like a wide range of plant pathogenic bacteria and fungi, plant-parasitic nematodes secrete a mixture of cell wall degrading enzymes (CWDEs) to break down the structural barrier formed by the plant cell wall (e.g. reviewed in Preston *et al.*, 2005; Ten Have *et al.*, 2002; Vanholme *et al.*, 2004). The repertoire of CWDEs in nematodes essentially targets all major cell wall components but their efficacy may be limited due to the compactness of the plant cell wall.

Expansins form a diverse protein superfamily in plants and play a role in various biological processes in which re-arrangement of plant cell wall polysaccharides is involved (Cosgrove, 2000a). They are proposed to weaken non-covalent interactions between cellulose and hemicellulose polymers thereby inducing plant cell wall extension (McQueen-Mason and Cosgrove, 1994). The first proteins isolated from cucumber cell walls that were shown to have cell wall extension activity were named α-expansins. A second group of proteins with similar type of activity albeit on cell walls of monocots, the group-I grass pollen allergens and related proteins in vegetative tissues, were classified as β-expansins (Cosgrove, 1997a; Lee *et al.*, 2001). Both α- and β-expansins consist of two domains, including a domain that is distantly related to the catalytic domain of glycoside hydrolase family 45 (GHF45) and a domain that is distantly related to group-II grass pollen allergens (a putative polysaccharide-binding domain) (Cosgrove, 2000a). Plant expansins lack hydrolytic activity, but they are thought to act synergistically with cellulases by making the plant cell wall polysaccharides more accessible to enzymatic attack (Cosgrove, 2000a).

The presence of expansins was believed to be restricted to land plants (Cosgrove, 2000a), but more recently, a small number of expansin-like sequences were identified from other organisms (Darley *et al.*, 2003; Laine *et al.*, 2000; Saloheimo *et al.*, 2002). In addition, a functional β-expansin with cell wall extension activity was found to be secreted by the potato cyst nematode *Globodera rostochiensis* (Kudla *et al.*, 2005; Qin *et al.*, 2004). *G. rostochiensis* also secretes a range of cellulolytic and pectinolytic enzymes at the onset of parasitism. It was therefore proposed that the synergistic action of expansins and CWDEs secreted by cyst nematodes may facilitate their invasion of the host.

We questioned if the concerted action of expansins and CWDEs is a unique phenomenon in parasitism of cyst nematodes or that similar combinations of cell wall modifying proteins are also secreted by other plant-parasitic nematodes and even by other unrelated plant-pathogens. To address this question, we used the β-expansin of *G. rostochiensis* (Gr-EXPB1) to mine the expressed sequence tags (ESTs) database from *M. chitwoodi* for expansin-like sequences. In total, four expansin-like proteins were found with different domain structures in the currently available sequences. *In situ* hybridization of transcripts of the genes and antisera raised to two of these proteins showed their expression in the subventral oesophageal glands of infective juveniles. Further analysis of the expansin-like proteins suggested that they are likely to be in the secretions of these oesophageal gland cells, which also produce the CDWEs. When we widened the scope of our data mining strategy, we found evidence for the presence of expansin-like sequences in five other root-knot nematode species, a root-lesion nematode, a plant pathogenic oomycete and fungi. Our findings suggest that expansins secreted by parasites and pathogens are likely to play an important role in diverse plant-pathogen interactions.

Materials and Methods

Identification of expansin-like sequences in the genus *Meloidogyne*

A NemaGene Cluster translating basic local alignment search tool (TBLASTN; Altschul *et al.*, 1990) was performed on EST contigs of the root-knot nematode species *M. chitwoodi*, *M. arenaria*, *M. hapla*, *M. incognita*, *M. javanica* and *M. paranaensis* on www.nematode.net with the mature protein sequence of expansin Gr-EXPB1 (EMBL accession number CAC83611). The contig consensus sequences producing High-scoring Segment Pairs (HSPs) with E-values smaller than e^{-15} were selected for further analysis. BLAST analyses of the contig consensus sequences were performed on the non-redundant (*nr*) and est_others (non-mouse and non-human) databases at NCBI to identify matches in these public databases. The contig consensus sequences were further checked to assess if they included the full-length transcript. The proteins encoded by the full-length sequences were fed into SignalP 3.0 to predict the presence and the cleavage site of signal peptides for secretion using both the neural networks and hidden Markov models (www.cbs.dtu.dk/services/SignalP/; Bendtsen *et al.*, 2004). A ScanProsite search was used to screen mature protein sequences for N-glycosylation consensus sites (www.expasy.org/tools/scanprosite/?PS00001). Conserved domain searches were

carried out with the mature protein sequences at NCBI and InterProScan (Marchler-Bauer *et al.*, 2005; www.ebi.ac.uk/InterProScan/).

Nematodes

Meloidogyne chitwoodi (line E1083) was propagated on tomato (*Lycopersicum esculentum*) cultivar Moneymaker in a greenhouse. Infected tomato roots were harvested and thoroughly, but carefully, rinsed with tap water to obtain clean roots with intact egg masses. These roots were cut into pieces and put in a mist chamber to hatch the nematodes at 19°C (Oostenbrink, 1960). Hatched nematodes were transferred to a double cotton wool filter (Hygia Favorit II) to select living individuals, which were collected after approximately five hours.

In situ hybridization microscopy

Sense and anti-sense probes were generated from library clones coding for expansin-like proteins. The probes were obtained by linear amplification and labelled using digoxigenin-dUTP in a transcription reaction (Roche Diagnostics, Mannheim, Germany). Template and primers used for amplification are listed in Table 1. Whole mount *in situ* hybridization procedure was performed as described by de Boer *et al*. (1998)(De Boer *et al.*, 1998). Nematodes were fixed overnight in 2% (w/v) paraformaldehyde at 4°C, followed by 4 hours at room temperature and cut into two to five pieces. Overnight hybridization was done at 50°C. Nematodes were labelled with alkaline phosphatase-conjugated sheep anti-digoxigenin Fab fragments (Roche Diagnostics). Alkaline phosphatase activity was detected overnight at 4°C with nitrobluetetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP) as substrates. The nematodes were examined using differential interference contrast microscopy (Leica, Deerfield).

Table 1 Oligonucleotide primers used to prepare probes for *in situ* hybridization from an EST member of a contig encoding an expansin-like protein

Name	EST	Forward primer (5' - 3')	Reverse primer (5' - 3')
Mc-EXP1 (probe a)	ri53f07	AGAATCAAGCAAAGGTACTAG	ATTCTGCAGCTACTTCACC
Mc-EXP1 (probe b)	ri82e04	AAGGATAATGGTGAAGTAGC	TTGAGGTCAAATAGTTGGTC
Mc-EXP2	ri20c02	GCGACAAAACTTACACGGTAG	CTCAACTTGTGCACTTCCTC
Mc-EXP3	ri22f04	TAAGCTATAATCAAGGCTGG	CAGCATGAGTTTTATCGCAAC
Mc-EXP4	ri12h03	CTAAAGATGTATGGGTACGTTC	GTTACGTCCATTATTGGAATTG

Antisera production and immuno assays

Based on the antigenic potential along two expansin-like protein sequences (Lasergene, DNAstar), named Mc-EXP1 and Mc-EXP2, two antigenic peptide sequences were identified and synthesized for each protein (Eurogentec S.A., Seraing, Belgium). For Mc-EXP1 the synthetic peptides included the regions K_{164}TTPSTNEEKATKST$_{178}$ (EXP1-1) and S_{70}NRNMRIVGNNHYRTP$_{85}$ (EXP1-2), while in the case of Mc-EXP2 the regions L_{178}ADPSRGKYLWRDFS$_{192}$ (EXP2-1) and F_{94}QGHFTYYNDAGYRAC$_{109}$ (EXP2-2) were selected. Two rabbits were immunized with either a mixture of the synthetic peptides EXP1-1 and EXP1-2 or EXP2-1 and EXP2-2. Pre-immune sera and antisera were obtained from each rabbit during 12 weeks post immunization (Eurogentec S.A., Seraing, Belgium).

To check for a potential cross reactivity of the antisera with other *M. chitwoodi* sequences, a TBLASTN search was performed with the four synthetic peptide sequences on the ESTs from *M. chitwoodi* in the est_others database at NCBI. The search did not show any matches other than the ESTs present in the contigs MC01203 and MC01774, which represent *Mc-EXP1* and *Mc-EXP2*. Neither predicted N-glycosylation sites nor other posttranslational modification sites were present in the amino acid region covered by the synthetic peptides suggesting that the corresponding epitopes in the Mc-EXP1 and Mc-EXP2 proteins are not to be shielded from the antibodies by posttranslational modifications.

A dot immunobinding assay was performed to test if the rabbit antisera recognized the synthetic peptides used in the immunization procedure. In this procedure, the pre-immune sera were used as negative control for the antisera from each rabbit. The secondary antibody, an alkaline phosphatase (AP)-conjugated goat anti-rabbit IgG (H+L) (Jackson ImmunoResearch Europe Ltd., Suffolk, UK), was also tested alone to check if this antibody non-specifically bound to the synthetic peptides. Approximately 4 ng, 0.4 ng and 40 pg of synthetic peptide was spotted on a nitrocellulose filter (Schleicher & Schuell, Dassel, Germany) and fixed to the filter by baking at 100°C for 20 minutes. Each incubation step mentioned hereafter was done at room temperature on a roller bench. Between each incubation, the blots were washed at room temperature for four times during one minute in phosphate buffered saline (PBS) plus 0.1% Tween 20. The filters were blocked with PBS plus 0.1% Tween 20 and 5% non-fat dry milk for 30 minutes. The pre-immune and antisera were diluted 1:100 in 10 ml PBS plus 0.1% Tween 20, 0.1% bovine serum albumin (BSA) and 0.05% NaN_3 and added to the filters with peptide spots for first antibody binding for 1 hour. This step was followed by secondary antibody binding with the AP-conjugated goat anti-rabbit IgG (H+L), which was diluted 1:5,000 in 10 ml PBS plus 0.1% Tween 20, 0.1% BSA and 0.05% NaN_3 and incubated with the filter for 1 hour. Alkaline phosphatase activity was detected at room temperature with NBT and

BCIP as substrates. The pre-immune sera and the secondary conjugated antibody did not bind to the synthetic peptides on dot blot and could therefore be used to assess the specific binding of the antisera raised to these peptides. The Mc-EXP1 antiserum that recognized both synthetic peptide EXP1-1 (up to 40 pg) and EXP1-2 (up to 0.4 ng) (data not shown) was used in the immunofluorescence microscopy experiments on ppJ2-s of *M. chitwoodi* (De Boer *et al.*, 1996b). In case of Mc-EXP2, the antiserum used in immunofluorescence microscopy only bound to synthetic peptide EXP2-1 (4 ng).

Immunofluorescence microscopy

The rabbit antisera to Mc-EXP1 and Mc-EXP2 were used in indirect immunofluorescence microscopy to screen for a reaction with *M. chitwoodi* ppJ2-s. Fixation, cutting and permeabilization of nematodes were performed according to the *in situ* hybridization protocol (De Boer *et al.*, 1998). The nematodes were further processed through a series of methanol and acetone treatments as described previously (De Boer *et al.*, 1996a). Labelling of the *M. chitwoodi* ppJ2-s was performed in small reaction tubes containing 10 µl of nematode suspension. To this suspension, 90 µl of either 1:50, 1:100 or 1:500 diluted antiserum was added. As negative controls, pre-immune sera were added to nematode suspensions in the same amount and dilutions as the antisera. We used fluorescein (FITC)-conjugated goat anti-rabbit IgG (H+L) (Jackson ImmunoResearch Europe Ltd., Suffolk, UK) as secondary antibody, which was also tested separately to check for aspecific binding to the nematodes. As positive control, the monoclonal antibody MGR007 was included in 1:50 and 1:100 dilutions. MGR007 reacts specifically with the body wall muscles of *G. rostochiensis* and several other cyst and root-knot nematode species (De Boer *et al.*, 1996a). The secondary antibody used against MGR007 was fluorescein (FITC)-conjugated rat anti-mouse IgG (H+L) (Jackson ImmunoResearch Europe Ltd., Suffolk, UK). The nematodes were examined with an Orthoplan Leitz EpiFluorescence microscope using filter 4.

Results

Identification and characterization of expansin-like *Mc-EXP1*

A total of 12,218 ESTs generated from four cDNA libraries from *M. chitwoodi* were assembled into contigs in the Washington University Nematode EST Project. These contigs were queried with the sequence of the β-expansin Gr-EXPB1 from *G. rostochiensis* to identify homologous sequences in *M. chitwoodi*. Contig MC01203

showed similarity to the N-terminal amino acid region 49 to 160 of Gr-EXPB1. This matching region comprises the carbohydrate-binding module family II (CBM2), the linker and the first ten amino acids of the expansin domain of Gr-EXPB1 (Kudla *et al.*, 2005; Qin *et al.*, 2004). Contig MC01203 includes 25 ESTs of which two originate from the egg stage and 23 from the ppJ2 stage of *M. chitwoodi*. These latter 23 ESTs constitute 0.57% of the total number of ESTs (4,067) generated from the ppJ2 cDNA library. MC01203 is in the top 20 of most abundantly expressed members of the secretome of ppJ2-s of *M. chitwoodi* (see: Table 4 in Chapter 2 of this thesis).

The full-length sequence, obtained from MC01203 clone inserts and hereafter named *Mc-EXP1*, consists of 1219 nt including a polyA-tail. *Mc-EXP1* codes for a protein of 346 amino acids with a molecular weight of 37.3 kDa and a pI of 8.31 (Table 2). SignalP analysis predicted an N-terminal signal peptide for secretion with its most likely cleavage site between amino acids 21 and 22.

A BLASTP search with the Mc-EXP1 sequence as a query on the *nr* database further confirmed its similarity with the expansin Gr-EXPB1, and with the putative expansin Gr-EXPB2 from *G. rostochiensis* (EMBL accession CAC84564; E-values $2e^{-25}$ and $9e^{-23}$ respectively). Amino acid region 47 to 156 of Mc-EXP1 shares 30% identity with amino acids 49 to 160 of expansin Gr-EXPB1, which entails the CBM2. This N-terminal region of Mc-EXP1 (aa 37 to 119) also shares 31 to 42% identity with the CBMs in β-1,4-endoglucanases from several cyst and root-knot nematode species (Fig. 1). These CBMs are all classified as bacterial family 2 CBMs, which are characterized by two highly conserved cysteine residues positioned at each extremity of the domain (Gilkes *et al.*, 1991a). These conserved cysteines are also present at each extremity of the N-terminal domain of Mc-EXP1. Next to the two conserved cysteines, there are four conserved tryptophans (W) in bacterial family 2 CBMs of which two are involved in cellulose binding (Gilkes *et al.*, 1991b). These latter two residues are also conserved in CBMs of plant-parasitic nematodes found to date, however, none of these two tryptophans is present in the CBM of Mc-EXP1 (Fig. 1). A third characteristic of bacterial family 2 CBMs are the conserved asparagine (N) and glycine (G) residues, of which most are also conserved in the CBM of Mc-EXP1 (Fig. 1).

Table 2 Overview of sequence characteristics of expansin-like proteins identified from *M. chitwoodi*. Listed are the number of expressed sequence tags (ESTs) representing each expansin-like protein in the EST dataset from *M. chitwoodi*. The figures of the deduced protein sequences include the predicted signal peptide for secretion. Similarity searches were performed using BLASTP on the non-redundant database at NCBI using the protein sequences minus signal peptide for secretion as search queries

Name	ESTs	Length (nt)	Protein (aa)	SP cleavage site	MW (kDa)	pI	Similarity	E-value
Mc-EXP1	25	1219	346	21/22	37.3	8.31	EXPB2 protein CAC84564 (*Globodera rostochiensis*)	$2e^{-25}$
Mc-EXP2	1	731	215	18/19	23.7	5.6	Hypothetical protein CAC42207 (*Amycolatopsis mediterranei*)	$3e^{-28}$
Mc-EXP3	1	679	132	22/23	14.8	8.87	Unknown AAD32751 (*Streptomyces lavendulae*)	$7e^{-25}$
Mc-EXP4	63	963	212	20/21	24.3	8.86	Putative avirulence protein CAC27774 (*Meloidogyne incognita*)	$7e^{-27}$

Abbreviations: (nt) nucleotides, (aa) amino acids, (SP) signal peptide, (MW) molecular weight, (kDa) kilodalton, (pI) isoelectric point.

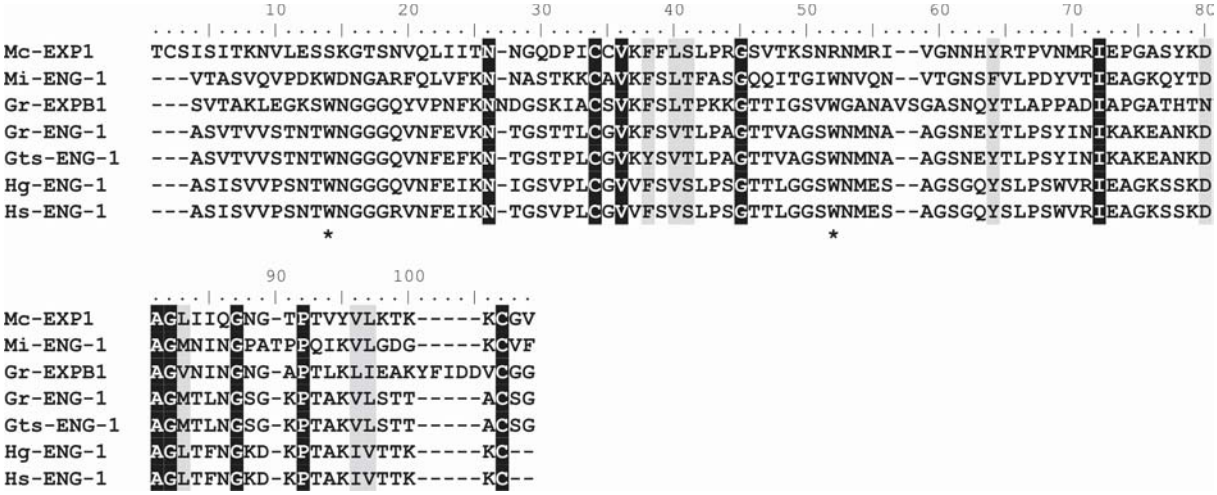


Fig. 1 Alignment of the putative carbohydrate-binding module (CBM) of Mc-EXP1 from *M. chitwoodi* with CBMs from other plant-parasitic nematode species. Mi-ENG-1: a β-1,4-endoglucanase from *M. incognita* (GenBank accession no. AAD45868), Gr-EXPB1: an expansin from *Globodera rostochiensis* (CAC83611), Gr-ENG-1: a β-1,4-endoglucanase from *G. rostochiensis* (AAC63988), Gts-ENG-1: a β-1,4-endoglucanase from *Globodera tabacum solanacearum* (AAD56392), Hg-ENG-1: a β-1,4-endoglucanase from *Heterodera glycines* (AAC48327), Hs-ENG-1: a β-1,4-endoglucanase from *Heterodera schachtii* (CAC12958). Residues identical in all CBMs are shaded black and residues that are similar in all CBMs are shaded grey. The tryptophans (W) that are putatively involved in cellulose binding are indicated with an asterisk.

The C-terminal amino acid region 208 to 342 of Mc-EXP1 has 48% and 46% identity with the expansin domains of Gr-EXPB1 and Gr-EXPB2 respectively. Significant similarity was also found between the C-termini of Mc-EXP1 and the putative avirulence protein MAP-1 of *Meloidogyne incognita* (EMBL accession CAC27774; E-value 2e-18 and identity 41%) (Semblat *et al.*, 2001). Furthermore, the C-terminus in Mc-EXP1 was similar to a hypothetical protein from *Amycolatopsis mediterranei* (EMBL accession CAC42207; 33% identity) and an unknown protein from *Streptomyces lavendulae* (EMBL accession AAD32751; 34% identity). The sequence identity between the C-terminus of Mc-EXP1 and plant expansins ranged from 30% to 35%. Remarkably, whereas Gr-EXPB1 shares highest similarity with (putative) β-expansins from plants, Mc-EXP1 shares highest similarity with α-expansins from plants (Table 3; Kudla *et al.*, 2005). A local alignment of Mc-EXP1 with Gr-EXPB1, two (putative) β-expansins and two α-expansins from plant species indicated the presence of a series of conserved amino acid residues, among which are a number of conserved cysteines previously marked as key signatures of the expansin family (Cosgrove, 2000b). Three of the six cysteines that are conserved in plant expansins (C1, C3 and C5 in Fig. 2) are also present in Mc-EXP1 and Gr-EXPB1. Three other cysteines are conserved in Mc-EXP1 and Gr-EXPB1 at amino acid position 72, 94 and 97 (in the alignment of Figure 2). The cysteine at position 94 is also present in the plant expansins, but is not indicated as a characteristic conservation (Cosgrove, 2000b). The HFD motif, another key motif of the plant expansin family (Cosgrove, 2000b), is present as $H_{312}VD_{314}$ in Mc-EXP1. Again, Mc-EXP1 is showing more similarity to the corresponding motifs in Gr-EXPB1 and in Hg-EXPB1 (Kudla *et al.*, 2005). The CBM and the expansin-like domain of Mc-EXP1 are linked by approximately 85 amino acids, which are relatively rich in serine (S), asparagine (N) and threonine (T). Five consensus N-glycosylation sites are present in Mc-EXP1 of which two are located in the linker sequence (N_{130} to T_{133} and N_{181} to L_{184}) and three in the expansin-like domain (N_{213} to G_{216}, N_{294} to L_{297} and N_{334} to I_{337}).

A further conserved domain search with Mc-EXP1 indicated additional similarity between the C-terminus of the protein and the rare lipoprotein A (RlpA)-like double-psi beta-barrel domain (DPBB_1; pfam03330; E-value of e^{-6} in CDART). The DPBB fold is classified as the Barwin-like endoglucanases superfamily in the Structural Classification of Proteins (SCOP; Andreeva *et al.*, 2004; Murzin *et al.*, 1995).

Table 3 Most significant (putative) expansin protein matches found with BLASTP for (A) the expansin domain of expansin Gr-EXPB1 from *G. rostochiensis* and (B) expansin-like domain of Mc-EXP1 from *M. chitwoodi*. Gr-EXPB1 matches with (putative) beta-expansins from plants, whereas Mc-EXP2 matches with (putative) alpha-expansins from plants

(A) Gr-EXPB1	Accession number	%ID	%Sim
ATEXPB5 (*Arabidopsis thaliana*)	NP_191616	37% (36/95)	47% (45/95)
PPAL beta-expansin-like protein (*Nicotiana tabacum*)	AAG52887	37% (44/117)	47% (55/117)
(B) Mc-EXP1			
Gr-EXPB1 (*Globodera rostochiensis*)	CAC83611	47% (55/116)	60% (70/116)
Cs-EXPA1 expansin S1 precursor (*Cucumis sativus*)	AAB37746	31% (37/119)	42% (50/119)
EXPA1 expansin (*Plantago major*)	CAJ38385	33% (34/102)	44% (45/102)
Expansin 2 (*Zinnia elegans*)	AAF35901	31% (37/118)	43% (51/118)
OsEXPA26 alpha- expansin (*Oryza sativa*)	AAL24497	35% (39/109)	43% (47/109)
EXP1 alpha-expansin precursor (*Gossypium hirsutum*)	AAM46997	32% (36/111)	39% (44/111)

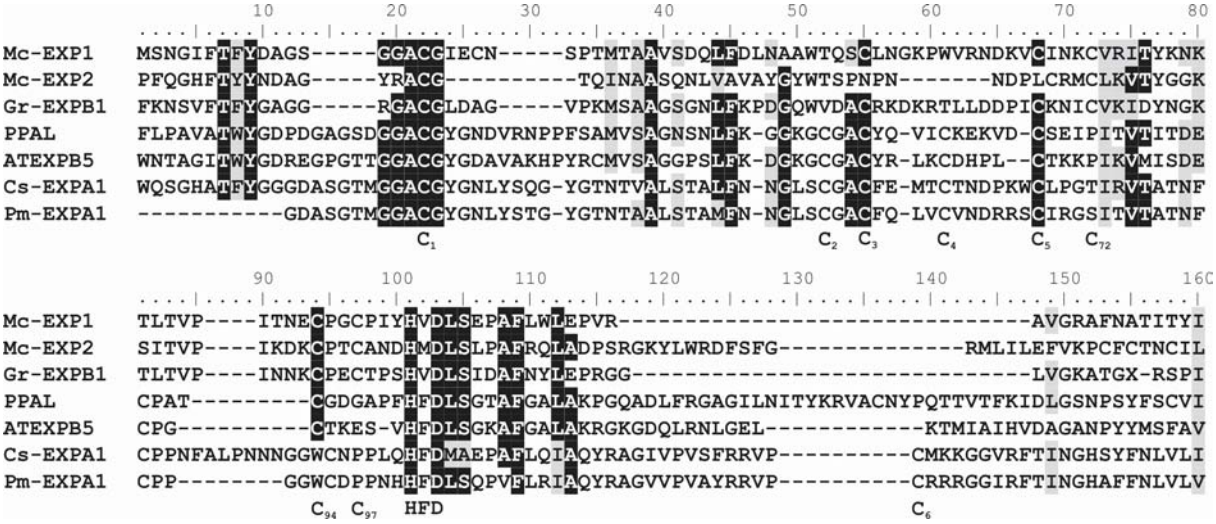


Fig. 2 Alignment of the expansin(-like) domain of Mc-EXP1 and Mc-EXP2 from *M. chitwoodi* with Gr-EXPB1: a functional expansin from the potato cyst nematode *Globodera rostochiensis* (GenBank accession no. CAC83611), PPAL: a pistil-specific β-expansin-like protein from *Nicotiana tabacum* (AAG52887), ATEXPB5: a putative β-expansin from *Arabidopsis thaliana* (NP_191616), Cs-EXPA1: expansin S1 precursor from *Cucumis sativus* (AAB37746) and Pm-EXPA1: an expansin from *Plantago major* (CAJ38385). Residues identical to those of Mc-EXP1 and/or Mc-EXP2 are shaded black and similar residues are shaded grey. The cysteines (C) indicated with C1 to C6 and the HFD motif are conserved in both α-expansins and β-expansins and are used as the key signatures of the expansin family (Cosgrove, 2000b). The cysteines indicated with C72, C94 and C97 are cysteines that are conserved in the expansin(-like) proteins from *M. chitwoodi* and *G. rostochiensis*. The alignment shows that only the H and D residues of the HFD motif are conserved in the expansin(-like) proteins from *M. chitwoodi* and *G. rostochiensis*.

Mc-EXP1 is expressed and translated in the subventral oesophageal glands

To localize the expression of *Mc-EXP1* in the nematode, *in situ* hybridizations were performed with two antisense probes matching either part of the CBM2 and linker, or part of the linker and expansin-like domain (probes a and b respectively in Table 1). Both probes designed on *Mc-EXP1* localized the gene transcription in the subventral oesophageal glands of ppJ2-s of *M. chitwoodi* (Fig. 3). No signal was observed with the corresponding sense probes.

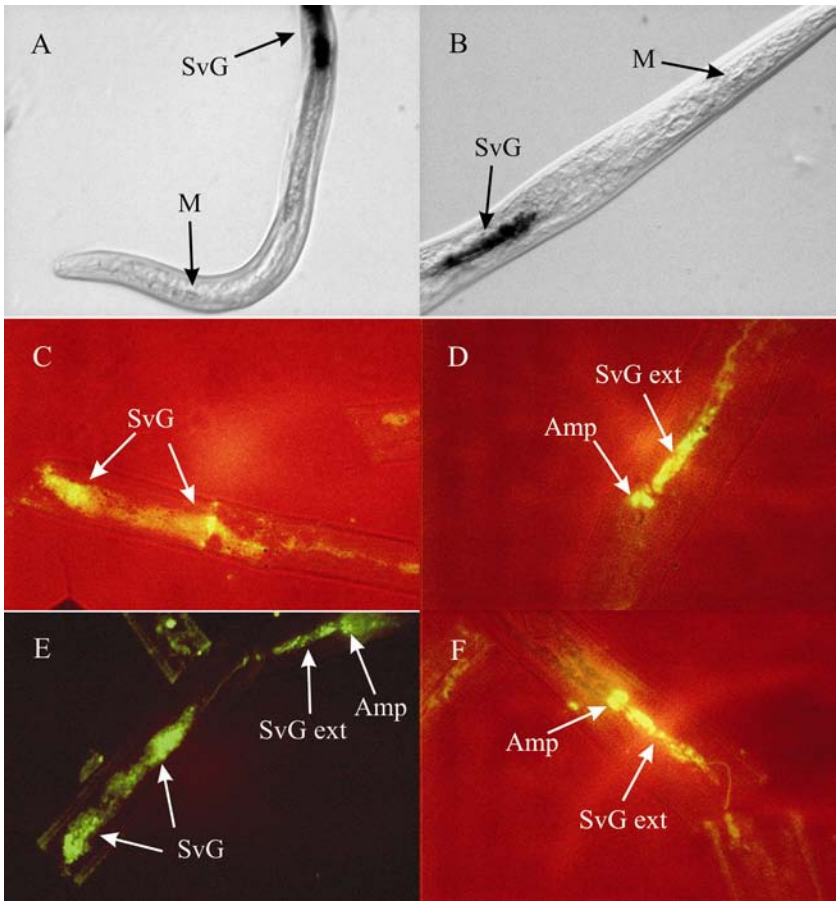


Fig. 3 *In situ* localization of expansin-like proteins from *M. chitwoodi* using *in situ* hybridization and immunofluorescence microscopy on pre-parasitic second stage juveniles. (A + B) Localization of expansin-like protein encoding transcripts using digoxigenin-labelled cDNA probes and an alkaline phosphatase-conjugated secondary antibody to digoxigenin. Sections of the nematode incubated with antisense probe designed on *Mc-EXP1* (A) and *Mc-EXP3* (B) showing a specific alkaline phosphatase staining of the subventral oesophageal gland region. (C - F) Immunofluorescence labelling of the subventral oesophageal glands with rabbit antisera to Mc-EXP1 (C and D) and Mc-EXP2 (E and F). Arrows point at the subventral oesophageal gland cell lobes (SvG), subventral gland cell extensions (SvG ext) and ampullae (Amp), and the metacorpus (M).

Immunofluorescence microscopy with specific antisera to Mc-EXP1 specifically localized the protein in the subventral oesophageal glands of *M. chitwoodi* ppJ2-s. A strong fluorescence was observed in the subventral oesophageal gland extensions and ampullae and to a lesser extent in the gland lobes (Fig. 3). The fluorescence signal in the gland extensions showed a clear granular pattern with apparent localization within the secretory granules. No fluorescent labelling was observed in the subventral oesophageal glands of nematode pieces incubated with pre-immune serum. In addition, no fluorescent signal was observed in nematode pieces treated solely with the FITC-conjugated goat anti-rabbit second antibody. The presence of a signal peptide for secretion and the localization of the Mc-EXP1 transcripts and protein in the subventral oesophageal glands of *M. chitwoodi* ppJ2-s strongly suggest that it is secreted by the nematode.

Identification and characterization of expansin-like *Mc-EXP2*

The second matching contig, named MC01774, includes only one EST (GenBank accession number CB831016) from the egg stage of *M. chitwoodi* and shares 36% identity with the C-terminal expansin domain (amino acid region 156 to 254) of Gr-EXPB1. We obtained a 731 nt sequence, hereafter named *Mc-EXP2*, by additional sequencing of the insert of library clone CB831016. The largest open reading frame in *Mc-EXP2* codes for a protein of 215 amino acids with a molecular weight of 23.7 kDa (Table 2). SignalP predicted the presence of an N-terminal signal peptide for secretion with its most likely cleavage site between alanine and cysteine at amino acid positions 18 and 19. The C-terminal region of Mc-EXP2 (amino acid 93 to 184) shares 36% identity and 51% similarity with the expansin-like domain of Mc-EXP1.

A BLASTP search on the *nr* database with Mc-EXP2 also showed similarity between this C-terminal region and a hypothetical protein from *A. mediterranei* (EMBL accession CAC42207; 63% identity) and an unknown protein from *S. lavendulae* (EMBL accession AAD32751; 55% identity). In addition, this C-terminal region of Mc-EXP2 shares approximately 50% identity with two hypothetical proteins (GenBank accession numbers EAL65201 and EAL65166) from the slime mould *Dictyostelium discoideum* and approximately 40% identity with hypothetical proteins from a range of fungal species. The highest similarity with plant expansins was found with the α-expansin PpEXPA6 from the moss *Physcomitrella patens* (Genbank accession AAN08122; 32% identity and 41% similarity) and with the α-expansin OsEXPA26 from *Oryza sativa* (Genbank accession AAL24497; 31% identity and 46% similarity). A local alignment indicated that Mc-EXP2 contains a similar series of conserved cysteines as Mc-EXP1 and Gr-EXPB1, only the cysteine at position C3 is lacking in Mc-EXP2 (Fig. 2). As was found for Mc-EXP1 and the expansins in *G.*

rostochiensis and *H. glycines*, only the histidine (H) and aspartic acid (D) residues from the HFD motif are conserved in Mc-EXP2 ($H_{167}MD_{169}$).

A further conserved domain search pointed at the signatures of two other protein domains in Mc-EXP2. Similarity was found between the Mc-EXP2 C-terminal amino acid region 150 to 185 and part of the DPBB_1 domain (E-value $1e^{-8}$ in CDART). In addition, the N-terminal 45 amino acids of the mature Mc-EXP2 showed similarity with a so-called Lysin-motif domain (LysM; CDART accession 00118.2; 50% identity and 70% similarity). In the linker region spanning amino acids 67 to 93 between the putative LysM and expansin domain, an N-glycosylation site is predicted at amino acid position N_{83} to I_{86}.

Mc-EXP2 was identified in a library made from nematode eggs. Attempts to amplify *Mc-EXP2* from the cDNA library made from *M. chitwoodi* ppJ2-s failed. In addition, *in situ* hybridization on *M. chitwoodi* ppJ2-s with an antisense probe spanning the N-terminal putative LysM domain did not result in a hybridization signal either.

Mc-EXP2 proteins are subventral oesophageal gland specific in *M. chitwoodi* ppJ2-s

Immunofluorescence microscopy with the Mc-EXP2 antiserum resulted in a strong fluorescent labelling in the subventral oesophageal glands, including the extensions and ampullae of the gland cells (Fig. 3). No fluorescent labelling was observed in the subventral oesophageal glands of nematode pieces incubated with pre-immune serum. In some nematode pieces treated with Mc-EXP2 antiserum, the amphids or genital primordium were fluorescently labelled, but this was also observed in case of the treatment with pre-immune serum (not shown). Therefore, we conclude that the Mc-EXP2 protein is specifically produced in the subventral oesophageal glands of *M. chitwoodi* ppJ2-s.

Mc-EXP1 and *Mc-EXP2* are members of a diverse family

Further screening of ESTs from *M. chitwoodi* with Gr-EXPB1 resulted in the identification of two additional contigs encoding expansin-like proteins. Contig MC01839 contains one EST (Genbank accession number CB831222) from the egg stage of *M. chitwoodi*. The full-length sequence (*Mc-EXP3*) obtained from the corresponding cDNA library clone is 679 nt including a polyA tail and is coding for an open reading frame of 132 amino acids (Table 2). SignalP analysis predicted an N-terminal signal peptide for secretion with its most likely cleavage site between amino acids 22 and 23. No consensus sites for N-linked glycosylation were predicted for Mc-EXP3. A BLASTP search on the *nr* database with the mature Mc-EXP3 protein

showed highest similarity with an unknown protein from *S. lavendulae* (Genbank accession AAD32751; 47% identity) and a hypothetical protein from *A. mediterranei* (Genbank accession CAC42207; 46% identity). The highest similarity with expansin-like sequences, besides Gr-EXPB1, was found with the expansin-like protein *Dd*EXPL1 (Genbank accession XP_647352) from the slime mould *D. discoideum* (31% identity). *D. discoideum* has a cellulosic cell wall and *Dd*ExpL1, together with other expansin-like genes, was found to be expressed throughout the development of this organism (Darley *et al.*, 2003). Mc-EXP3 contains a series of conserved cysteines, which is identical to Gr-EXPB1 and Mc-EXP1, and the HFD signature motif is present as an $H_{103}AD_{105}$ motif in Mc-EXP3 (data not shown). A conserved domain search with the mature Mc-EXP3 indicated similarity in amino acid region 6 to 104 with the RlpA-like DPBB_1 domain (pfam03330) with an E-value of $9e^{-8}$ in CDART. *In situ* hybridization with anti-sense probe derived from *Mc-EXP3* resulted in specific staining of the subventral oesophageal glands in *M. chitwoodi* ppJ2-s, while no staining was observed with the corresponding sense probe (Fig. 3).

Contig MC01221 also matched Gr-EXPB1 and contains 63 ESTs originating from both the egg (8 ESTs) and ppJ2 (55 ESTs) stage of *M. chitwoodi*. Contig MC01221 belongs to the top 10 of most abundantly represented transcripts in the *M. chitwoodi* EST dataset (see Table 3 in Chapter 2 of this thesis). By additional sequencing of inserts of clones belonging to this contig, we obtained the full-length sequence of 963 nt, including the polyA tail. The largest open reading frame in this sequence codes for a protein of 212 amino acids, hereafter named Mc-EXP4, with a molecular weight of 24.3 kDa (Table 2). The SignalP analysis predicted the presence of an N-terminal signal peptide for secretion with its most likely cleavage site between amino acid position 20 and 21. A BLASTP search with the mature Mc-EXP4 sequence as query on the *nr* database showed highest similarity to the C-terminus (amino acid 305 to 453) of the putative avirulence protein MAP-1 from *M. incognita* (E-value $7e^{-27}$, 47% identity and 61% similarity) (Semblat *et al.*, 2001). In addition, significant similarity was found with the expansin domain of expansins Gr-EXPB1 (31% identity) and Gr-EXPB2 (34% identity). Mc-EXP4 includes three (C1, C3 and C5) of the six conserved cysteines that are characteristic for plant expansins and, in addition, the HFD motif is present as $H_{180}LD_{182}$ (data not shown). The conserved domain search and InterProScan with Mc-EXP4 did not detect any of the conserved domains found in Mc-EXP1, Mc-EXP2 and Mc-EXP3. Despite its high abundance, no hybridization signal was detected by *in situ* hybridization in ppJ2-s of *M. chitwoodi* with an antisense probe matching nucleotide region 458 to 593.

Expansin-like proteins are wide-spread among members of the order Tylenchida

Searching EST datasets from other root-knot and root-lesion nematode species with the expansin domain of Gr-EXPB1 resulted in the identification of three matching contigs from *M. arenaria* (MA00256, MA00532, MA03038), five matching contigs from *M. hapla* (MH02133, MH02764, MH07582, MH07998, MH10906), two matching contigs from *M. incognita* (MI00267, MI01406), two matching contigs from *M. javanica* (MJ00856, MJ04592), three matching contigs from *M. paranaensis* (MP00467, MP00635, MP00661), and one matching EST from *Pratylenchus vulnus* (GenBank accession number CV199817) (Table 4). The level of identity between the expansin domain in Gr-EXPB1 and the matching contigs and EST ranged from 41% to 67%. For the translations of the *P. vulnus* EST and all consensus sequences of the root-knot nematode contigs, except for MH07582, MH07998, MJ04592, and MJ00856, an N-terminal signal peptide for secretion was predicted. In addition, for none of the sequences we have found evidence for ancillary domains next to the expansin domain. The alignment of the expansin domain of Gr-EXPB1 with the root-knot nematode expansin-like domains indicated that these domains share a similar series of conserved cysteines along the protein backbone (Fig. 4). As was found for the *M. chitwoodi* expansin-like proteins and Gr-EXPB1, only the H and D of the HFD motif found in plant expansins around position 110 are conserved in the root-knot nematode sequences (Fig. 4).

Further inspection of the *M. paranaensis* contigs MP00467, MP00635 and MP00661 revealed that the transcripts code for proteins that only differ at three amino acid positions in the predicted signal peptide for secretion, while the serine (S) at amino acid position 65 in contigs MP00635 and MP00661 is changed into asparagine (N) in contig MP00467 (data not shown). Remarkably, the alignment of the contig consensus sequences showed that the mature expansin-like proteins in MI01406 from *M. incognita* and contig MP00635 from *M. paranaensis*, two closely related root-knot nematode species (*M. Holterman, personal communication*), are 100% identical at the amino acid level (Fig. 4). Even in their predicted N-terminal signal peptides for secretion, a sequence that is being cleaved off before secretion, these two proteins only differ at two amino acid positions. The coding sequences of these two contigs share 97% identity at the nucleotide level.

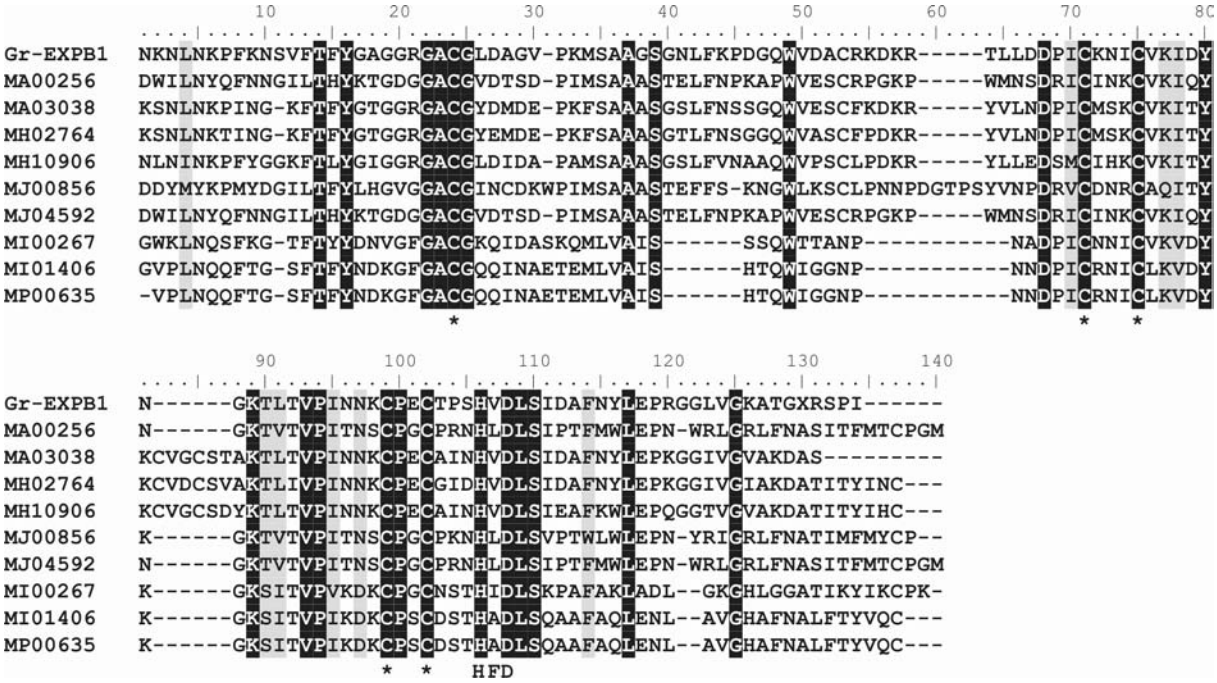


Fig. 4 Alignment of the expansin domain of Gr-EXPB1 from potato cyst nematode *Globodera rostochiensis* (GenBank accession no. CAC83611) with the amino acid sequences deduced from contig consensus sequences from the root-knot nematode species *Meloidogyne arenaria* (MA00256, MA03038), *M. hapla* (MH02764, MH10906), *M. javanica* (MJ00856, MJ04592), *M. incognita* (MI00267, MI01406) and *M. paranaensis* (MP00635). Residues identical in all sequences are shaded black and those that are similar in all sequences are shaded grey. The cysteines (C) conserved in all the expansin(-like) protein sequences are indicated with an asterisk. The position of the HFD motif (conserved in α-expansins and β-expansins and used as key signature of the expansin family (Cosgrove, 2000b)) is indicated, showing that only the H and D residues are conserved in the sequences from root-knot and cyst nematode species.

Expansin-like proteins in other cellulolytic organisms

The expansin-like sequences identified in this study share more sequence similarity with two hypothetical proteins from the saprophytic soil bacteria *A. mediterranei* and *S. lavendulae* than with plant expansins. This finding suggests that the distribution of expansins outside the plant kingdom may not be limited to the phylum Nematoda. To investigate if expansin-like proteins are also produced in other cellulose degrading micro-organisms, we conducted TBLASTN searches on fungal and oomycete sequence data using the expansin domain of Cs-EXPA1 from *C. sativus* as query (http://www.broad.mit.edu/annotation/fgi/ and http://www.broad.mit.edu/annotation/genome/ phytophthora_infestans/). The search on the fungal sequence database resulted in the identification of expansin-like proteins from *Magnaporthe grisea* [MGG_07556.5], several *Fusarium* species (*F. graminearum* [FGSG_10922.3], *F. verticillioides* [FVEG_10241.3], and *F. oxysporum* [FOXG_11583.2]) and several *Aspergillus* species (*A. fumigatus* [Afu5g08030], *A. nidulans* [AN7735.3], *A. niger* [est_GWPlus_C_140158], *A. clavatus* [ACLA_012490], *A.*

terreus [ATEG_08323.1], and the closely related *Neosartorya fischeri* [NFIA_079270]) (Table 4). All these fungal expansin-like proteins share 25 to 30% identity with plant expansin domains. In addition, these proteins, except those from *A. niger* and *A. terreus*, were predicted to contain an N-terminal signal peptide for secretion, another characteristic they share with plant expansins. Finally, we found a sequence (PITG_10702.1) from the plant pathogenic oomycete *Phytophthora infestans* that shares approximately 25% sequence identity with plant expansin domains and which also contains a predicted N-terminal signal peptide for secretion.

Table 4 Expansin-like sequences identified from the root-lesion nematode *Pratylenchus vulnus*, several fungal species and the plant pathogenic oomycete *Phytophthora infestans*. TBLASTN search queries were the expansin domain of β-expansin Gr-EXPB1 (CAC83611) from *G. rostochiensis* and the expansin domain of Cs-EXPA1 (AAB37746) from *Cucumis sativus*. Listed are the percentage of identity (%ID), the percentage of similarity (%Sim) and the presence (+) or absence (-) of a predicted signal peptide for secretion (SP)

Search query	Matches				
	Organism	Accession number	%ID	%Sim	SP
Gr-EXPB1	*Pratylenchus vulnus*	CV199817	54 (54/100)	59 (59/100)	+
Cs-EXPA1	*Magnaporthe grisea*	MGG_07556.5	29 (33/112)	43 (48/112)	+
	Fusarium graminearum	FGSG_10922.3	30 (33/108)	45 (49/108)	+
	F. verticillioides	FVEG_10241.3	29 (32/108)	43 (47/108)	+
	F. oxysporum	FOXG_11583.2	29 (32/108)	43 (47/108)	+
	Aspergillus fumigatus	Afu5g08030	30 (33/108)	42 (46/108)	+
	A. nidulans	AN7735.3	27 (30/108)	41 (45/108)	+
	A. niger	est_GWPlus_C_140158	26 (29/108)	42 (46/108)	-[a]
	A. clavatus	ACLA_012490	26 (29/108)	40 (44/108)	+
	A. terreus	ATEG_08323.1	24 (26/108)	39 (43/108)	-
	Neosartorya fischeri	NFIA_079270	29 (32/108)	41 (45/108)	+
	Phytophthora infestans	PITG_10702.1	26 (32/121)	40 (49/121)	+

[a] transcript misses 5'-end

Discussion

Expansin, expansin-like or expansin-related proteins in *M. chitwoodi*

In this paper we described the identification and characterization of four novel proteins from the root-knot nematode *M. chitwoodi* with similarity to the expansins from cyst nematodes and from various plant species. The proteins Mc-EXP1 and Mc-EXP2 are two-domain proteins with remote similarity to GHF45 members in one domain, and similarity either to group-II grass pollen allergens (Mc-EXP1) or to LysM domains (Mc-EXP2) in a separate domain. In contrast, Mc-EXP3 and Mc-EXP4 are single-domain proteins including only GHF45-like features. The nomenclature of the expansin superfamily, as was proposed by the ad-hoc working group on expansin nomenclature (Kende *et al.*, 2004), comprises four monophyletic classes designated as expansin A (EXPA), expansin B (EXPB), expansin-like family A (EXLA) and expansin-like family B (EXLB). Only proteins that contain both a domain with distant sequence similarity to the catalytic domain of members of GHF45 (GHF45-like domain) and a domain with remote sequence similarity to the so-called group-II grass pollen allergens (a putative carbohydrate-binding domain) are classified in this system as expansins or expansin-like proteins. Furthermore, a separate 'catch-all' category of expansin-related sequences has been proposed for proteins of non-plant organisms with structural similarity to both the GHF45-like domain and the group-II grass pollen allergens, but that are not in monophyly with the established plant gene families. Consequently, the current four designated classes in the nomenclature of expansins are reserved for plant proteins only. Two related pH-dependent activities, namely long-term irreversible extension of plant cell walls and enhanced stress relaxation, are further used as diagnostic markers for expansins. However, only a small proportion of the plant genes that are currently being classified into the four categories above have been positively tested for these activities. This is in part due to the fact that the production of active recombinant expansins has proven largely unsuccessful (Li *et al.*, 2003). Gr-EXPB1 from *G. rostochiensis* has all the features of expansin B-s (β-expansins) including a specific cell wall extension activity on wheat coleoptiles, and was therefore designated by us as a β-expansin. In the case of Mc-EXP1, Mc-EXP2, Mc-EXP3 and Mc-EXP4, cell wall extension activity could not be tested in extensometer assays because attempts to produce recombinant proteins have failed to date. The *M. chitwoodi* sequences harbour most of the signature motifs of α- and β-expansins, but as long as we are not able to test the activity of these proteins on plant tissues, the sequences will be designated as expansin-like proteins.

Are *M. chitwoodi* expansin-like proteins α- or β-expansins?

Since *M. chitwoodi* has a very broad host range, comprising both monocots (including grasses) and dicots, it is likely to benefit from both α- and β-expansins. The latter because the α-expansins are more effective on cell walls of dicots (Type I) whereas β-expansins have more effect on grass cell walls (Type II) (Cosgrove, 2000b; Cosgrove *et al.*, 1997a). It is difficult to predict whether the expansin-like proteins found in *M. chitwoodi* are functionally more similar to α- or β-expansins solely based on amino acid sequences. Mc-EXP1 and Mc-EXP2 from *M. chitwoodi* share highest sequence similarity with α-expansins from plants. But in contrast to α-expansins in plants, which are not glycosylated, N-glycosylation consensus sites were found in both Mc-EXP1 (five) and Mc-EXP2 (one). This suggests that Mc-EXP1 and Mc-EXP2 are more similar to the β-expansins, which are known to be glycosylated. However, the degree of glycosylation may not hold as a discriminative trait for expansins outside the plant kingdom. Conclusive evidence on the type of activity of Mc-EXP1 and Mc-EXP2 requires the production of active recombinant protein and subsequent cell wall extension assays on the different types of cell walls.

Domain structure variants

The expansin-like proteins in *M. chitwoodi* represent three different protein domain variants. Mc-EXP1 and Mc-EXP2 both contain two domains, i.e. an expansin-like domain C-terminally linked to either (i) a bacterial family 2 CBM (Mc-EXP1) or (ii) a LysM-like domain (Mc-EXP2). In contrast, Mc-EXP3 and Mc-EXP4 consist of an expansin-like domain alone. Compared to the domain structure of plant expansins, the expansin-like domain and the putative CBM of Mc-EXP1 are in reverse orientation. Thus, the domain structure of Mc-EXP1 is similar to that of Gr-EXPB1. However, the signature tryptophans in bacterial family 2 CBMs, of which two are involved in cellulose binding, are not present in the predicted CBM2 of Mc-EXP1 and therefore it remains to be tested if Mc-EXP1 indeed binds to cellulose.

In case of Mc-EXP2, the N-terminal region resembles the LysM domain. The latter is a widespread domain, mainly found in bacterial proteins, among which is a variety of bacterial cell wall degrading enzymes and proteins involved in bacterial pathogenesis (Bateman and Bycroft, 2000). The LysM domain may have a general peptidoglycan, a structural component of bacterial cell walls, binding function. In bacterial cell wall degrading enzymes it probably serves to anchor the catalytic domains to their substrates. LysM domains are not only part of bacterial proteins but are also present in a number of eukaryotic proteins, e.g. in chitinases from *C. elegans* (Ponting *et al.*, 1999). The function of the putative LysM domain in Mc-EXP2 is unclear and remains to be investigated.

As mentioned above, the proposed definition of expansins includes the prerequisite that the protein should contain both a GHF45-like domain (expansin domain) and a domain with remote sequence similarity to the so-called group-II grass pollen allergens, i.e. a putative CBM. In the light of this definition, Mc-EXP2, Mc-EXP3 and Mc-EXP4 cannot be designated as expansins since they lack a CBM. The absence or presence of a CBM will most probably influence the substrate specificity and affinity of the protein, but one can question whether the presence of an ancillary CBM is crucial for the expansin-like activity of the proteins. Firstly, similar domain architectures, i.e. a catalytic domain with and without a CBM, were also found for β-1,4-endoglucanases identified from *M. chitwoodi* (Chapter 3 of this thesis) and other plant-parasitic nematode species (Rosso *et al.*, 1999; Smant *et al.*, 1998; Uehara *et al.*, 2001). For the β-1,4-endoglucanases of *G. rostochiensis* and *H. glycines*, the absence of the CBM did not influence the activity of the enzyme on carboxymethylcellulose (Smant *et al.*, 1998). Secondly, the β-expansins (group-I allergens) found in grass pollen do not bind tightly to the pollen wall but do have an expansin-like activity. This suggests that strong binding is not essential for expansin activity (Cosgrove *et al.*, 1997a). Lack of high cell wall affinity means that the expansin protein can diffuse to adjacent cells. Cosgrove (1999) postulated that the latter may actually be important for proper protein function when the aim is to secrete a protein that will loosen the walls of neighbouring cells, e.g. when a pollen tube penetrates stigma and style (Cosgrove, 1999). The process of pollen tube penetration is functionally similar to the intercellular migration of root-knot nematodes through the plant roots, therefore, the tight binding of expansin-like proteins secreted by these nematode species may also not be required. Only activity assays with the Mc-EXPs can reveal whether the absence of a CBM influences the expansin-like activity of the protein. The results of these experiments can be used to see whether the current definition of expansins needs revision.

Conserved expansin motifs

Regularly spaced conserved cysteines (C) and the HFD motif around amino acid position 110 are used as the key signatures of the plant expansin superfamily (Cosgrove, 2000b). A series of conserved cysteines is also present along the backbone of the root-knot nematode expansin-like domains. In case of the HFD motif, only the H and D residues are conserved in the expansin-like proteins of root-knot nematodes. In most of these sequences, the aromatic phenylalanine (F) is changed into one of the aliphatic amino acids valine (V), isoleucine (I), or leucine (L). A similar conservation as well as an F to V conversion was found in the β-expansin Gr-EXPB1 from *G. rostochiensis*, for which cell wall expansion activity on type II primary cell walls was found (Kudla *et al.*, 2005; Qin *et al.*, 2004). This indicates that the F to V

conversion does not abolish the typical expansin activity of the protein, which implies that the functional significance of the HFD motif is obscure. Clarification of its functional significance requires biochemical characterization of the root-knot nematode expansin-like proteins, e.g. by site directed mutagenesis followed by activity assays.

Expansins and expansin-like proteins in plant-nematode interactions

For a long time, expansins as a group seemed to be restricted to land plants (Cosgrove, 2000a). More recently, a small number of expansin-like sequences was identified in the slime mould *D. discoideum* (Darley *et al.*, 2003; Li *et al.*, 2002). The Dictyostelia are positioned at the base of the division between the Animalia and the Fungi and make a cellulosic cell wall, which is a characteristic they share with plants (Baldauf *et al.*, 2000; Darley *et al.*, 2003). The finding of expansin(-like) sequences in both Dictyostelia and plants led to the suggestion that expansins must have been present in a common ancestor of these organisms, but were lost in organisms that do not possess cellulose-based cell walls (Darley *et al.*, 2003). This hypothesis was rejected by our previous finding of the functional β-expansin Gr-EXPB1 in cyst nematodes (Kudla *et al.*, 2005; Qin *et al.*, 2004). The data presented in this study show that expansins and expansin-like proteins are not a rare evolutionary innovation of cyst nematodes. It now seems that the evolutionary most advanced biotrophic root-knot nematode and the more basal necrotrophic root-lesion nematode species include homologs of Gr-EXPB1 in their repertoire of cell wall modifying proteins. The expansins and expansin-like proteins are likely to be co-secreted from the oesophageal gland cells of these parasites to collectively weaken and breakdown the cell walls of host plants during plant invasion.

Expansins and expansin-like proteins are likely to be involved in diverse plant-microbe interactions

A limited number of proteins with similarity to expansins has been found in a diversity of taxa outside the plant kingdom, including the blue mussel *Mytillus edulis* (Xu *et al.*, 2000), the fungus *Trichoderma reesei* (Saloheimo *et al.*, 2002) and the bacterial species *Clavibacter michiganensis* ssp. *sepedonicus* (Laine *et al.*, 2000), *Bacillus subtilis* and *Xylella fastiosa* (Li *et al.*, 2002). The recent sequencing initiatives on fungal and oomycete genomes provide a whole new resource to be mined for expansins and expansin-like proteins. We used this resource to identify expansin-like sequences from the fungi *M. grisea*, several *Aspergillus* and *Fusarium* species, and from the plant pathogenic oomycete *P. infestans*. Furthermore, the best hits in the *nr* database for the Mc-EXPs and Gr-EXPB1 are two hypothetical proteins from the saprophytic soil

bacteria *A. mediterranei* and *S. lavendulae*. Despite the lack of experimental evidence for the typical cell wall-extension activity of the expansin-like proteins identified in this study we believe that expansin activity is likely to feature in many plant-microbe interactions.

5

A symbiont-independent endo-1,4-β-xylanase from the plant-parasitic nematode *Meloidogyne incognita*

Makedonka Mitreva-Dautova, Erwin Roze, Hein Overmars, Leo de Graaff, Arjen Schots, Johannes Helder, Aska Goverse, Jaap Bakker and Geert Smant

Molecular Plant-Microbe Interactions (2006) 19 (5), pp. 521-529

SUMMARY

Substituted xylan polymers constitute a major part of the hemicellulose fraction of plant cell walls, especially in monocotyledons. Endo-1,4-β-xylanases (EC 3.2.1.8) are capable of hydrolyzing substituted xylan polymers into fragments of random size. Many herbivorous animals have evolved intimate relationships with endosymbionts to exploit their enzyme complexes for the degradation of xylan. Here we report the first finding of a functional endo-1,4-β-xylanase gene from an animal. The gene (*Mi-xyl1*) was found in the obligate plant-parasitic root-knot nematode *M. incognita*, and encodes a protein that is classified as a member of glycoside hydrolase family 5. The expression of *Mi-xyl1* is localized in the subventral oesophageal gland cells of the nematode. Previous studies have shown that *M. incognita* has the ability to degrade cellulose and pectic polysaccharides in plant cell walls independent of endosymbionts. Including our current data on *Mi-xyl1*, we show that the endogenous enzyme complex in root-knot nematode secretions targets essentially all major cell wall carbohydrates to facilitate a stealthy intercellular migration in the host plant.

Genbank accession number: Mi-xyl1, AF224342.

INTRODUCTION

Plant cells have an extracellular matrix known as the cell wall which is a highly organized composite of diverse polysaccharides, proteins and aromatic compounds (Cosgrove, 1997a). In current models of the cell wall its architecture is described as consisting of several independent but interacting structural networks (Carpita and Gibeaut, 1993). The primary structural network is made of cellulose microfibrils interlocked by heterogeneous polymeric glycans and xylans. This cellulose/hemicellulose scaffold is embedded in a matrix of pectic polysaccharides, which may be reinforced by covalent interactions with aromatic compounds such as the phenylpropanoids. An important function of plant cell walls is to dictate size and shape of the protoplasts during cell growth and differentiation. To further protect the cell against invading organisms, the cell wall may become impregnated with structural proteins and lignins, or covered with local deposits of callose. Alternatively, diffusible fragments cleaved from cell wall polysaccharides may act as elicitors of specific disease resistance responses to invading pathogens and parasites (Boudart *et al.*, 1998).

The proportion of hemicellulose in terrestrial plants may exceed thirty percent of the total dry weight (Cosgrove, 1997a). In type II cell walls of the Gramineae and related monocot families substituted xylans form the main hemicellulose polysaccharide (Carpita, 1996), whereas it accounts for only 5% of the primary cell wall in dicots and non-commelinoid monocots. Endo-1,4-β-xylanases (EC 3.2.1.8) are capable of hydrolyzing the backbone polymers of substituted xylans into smaller pieces of random size (Kulkarni *et al.*, 1999). Presently, most of the endo-1,4-β-xylanases belong to the family 10 and 11 in the numerical classification of glycoside hydrolases CAZy (Carbohydrate-Active Enzymes; Henrissat and Bairoch, 1996). However, the glycoside hydrolase families (GHFs) 5, 8 16, 26, 43, and 62 also include relatively poorly characterized bacterial endo-1,4-β-xylanases. Recently, Larson et al. (Larson *et al.*, 2003) resolved the first crystallographic structure of an endo-1,4-β-xylanase not belonging to either GHF10 or 11.

There are currently a few reports of plant genes encoding endo-1,4-β-xylanase activity, and the functions of these enzymes are associated with a variety of developmental processes (Banik *et al.*, 1996; Bih *et al.*, 1999; Chen and Paull, 2003; Suzuki *et al.*, 2002). In contrast, endo-1,4-β-xylanases have been reported in a wide range of micro-organisms that are typically associated with the degradation of plant cell walls (Beg *et al.*, 2001). For instance, the hydrolysis of xylan polymers of plant cell walls in ruminants depends on enzymes from endosymbiotic bacteria and fungi (Wubah *et al.*, 1993). Similarly, bacteria and fungi are known to contribute to the

endo-1,4-β-xylanase activity in the digestive tracts of xylophagous wood boring insects (Brennan *et al.*, 2004). Some bacterial and fungal pathogens of plants also secrete endo-1,4-β-xylanase suggesting that it may contribute to the infection process (Walton, 1994). Conversely, there are also reports of fungal endo-1,4-β-xylanases that have an adverse effect on the colonization of the plant as they act as elicitors of a cultivar specific local hypersensitive response in tomato, tobacco and pepper plants (Yano *et al.*, 1998).

Here we report the characterization of the first functional animal endo-1,4-β-xylanase gene from the southern root-knot nematode *Meloidogyne incognita.* This enzyme is classified as a member of GHF5. The root-knot nematodes (*Meloidogyne* spp.) are obligate endoparasites of thousands of different plant species, including both monocotyledons and dicotyledons (Williamson and Hussey, 1996). These microscopic worms use a hollow, protrusible stylet to penetrate plant cell walls, release oesophageal gland secretions and withdraw nutrients from the plant cell cytoplasm. The initial penetration of the root epidermis takes place in the elongation and differentiation zone after which the nematode migrates parallel to the vascular bundle through the cortex towards the root tip. In the quiescent centre and the apical initial region the nematode makes a U-turn through the procambium into the vascular cylinder, where it establishes a feeding site (Sijmons *et al.*, 1991). Amazingly, this extensive intercellular migration of the nematode does not inflict any detectable damage to the cells in the cortex and the vascular cylinder. To facilitate this stealthy invasion process the nematode partially dissolves the middle lamella and the primary cell wall between columns of cells. Stylet secretions are believed to have an important role in this parasitic behaviour and are therefore subject of intense study (Davis *et al.*, 2004; Davis *et al.*, 2000). It has recently been discovered that nematodes use cellulases (De Meutter *et al.*, 1998; Gao *et al.*, 2004b; Kikuchi *et al.*, 2004; Rosso *et al.*, 1999; Smant *et al.*, 1998; Uehara *et al.*, 2001; Yan *et al.*, 2001), exo-polygalacturonases (Jaubert *et al.*, 2002), and pectate lyases (De Boer *et al.*, 2002a; Doyle and Lambert, 2002; Popeijus *et al.*, 2000b) to facilitate invasion of the host plant. To date, plant cell wall degradation by animal enzymes has been studied in most detail in nematode-plant interactions covering both covalent and non-covalent interactions between cell wall polymers (Qin *et al.*, 2004). Nematodes were the first animals shown to produce cell wall degrading enzymes independent of endosymbionts. With our current finding of an endogenous functional endo-1,4-β-xylanase gene we show that nematodes have the capacity to degrade the backbones of essentially all major carbohydrate polymers in the (hemi)cellulose and pectin networks in plant cell walls.

Materials and methods

Nematodes

Meloidogyne incognita was propagated on greenhouse cultures of tomato cultivar Moneymaker at 20-25°C. Eggs were harvested approximately 12 weeks after inoculation using a 0.5% (w/v) NaOCl-solution (Hussey and Barker, 1973). Freshly hatched second stage juveniles (J2-s) were collected from eggs on a cotton wool filter, purified using 70% (w/v) sucrose, and stored at –80°C until further processing.

Cloning of *Mi-xyl1*

A partial DNA sequence, named *MD0915*, was generated by random single-pass 5'-end sequencing of approximately 1,000 clones from a cDNA library constructed by Dautova et al. (Dautova *et al.*, 2001). *MD0915* exhibited some similarity to bacterial xylanases and therefore its sequence was extended at the 3'-end by sequencing from the universal SP6 priming site upstream, using the clone from which the initial sequence originated. The open reading frame included in the complete insert of the clone showing the similarity with xylanases lacked a start codon and further amplification reactions were performed to extent the 5'-end of the cDNA sequence. For this purpose, a vector specific forward primer (5'–ATATCTGCAGAATTCGCTAG–3') and an internal reverse primer (5'–GAGCACATTCGGGTGCCA–3') were combined with a plasmid prep of the cDNA library in a polymerase chain reaction (PCR) using *Pwo* DNA-polymerase according to the manufacturer's protocol (Roche Molecular Diagnostics, Indianapolis, IN, U.S.A.).

DNA sequence analysis

Sequencing reactions were performed using the Thermo-Sequenase fluorescent-labelled primer cycle sequencing kit with 7-deaza-dGTP (Amersham-Pharmacia, Buckinghamshire). Basic Local Alignment Search Tool algorithm (BLAST-P2) at the National Centre for Biotechnology Information (NCBI) was used to compare the deduced protein sequence with the non-redundant sequence database of GenBank (03262005). The presence of an N-terminal signal peptide for secretion was predicted at the Signal-P server (Bendtsen *et al.*, 2004).

In situ hybridization microscopy

Linear amplification was used to obtain digoxigenin-11-dUTP labelled sense and antisense cDNA probes according to the manufacturer's protocol (Roche Molecular Diagnostics). The probes were amplified from nucleotide region 625 – 1095 nt of *Mi-*

xyl1 using the primers 5′–TTATGGCACCCGAATGTGC–3′ and 5′–CTGTGATGCACACTTAACG–3′ in two separate reactions (De Boer *et al.*, 1998). Sense and antisense probes (nt 83 to 410) of *Mi-eng-1* were amplified using the primers 5′-TTGGATCCGCTGCTCCTCCATATGG-3′ and 5′-GAGCATTGTGGTCGTGAAAGTC-3′. Freshly hatched nematodes were fixed overnight in 2% (w/v) paraformaldehyde and cut into 3 to 4 pieces. Alkaline phosphatase activity was detected with X-phosphate and 4-nitrobluetetrazoliumchloride (Roche Molecular Diagnostics). The juveniles were examined using differential interference contrast microscopy (Leica, Deerfield).

Heterologous expression in *Escherichia coli*

The sequence (nt 117 – 1045) encoding the mature protein of *Mi-xyl1* was amplified with *Pwo* DNA-polymerase using oligonucleotide primers 5′–GTCGACGAATTCGATAATATAGCAAAAATAAATTCTG–3′ and 5′–CAAGCTTGGATCCCCTAAAATTTATAGAATATTGTTG–3′ according to the manufacturer's protocol. These primers introduced *Eco* RI and *Bam* HI restriction sites (underlined) at the 5′– and 3′–ends of the cDNA sequence respectively to facilitate cloning into the pMAL-p2 expression vector (New England Biolabs, Beverly). Induction of expression and analysis of the recombinant protein were performed according to the manufacturer.

DNA blot analysis

Genomic DNA was isolated from second stage juveniles with alkaline/SDS lysis and phenol/chloroform extraction (Yan *et al.*, 1998) and blotted on a positively charged Nylon Membrane (Roche Molecular Diagnostics). A cDNA probe was synthesized from the 172 – 483 nt region of *Mi-xyl1* using primers 5′–GGTGGTTCTAGTGCTTGG–3′ and 5′– AATTCATCAACAGCAGATTG–3′, and digoxigenin-11-dUTP in a PCR (Roche Molecular Diagnostics). Pre-hybridization (in DIG Easy Hyb at 38°C for 30 min.), hybridization (in DIG Easy Hyb at 38°C overnight), stringency washes (in 0.1% (w/v) SSC and 0.1% (w/v) SDS at 68°C), and immunodetection were performed as recommended by the manufacturer (Roche Molecular Diagnostics).

Enzyme activity assays

A semi-quantitative cup plate assay with birch wood xylan (Sigma, Zwijndrecht, The Netherlands) was used to determine the specific hydrolase activity of the *MBP::Mi-XYL1* fusion protein (Keen *et al.*, 1996). The effect of the pH on enzyme activity was determined by measuring the release of reducing sugars (Bailey *et al.*, 1992). For this

purpose birch wood xylan was suspended at a concentration of 1% (w/v) in 0.05 M buffers based on acetate (pH 3.6-5.4), phosphate (pH 6.0-7.0), Tris/HCl (pH 7.0 – 9.0) or carbonate bicarbonate (pH 10.0-11.0). The buffered substrate (0.9 ml) was incubated with *E. coli* lysates harbouring either the *Mi-xyl1* fusion gene or the empty vector (0.1 ml) for 24 hours at 30°C. Aliquots of 0.2 ml were boiled with 0.3 ml dinitrosalicylic acid to determine the amount of reducing sugars. Each sample was diluted five times prior measurement at 540 nm using a Shimadzu UV-160 recording spectrophotometer (Shimadzu Coop., Kyoto, Japan) with the monomer D-xylose as standard. The absorbance was converted to activities expressed as nanomol reducing sugars released in 1 ml for a time period of 1 sec.

Hydrophobic cluster analysis

HCA plots were made using the program HCA-PLOT V3.0 (Doriane, Le Chesnay, France). In these plots, the amino acid sequence of the proteins is drawn on a duplicated helical net using the standard one-letter code except for P, G, T and S which are represented by stars, diamonds, squares and pointed stars, respectively. Clusters of hydrophobic residues (V, I, L, F, W, M, and Y) are automatically drawn on the bi-dimensional helical representation. Analysis of the plots was performed as described in the results (Lemesle-Varloot *et al.*, 1990). The resolved protein structures of both an endoglucanase (PDB Accession 1EGZ) and a xylanase (PDB Accession 1NOF) from *P. chrysanthemi* were used as coordinating templates for the HCA plots.

Phylogenetic analysis

A BLAST-P2 run was executed on the non-redundant database (26032005) of NCBI using the open reading frame encoded by *Mi-xyl1* as query. All matching sequences from the output with an E-value lower than $1.e^{-5}$ and classified as GHF5 in CAZy were subsequently used to construct phylogenetic trees based on maximum parsimony (ProtPars in Phylip at the Pasteur Institute, Paris) and maximum likelihood (Puzzle) principles. An initial alignment of the matching sequences was made in ClustalW version 1.8 (EBI, Cambridge, UK). The alignment was optimized by removing unambiguously aligned positions and gaps. A character set including a total of 266 amino acid residues for each taxon was used to infer the trees. For maximum parsimony analysis the character set was jumbled 7 times and relative support for the nodes was assessed using 1000 bootstrap replicates with 13 random number seeds. For the maximum likelihood analysis, trees were inferred using the BLOSUM62 model of substitution with 1000 quartet puzzling steps. Three models of rate heterogeneity were used: uniform, two rate (1 variable and 1 invariable) and gamma distributed.

Results

Identification of an endo-1,4-β-xylanase gene from *M. incognita*

Random sequencing of a cDNA library from *M. incognita* pre-parasitic second stage juveniles (J2-s) generated an expressed sequence tag (EST) of 740 bp with similarity to an endo-1,4-β-xylanase of *Aeromonas cavia* (GenBank accession number AAB63573). Further extension of the cDNA sequence at the 5′- and 3′-end revealed a full-length transcript, named *MD0915*, of 1220 nt. The largest open reading frame present in *MD0915* consisted of 331 amino acids. The first 21 N-terminal amino acids are predicted by the SignalP set for eukaryotic sequences to encode a signal peptide for secretion with the most likely cleavage site of the signal peptide between C_{21} and D_{22}. The predicted molecular mass of the mature protein is 34,848 Da with an isoelectric point of 7.69. Two consensus sites for N-linked glycosylation (N_{319} to T_{321} and N_{323} to T_{325}) are present close to the carboxy terminus of the open reading frame.

MD0915 encodes a protein that belongs to GHF5

Similarity searches in non-redundant protein databases showed that the largest open reading frame encoded by *MD0915* is significantly similar to members of two GHFs (Table 1). The first group (i) of matches (Bit Scores ranging from 169 to 146, with E-values ranging from e^{-41} to e^{-21}) are the (putative) endo-β-1,4-xylanases of various bacterial origins, which are all classified as members of the GHF5 in the CAZy database of glycoside hydrolases (Henrissat and Bairoch, 1996). The second group (ii) of matches (Bit Scores ranging from 197 to 152, with E-values ranging from e^{-49} to e^{-6}) represents hypothetical protein sequences from genome sequencing projects on various bacteria that are annotated as possible xylan degrading enzymes from the GHF30. However, these latter sequences have not been included in the CAZy classification. The level of significance of the matches from groups (i) and (ii) are overlapping, which suggested that the protein encoded by *MD0915* could either be member of GHF5 or GHF30. The inferred protein structure of both GHF5 and GHF30 members is an eightfold β-α barrel indicating their ancient ancestry. The variety of enzyme activities in GHF5 (i) includes cellulase (EC 3.2.1.4), endo-1,4-β-xylanase, and a wide range of enzymes with other substrates. The annotation in the second group (ii) of matches derives from one consensus build on five protein sequences (NCBI Conserved Domain Database accession number COG5520). A closer look at the alignment underlying this consensus revealed that three out of five proteins (GenBank accession numbers AAK76863, AAK76864, and CAB13698) are actually classified in CAZy as putative endo-1,4-β-xylanase members of GHF5, and not of GHF30. Consistent with the original definition of COG5520 the remaining two

protein sequences (GenBank accession numbers AAK23733 and AAL23246) underlying the consensus are indeed classified as putative glycosylceramidase members of GHF30. The significance of the similarities between the open reading frame in *MD0915* and GenBank Accessions Numbers AAK76863, AAK76864, and CAB13698 range from e^{-49} to e^{-33} in BLAST-P2 searches, whereas the level of significance of the matches with AAK23733 and AAL23246 are in the range of e^{-9} to e^{-5}. Based on this finding we conclude that the sequences that match *MD0915* belong to two distinct functional groups: the endo-1,4-β-xylanases in GHF5 and, more distantly, the glycosylceramidases in GHF30. Consequently, *MD0915* encodes a protein that is classified as a putative endo-1,4-β-xylanase member of GHF5, and is therefore named *Mi-xyl1*.

Table 1. Summary of BLAST-P2 search using Mi-XYL1 as query in *nr* protein databases of GenBank ([1] the primary GenPept accession number (Acc. no.) of the hits; [2] Bit score of the hits in BLASTP; [3] percentage of identity (ID) and similarity (sim); [4] primary annotation; [5] classification in glycoside hydrolase carbohydrate-active enzymes (CAZy); GHF = glycoside hydrolase family; n.d. = not determined)

Group	Acc. no.[1]	Source organism	Score (bits)[2]	ID/sim[3]	Annotation[4]	CAZy[5]
(i)	AAL16415	*Pectobacterium chrysanthemi*	169	33/52	Endo-xylanase	GHF5
	AAM40172	*Xanthomonas campestris pv. campestris*	157	31/50	Xylanase	GHF5
	AAB63573	*Aeromonas punctata*	152	33/51	Xylanase D	GHF5
	BAB39494	*Ruminococcus albus*	146	32/49	XynC	GHF5
	AAN07016	*Bacillus subtilis*	146	31/51	YnF	GHF5
(ii)	AAK76864	*Clostridium acetobutylicum*	197	37/52	Possible xylan degrading domain (GHF30-like)	n.d.
	ZP_00312042	*C. thermocellum*	152	31/51	COG5520: O-Glycosyl hydrolase (GHF30-like)	n.d.
	AAK76863	*C. acetobutylicum*	152	31/49	Possible xylan degrading domain (GHF30-like)	n.d.

Mi-xyl-1 encodes a functional endo-1,4-β-xylanase

Mi-xyl1 was expressed as a translational fusion-protein with the maltose binding protein *malE* (MBP; ~40 kDa) in the periplasm of *E. coli*. The recombinant protein appeared on Western blot as a single band of approximately 75 kDa (Fig. 1A and B), which is in agreement with the sum of the predicted molecular weight of the mature Mi-XYL1 and the maltose binding protein. Lysates of *E. coli* expressing the MBP::Mi-XYL1 fusion protein tested positive in a cup plate assay with birch wood xylan (Fig. 1C). A potential effect of the pH on the activity of MBP::Mi-XYL1 was assessed by measuring the amount of reducing sugars released from birch wood xylan per unit of time within a pH range from 3.6 to 9. The highest amount of reducing sugars at 30 °C

with an enzyme:substrate ratio of 1:10, was detected at pH 8.0, thus revealing the optimum pH for Mi-XYL1 activity (Fig. 1D). Lysates of bacteria producing the MBP protein from the expression vector pMAL-p2 alone did not produce a halo in the cup plate assays or any detectable level of reducing sugars.

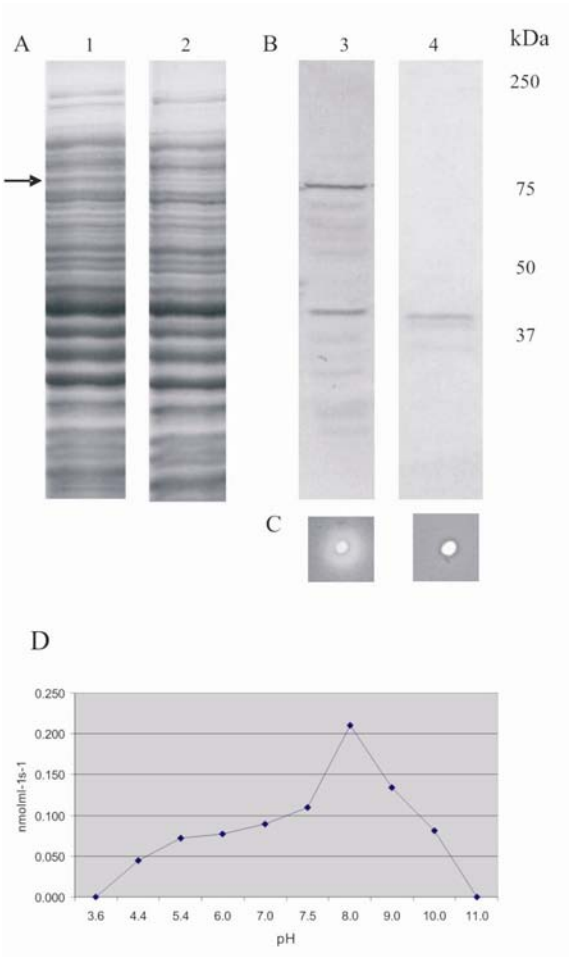


Fig. 1 A. Coomassie Brilliant Blue stained sodium dodecylsulphate/polyacrylamide gel of lysate of *Escherichia coli* cells expressing *Mi-xyl1* from the plasmid *pMAL-p2*. Lane 1 shows a protein extract from induced *E. coli* cells. The heterologous expression in bacteria was induced for 4 h at 37°C with 0.3 mM isopropylthiogalactoside. Lane 2 shows a protein extract from non-induced *E. coli* cells. **B**. Detection of MBP::Mi-XYL1 fusion protein in *E. coli* lysates on Western blot probed with an antibody to the maltose binding protein (MBP). Lane 3 shows the MBP::Mi-XYL1 fusion protein expressed in induced *E. coli* cells. Lane 4 shows the MBP alone in a protein extract from induced *E. coli* cells harbouring the empty *pMAL-p2* vector. Molecular mass markers are indicated on the *right* side of the panel. The arrow indicating the MBP::Mi-XYL1 fusion protein is on the *left* side of the panel. **C**. A semiquantitative cup plate enzyme activity assay of the MBP::Mi-XYL1 fusion protein using birch wood xylan in agar. Lane 3 shows the activity in a protein extract from *E. coli* cells expressing the MBP::Mi-XYL1 fusion protein from the *pMAL-p2* plasmid. Lane 4 shows the activity of the MBP protein alone in protein extracts from *E. coli* cells harbouring *pMAL-p2*. **D**. The pH effect on the activity of MBP::Mi-XYL1 on birch wood xylan. The effect of the pH on the activity of MBP::Mi-XYL1 is measured in the release of reducing sugars in nanomol per millilitre per second.

GHF5 cellulases and endoxylanases differ in β-sheets β1 and β3

All cellulase genes that have been found in *M. incognita* to date also belong to GHF5 (Rosso *et al.*, 1999), and we expected to find sequence similarities between the nematode endoxylanase members and cellulase members of GHF5. Surprisingly, the sequence alignments of Mi-XYL1 and Mi-ENG-1 (GenBank accession number AAD45686), or any other nematode cellulase in GHF5, did not reveal any statistically significant similarities. Hydrophobic cluster analysis was subsequently used to compare the folding patterns of two functional cellulase members of GHF5 (Mi-ENG-1 in *M. incognita* and endoglucanase Z in *Pectobacterium chrysanthemi*, GenBank accession number AAB53151), the best-matching functional endo-1,4-β-xylanase in GHF5 (*P. chrysanthemi*, GenBank accession number AAB53151), and Mi-XYL1 (Fig. 2). The alignment of the hydrophobic cluster analysis (HCA) plots was aided by the crystallographic data of the cellulase and the endoxylanase from *P. chrysanthemi* (PDB Accessions 1EGZ and 1NOF). The alignment showed that the overall eight-fold β-α barrel architecture is conserved in all four protein sequences, which facilitated the identification of two putative active site glutamic acids in the Mi-XYL1 sequence (E_{158} and E_{246}) flanking β-sheets β4 and β7. An asparagine residue directly amino terminal to the putative acid catalyst (E_{158}), which is conserved in most members including the other putative endoxylanases in GHF5, is replaced by serine (S_{157}) in Mi-XYL1. Several independent clones of MD0915 were analysed at this position to eliminate the possibility of a sequencing error. An alignment of all (putative) endoxylanase members of GHF5 showed that a histidine residue at the carboxy flank of the β3-sheet, which is highly conserved in most cellulase members of GHF5, is only present in two endoxylanases of GHF5 (GenBank Accessions Numbers CAE11246 and CAB13698 from *Bacillus* spp.). The conserved histidine in β3-sheet of most GHF5 members is replaced with a tryptophan in GHF5 endoxylanases, and resembles at this position GHF10 endoxylanases. The β3-sheet in Mi-XYL1, however, includes neither a histidine nor a tryptophan. Furthermore, there is almost an absolute conservation in the amino acid motif ($G_{38}F_{39}G_{40}G_{41}$ in Mi-XYL1) that makes up the first β-sheet in the endoxylanase members of the GHF5 (Supplemental Fig. 1), which is absent in the cellulase members in this family.

Mi-xyl1 represents an evolutionary outgroup among GHF5 endoxylanases

A phylogenetic analysis of all endoxylanase members of GHF5 (Fig. 3) revealed three clusters of closely related sequences separate from the protein Mi-XYL1. The first cluster comprises two nearly identical sequences originating from a genome sequencing project on *Clostridium acetobutylicum*. The second cluster with an average

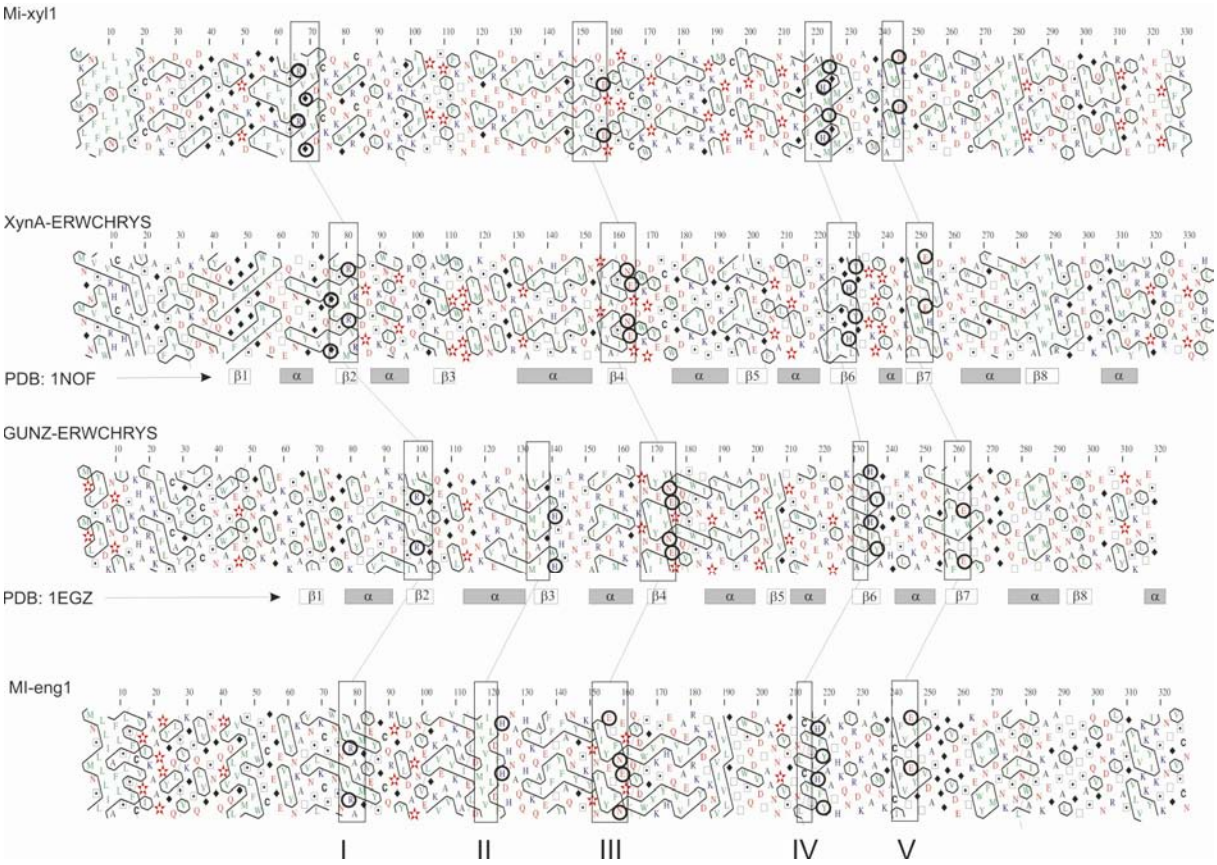


Fig. 2 Alignment of hydrophobic cluster analysis (HCA) plots of Mi-XYL1 from *M. incognita* and the catalytic domains of a cellulase (Mi-ENG-1) from *M. incognita,* an endoxylanase (XynA) and a cellulase (GUNZ) from *P. chrysanthemi* showing the overall eightfold beta-alpha barrel structure of GHF5 family members. The secondary structure elements from the Protein Data Base accessions of XynA (accession 1NOF) and GUNZ (accession 1EGZ) from *P. chrysanthemi* are included to optimize the alignment of the HCA plots. The residues encircled by a bold line in boxes I, II, III, IV, and V are described as hallmark features of GHF5 enzymes (Henrissat and Bairoch, 1996). Box II, including a histidine putatively conserved in GHF5, is missing in the endo-1,4-β-xylanases Mi-XYL1 and XynA, whereas the G-F-G-G tetrapeptide that makes up the first β-sheet in the catalytic domain of XynA is conserved in Mi-XYL1 and in other active endoxylanases of GHF5 (see also Supplemental Fig. 1).

maximum likelihood distance of 0.150 includes the sequences from *Bacillus subtilis* and *B. amyloliquefaciens*, and two sequences from *Aeromonas punctata*. *B. subtilis* and *B. amyloliquefaciens* are soil-dwelling saprophytes, whereas *A. punctata* occurs in surface and waste water. The third cluster with an average maximum likelihood distance of 0.368 includes three sequences from plant-pathogenic bacteria (*P. chrysanthemi* (GenBank accession numbers AAL16415 and AAB53151) and *Xanthomonas campestris* (GenBank accession number AAM40172)). While the average maximum likelihood distance in the complete dataset is 0.83387, the distances of Mi-XYL1 to the other taxa range from 1.14735 to 1.27548.

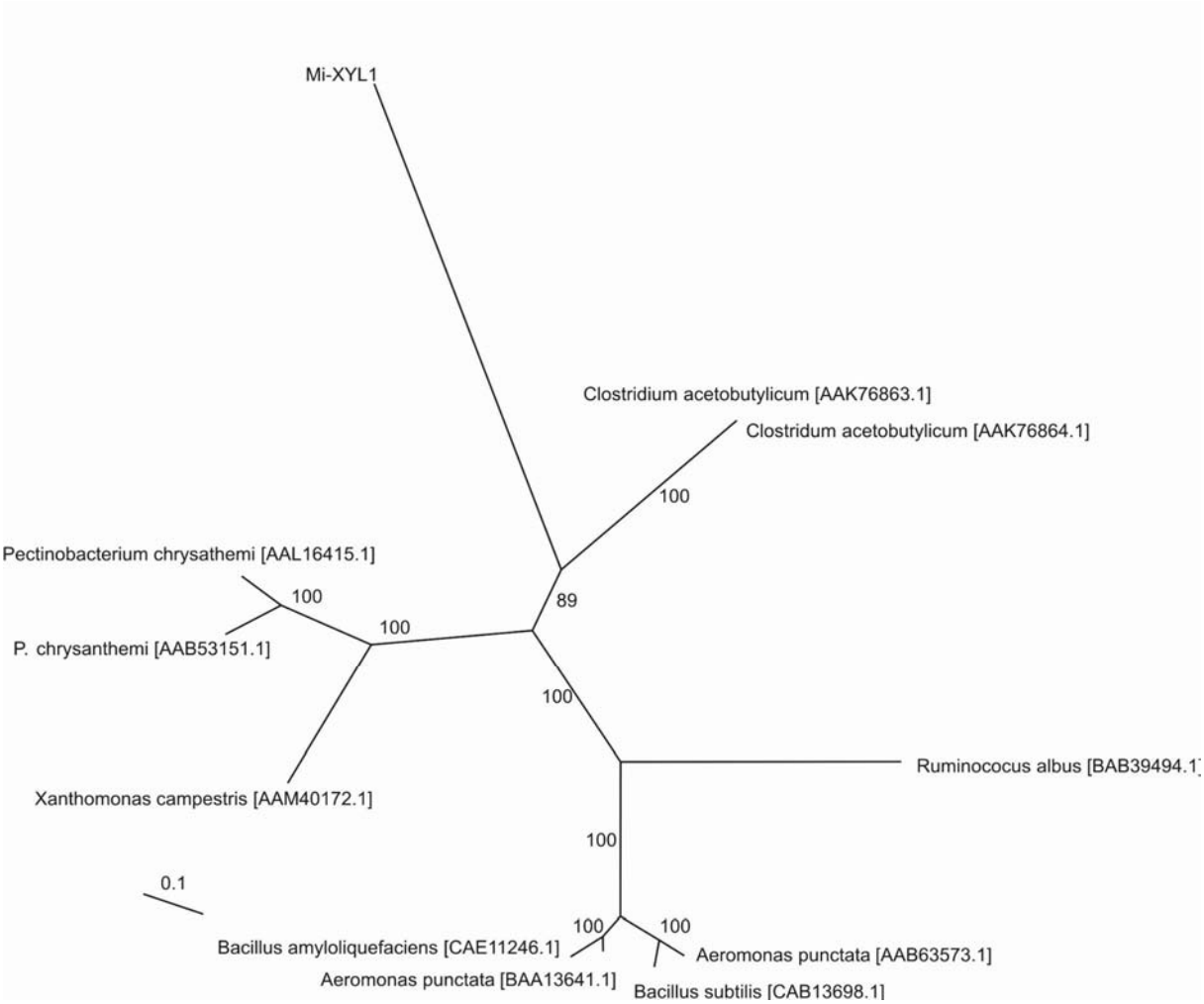


Fig. 3 Phylogenetic analysis of (putative) GHF5 endo-1,4-β-xylanases (with primary accession numbers in GenBank between brackets) using Maximum Likelihood methods. Branch lengths indicate maximum likelihood distance and the relative support for each branching point is indicated at the nodes as a percentage.

Mi-xyl1 is expressed in the subventral oesophageal glands

Mi-xyl1 transcripts were detected in whole mount sections of pre-parasitic J2-s of *M. incognita* using digoxigenin-11-dUTP labelled antisense probes. Transcription of the *Mi-xyl1* gene was confined to the two subventral oesophageal glands (Fig. 4A). The transcripts were detected within the gland cell lobe, and not in the gland extensions. As positive control, *in situ* hybridization was performed with sense and antisense probes designed on the subventral oesophageal gland specific cellulase gene *Mi-eng-1* (Rosso et al. 1999). Labelling of the subventral glands was stronger in the case of *Mi-eng-1* (Fig. 4B), indicating a higher transcriptional level compared to *Mi-xyl1*. In control experiments with sense probes from the *Mi-xyl1* and *Mi-eng-1* gene no hybridization was detectable in nematode sections (Fig. 4C).

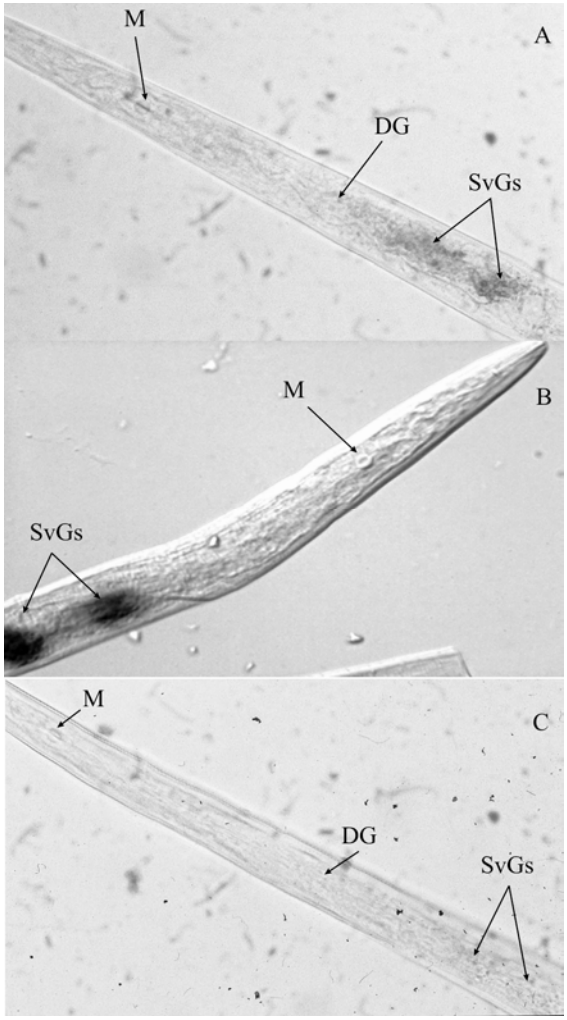


Fig. 4 Whole mount *in situ* localization of *Mi-xyl1* and *Mi-eng-1* transcripts in pre-parasitic second stage juveniles of *Meloidogyne incognita* using digoxigenin-labelled cDNA probes and an alkaline phosphatase-conjugated secondary antibody to digoxigenin. Sections of the nematode incubated with antisense probe designed on *Mi-xyl1* (**A)** and *Mi-eng-1* (**B)** showing a specific alkaline phosphatase staining of the subventral oesophageal gland cells (SvGs) in anterior sections of nematodes. No specific staining was observed in the dorsal oesophageal gland (DG), or in other tissues (M, indicates the metacorpus). **C**. Incubations of nematode sections with sense probe resulted in no specific staining.

Mi-xyl1 homologs are present in *M. javanica*

DNA blot hybridization using a 311 nt digoxigenin-11-dUTP labelled probe of *Mi-xyl1* revealed two distinct bands in genomic DNA of *M. incognita* (Fig. 5, lanes 1 and 2). The oligonucleotide primers used to amplify the probe from nematode cDNA were also used to amplify the same region of the coding sequence from the genomic DNA of *M. incognita*. Restriction analysis with *Bam* HI and *Eco* RI showed that the recognition sites of these enzymes are not present in the region of the genomic sequence of *Mi-xyl1* covered by the probe (data not shown), and that consequently the two bands on southern blot represent two similar but distinct genes. Genomic

DNA of five other nematode species was also analyzed for hybridizing bands using the same *Mi-xyl1* specific probe. A single hybridizing band was observed in genomic DNA of *M. javanica* digested with *Eco* RI (Fig. 5, lane 3), whereas two faint bands where observed in genomic DNA restricted with *Bam* HI (Fig. 5, lane 4). No hybridization was observed in genomic DNA from the plant-parasites *M. hapla*, *Globodera rostochiensis*, *Globodera pallida* and the free-living *Caenorhabditis elegans* (data not shown).

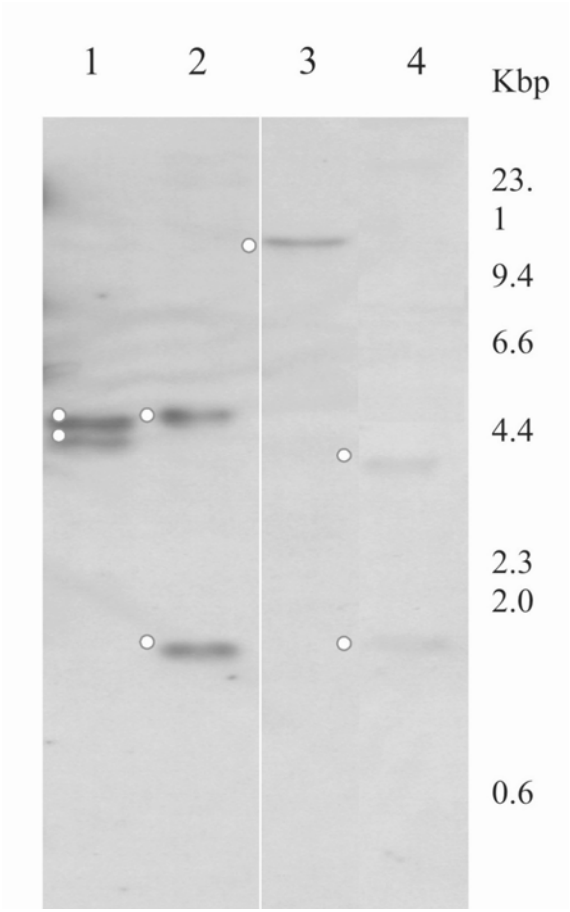


Fig. 5 Southern blot of *M. incognita* and *M. javanica* genomic DNA probed with *Mi-xyl1*. Genomic DNA from *M. incognita* (lanes 1 and 2) and *M. javanica* (lanes 3 and 4) was digested with *Eco* RI and *Bam* HI and hybridized with digoxygenin-dUTP labelled probe from *Mi-xyl1* on a nylon membrane. Two hybridizing bands were observed in *M. incognita* genomic DNA (Lanes 1 and 2). One band was observed in *Eco* RI digested *M. javanica* genomic DNA (lane 3), whereas in *Bam* HI digested genomic DNA two faint bands were detected (lane 4) indicating the presence of at least one homologous gene in *M. javanica* with a possible internal *Bam* HI site. The DNA size markers are indicated on the *right* side of the panel, the white dots indicate positions of the bands.

Discussion

Xylans and glucans in the hemicellulose fraction of terrestrial plants represent one-third of the total carbon resources in biomass on earth. Herbivorous animals have evolved means to breakdown these cell wall polymers either to use them as a carbon source or to gain access to the plant cell cytoplasm. For a long time it was generally believed that animals were not capable of producing cell wall degrading enzymes independent of bacterial or fungal symbionts. This view has been proven wrong as evidence for the production of endogenous cellulases, exo-polygalacturonases, and pectate lyases by several different arthropods, molluscs and nematodes has accumulated. However, no such evidence for the endogenous production of an endo-1,4-β-xylanase in animals has previously been provided.

Ruminants feeding on monocots make use of endo-1,4-β-xylanases produced by endosymbiotic bacteria and fungi (e.g. *Ruminococcus* spp. and *Clostridium* spp.; Kim *et al.*, 1999). There still is a broad consensus that these animals are not capable of producing endo-1,4-β-xylanase themselves. Xylan degrading activity has also been found in the digestive organs of a variety of other animals, e.g. in molluscs (Wang *et al.*, 2003), the rose chafer (Cazemier *et al.*, 1999), beetles (Pitman *et al.*, 2003), cockroaches (Scrivener *et al.*, 1998), termites (Matoub and Rouland, 1995), oribatid mites (Hubert *et al.*, 1999), *Anoplophora chinensis* (Dong *et al.*, 2002), and Antarctic krill (Turkiewicz *et al.*, 2000). Moreover, a cDNA fragment with similarity to endoxylanases was isolated from the digestive tract of the phytophagous beetle *Phaedon cochleariae* (Girard and Jouanin, 1999). The digestive tract of most of these animals is inhabited by cellulolytic and xylolytic microbes, and it is not clear if the origin of cell wall degrading enzymes in these animals is endogenous, from symbiotic micro-organisms, or both.

In this paper we show for the first time evidence for a functional endo-1,4-β-xylanase gene (*Mi-xyl1*) of animal origin. The *Mi-xyl1* gene is specifically transcribed in the subventral oesophageal gland cells of the plant-parasite *M. incognita*. Previous studies have shown the expression of other cell wall degrading enzymes in these single-celled secretory glands including cellulases, exo-polygalacturonases, and pectate lyases. It could be argued that an endosymbiotic eukaryote present in the cytoplasm of the gland cells produces the *Mi-xyl1* transcripts. However, the ultrastructure of the oesophageal gland cells of *M. incognita* was studied extensively with electron microscopy and no indications for a eukaryotic endosymbiont living in the cytoplasm of the gland cells were found (Hussey and Mims, 1990).

Most members of GHF5 have cellulase activity, including the two cellulases (*Mi-eng-1* and *Mi-eng-2*; Rosso *et al.*, 1999) previously identified in *M. incognita*. It is noted,

however, that some cellulases in GHF5 may also have xylanolytic activity (Gao *et al.*, 2004b). The primary amino acid sequence of Mi-XYL1 shows no significant similarity to Mi-ENG-1 and Mi-ENG-2, or in fact to any of the other cellulases in GHF5. This suggests that, although they are all classified as members of the same GHF5, the cellulases *Mi-eng-1* and *Mi-eng-2*, and *Mi-xyl1* have not arisen from a relatively recent gene duplication in *M. incognita*. Nevertheless, hydrophobic cluster analysis of GHF5 members allowed the identification of four stretches of conserved amino acids in the GHF5 endo-1,4-β-xylanases and cellulases, including the two glutamic acid residues putatively involved in the active site of the enzymes. A fifth stretch spanning the β3α3-loop, which was described as a hallmark feature of all GHF5 enzymes, seems not to be conserved in the endo-1,4-β-xylanase members of GHF5. In an attempt to further differentiate the xylanase and cellulase members of GHF5, Larson *et al.* (Larson *et al.*, 2003) suggested that a tryptophan replacing a histidine in this $\beta_3\alpha_3$-loop could define the xylose specificity of both GHF5 and GHF10 endo-1,4-β-xylanases. However, this tryptophan is not conserved in Mi-XYL1 (S_{107}) which makes this hypothesis less likely. The tetra-peptide G-F-G-G that constitutes the first β-sheet in all endo-1,4-β-xylanase members of GHF5, including Mi-XYL1, is lacking in other GHF5 members as well as in the xylan-degrading enzymes of all other GH families, and may therefore be a more signifying feature of the GHF5 endo-1,4-β-xylanases. For its greater evolutionary distance to all other GHF5 endo-1,4-β-xylanase members, Mi-XYL1 may serve as an outgroup to identify the key amino acids for GHF5 type of endo-1,4-β-xylanase activity. Experiments with site-directed mutagenesis of these amino acid residues in GHF5 endo-xylanases are currently being undertaken in our laboratory and may shed light on the basis of substrate specificity of GHF5 enzymes.

The bacterial endo-1,4-β-xylanase members of GHF5 are all multi-domain proteins including a catalytic domain of approximately 280 amino acids and a putative C-terminal carbohydrate binding module of approximately 90 amino acids. In contrast, Mi-XYL1 consists only of a catalytic domain of around 280 amino acids, and in that sense it reflects more the topology of most GHF11 members. Nematode cellulases often occur in two nearly identical types in which presence or absence of the ancillary cellulose binding domain is the only significant difference (Smant *et al.*, 1998). The second band that appeared on Southern blot of *M. incognita* genomic DNA probed with *Mi-xyl1* may represent a copy of the gene that includes a carbohydrate binding module. Preliminary results from an ongoing EST project on the plant-parasite *Meloidogyne chitwoodi* show that this root-knot nematode species has at least one similar endo-1,4-β-xylanase including a putative xylan binding domain (Chapter 3 of this thesis).

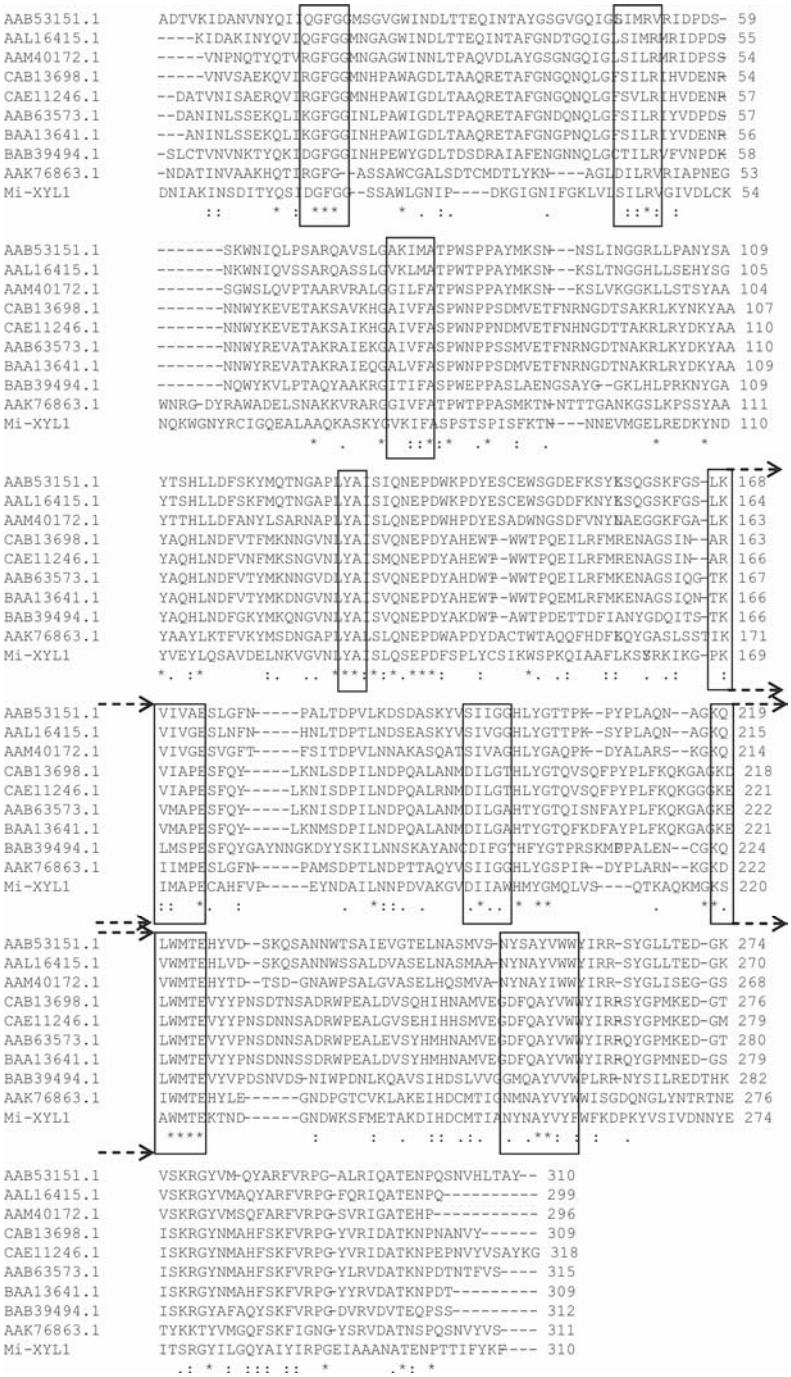
M. incognita is a polyphagous parasite of approximately seven hundred different plant species, both monocotyledons and dicotyledons. The polymer constitution of

the cell wall in each of these host plants varies extensively, thus requiring great flexibility in the substrate specificity of the enzymes in nematode stylet secretions. Southern blot analysis revealed the presence of a homologue of *Mi-xyl1* in the related root-knot nematode species *M. javanica*, which also thrives on both monocots and dicots, whereas in *M. hapla* – a specialist of dicots – the probe failed to produce a hybridizing band. Database searches revealed two matching ESTs out of 7,587 from *M. javanica* with approximately 96 % identity in 500 nucleotides (GenBank accession numbers CF350477 and CF350376), one match out of 12,218 ESTs from *M. chitwoodi* (GenBank accession number CB931318), and one short matching element with 86% identity in 50 nucleotides in two ESTs out of 5,018 from *M. arenaria* (GenBank accession numbers CF357210 and CF357155). No matches with *Mi-xyl1* were found in the 24,452 ESTs that are currently available from *M. hapla* (data not shown). Putative endo-1,4-β-xylanases also seem to be absent in other nematode species that parasitize dicotyledons only. Although these datasets are limited, we hypothesize for further studies that the production of endo-1,4-β-xylanases in plant-parasitic nematodes is correlated with parasitism on monocots.

Endo-1,4-β-xylanases are thought to have an essential role in the virulence of bacterial and fungal pathogens of monocots (Walton, 1994). Several studies involving targeted gene disruption, however, have not shown that endo-1,4-β-xylanases are required for the pathogenicity of these plant-pathogens (Gomez Gomez *et al.*, 2002; Keen *et al.*, 1996; Ray *et al.*, 2000; Wu *et al.*, 1997). Attempts to develop similar gene disruption methods for plant-parasitic nematodes in order to assess the role of cell wall degrading enzymes in parasitism have been fruitless. However, recent developments to knock-down genes by RNA interference in other plant-parasitic nematodes are currently being adopted to *M. incognita* (Fanelli *et al.*, 2005; Rosso *et al.*, 2005; Urwin *et al.*, 2002), and these methods may provide information about the influences of endo-1,4-β-xylanases and other cell wall degrading enzymes on the host plant range of plant parasitic nematodes.

ACKNOWLEDGEMENTS

We thank Dr. John T. Jones from the Scottish Crop Research Institute for critically reading the manuscript. The work was supported by the European Union grants NONEMA (QLK5-1999-1501) and DREAM (QLRT-1999-1462), and by the VENI-VIDI-VICI program of the Dutch Science Foundation (NWO).



Supplemental Fig. 1 Alignment of protein sequences retrieved from the non-redundant database of NCBI that have significant similarity (Score (Bits) > 50 and E-value < $1e^{-5}$) to Mi-XYL1 in BLASTP and are classified as members of GHF5. The boxed residues correspond to the eight β-sheets in Xylanase A of *P. chrysanthemi* (GenBank accession numbers AAL16415/ABB53151 and PDB accession 1NOF). The [:] indicate positives and the [*] indicate identities in the alignment generated with the BLOSUM62 matrix.

6

Summarizing discussion and concluding remarks

The products of parasitism genes in nematodes must be secreted to reach their targets at the nematode-plant interface. These nematode secretory proteins are therefore recognised to play an important role in the nematode-plant interaction and as a result have been subject of intense study for years. In this thesis, the results on the identification of members of the parasitome in plant-parasitic nematodes, with a focus on the root-knot nematode species *Meloidogyne chitwoodi* and *M. incognita*, are described and discussed. In **Chapter 1** a general introduction to the subject is given, which provides background on the main issues of this thesis. In **Chapter 2** the first overview of the secretome and candidate parasitome members from *M. chitwoodi* are reported, which was at the start of the project a molecularly unexplored root-knot nematode species. As a starting point, all publicly available expressed sequence tags (ESTs) were used, i.e. 12,218, from three life stages of *M. chitwoodi*. A pipeline of bio-informatics algorithms was applied to screen this dataset for secretory proteins, which resulted in the identification of 398 secretome members. Two-third of these 398 sequences has no significant similarity to other known proteins in the *nr* database.

In this study 'parasitome' is defined as 'the products of parasitism genes secreted by a range of secretory organs in the nematode, including the oesophageal, amphidial, and rectal glands, the intestine, and the hypodermis'. To point out candidate parasitism gene products within the *M. chitwoodi* secretome, the site of expression of the most abundant secretome members was localised using the *in situ* hybridisation procedure. We found that at least eight out of the twenty most abundantly represented members of the secretome are specifically expressed in the single-celled oesophageal glands. One tag hybridised in the dorsal oesophageal gland, seven in the two subventral oesophageal glands, two in the intestine, and one tag hybridised to the tail tip in the proximity of the phasmids. Four sequences showed similarity to putative parasitism genes from other nematode species, whereas seven represented pioneering sequences. In case of three oesophageal gland specific parasitism genes, the predicted subcellular localization in host plant cells, following cleavage of the leader peptide for protein secretion, was nuclear.

The results of this study, and those described in chapters 3, 4 and 5, show that the generation of ESTs indeed forms an invaluable resource for gene discovery, in this case (candidate) parasitism genes (McCarter *et al.*, 2000). In particular, our findings obtained with *in situ* hybridisation microscopy make clear that with a combination of ESTs from high quality cDNA libraries and bioinformatics it is not necessary to use pre-selection methods to enrich libraries for transcripts from the gland region of a nematode. These pre-selections were based on tissue specific hybridizations (e.g. suppressive subtractive hybridization; Huang *et al.*, 2004a), micro-aspiration of mRNA from specific regions in the parasite (Gao *et al.*, 2003; Huang *et al.*, 2003; Wang

et al., 2001), and signal sequence trapping (Wang *et al.*, 2001). Compared to these methods, EST generation from whole nematode libraries followed by bioinformatics analyses is relatively simple and more time efficient. Similarly, differential display methods, as were used by us and others in the past, seem less efficient than EST analysis. The consequence of using whole nematode cDNA libraries, without pre-selection, is that our strategy is most effective with larger numbers of ESTs. Five years ago the costs involved in the generation of ESTs were still a rate limiting factor. However, the generation of ESTs from (plant-)parasitic nematodes is an ongoing process which currently operates at a high pace. Fortunately, the ESTs that are being produced are made publicly available by efforts such as the Parasitic Nematode Sequencing Project (for an overview, see http://www.nematode.net/Species.Summaries/index.php (Wylie *et al.*, 2004)). Our strategy may be further improved by using signal peptide prediction software which is trained on datasets built from nematode secretory proteins (e.g. *Caenorhabditis elegans*), since there is evidence that the composition of the signal peptides may vary between eukaryotic phyla (Pearson *et al.*, 2005).

Our strategy to identify the products of parasitism genes forms a basis for further functional analysis of parasite proteins involved in the nematode-host interaction. Our efforts were focused on the genes involved in the very early stages of parasitism by infective second stage juveniles, including host invasion and suppression of early defence responses in the host. In order to extend this basis, it would be interesting to enrich the EST dataset analysed in Chapter 2 with ESTs from later parasitic juvenile stages of *M. chitwoodi*. The same analyses can then be applied to these ESTs in order to point out candidate parasitism genes which are for example involved in giant cell formation and maintenance. Alternatively, or in addition, candidate parasitism genes can be pointed out by *in silico* comparison of the ESTs (and their abundances) between different developmental stages. In this way, EST analysis also forms an alternative for the technically more complicated differential display methods such as cDNA-AFLP and micro array analyses (De Boer *et al.*, 2002b; Qin *et al.*, 2000).

In **Chapter 3** and **Chapter 4**, the focus is on the repertoire of cell wall modifying proteins (CWMPs) which were identified from *M. chitwoodi*. These proteins form a subset of the parasitome of plant parasitic nematodes and play an important role during the intercellular migration of root-knot nematodes through host plant roots. Our study resulted in the identification of the most elaborate repertoire of CWMPs found in a single plant-parasitic nematode species so far. It includes β-1,4-endoglucanases (cellulases), a β-1,4-endoxylanase, pectate lyases, polygalacturonases, a cellulose binding protein (all described in **Chapter 3**) and expansin-like proteins (**Chapter 4**).

Four β-1,4-endoglucanases were identified from *M. chitwoodi* and named *Mc-eng-1* to *Mc-eng-4*. Full-length sequences were obtained for *Mc-eng-1* to *Mc-eng-3*. The encoded cellulases were found to belong to glycoside hydrolase family 5 (GHF5) and were all predicted to contain an N-terminal signal peptide for secretion. *In situ* hybridisation localised the site of expression of *Mc-eng-3* in the subventral oesophageal secretory glands of *M. chitwoodi* pre-parasitic J2-s. Attempts to obtain the full length sequence of *Mc-eng-4* failed, but sequence similarity search results with the partial sequence revealed that this sequence encodes a cellulase that consists of a catalytic domain and a carbohydrate binding module (CBM). These results indicate that *M. chitwoodi* produces cellulases with two different domain architectures, either a single catalytic domain (Mc-ENG-1, Mc-ENG-2, and Mc-ENG-3) or a catalytic domain linked to a CBM (Mc-ENG-4). Our phylogenetic analysis of root-knot nematode cellulases, including Mc-ENG-1, Mc-ENG-2, and Mc-ENG-3, showed that there are probably no cellulases that are unique for one (group of) nematode species, but that there are several lines of cellulases common to root-knot nematodes.

The site of *Mc-cbp-1* (cellulose binding protein) expression was localised to the subventral oesophageal secretory glands of *M. chitwoodi* pre-parasitic J2-s. The Mc-CBP-1 protein consists of a bacterial family 2 CBM, which can mediate the binding to the cellulose substrate and/or may facilitate the non-hydrolytic disruption of cellulose fibres (Din *et al.*, 1991; Levy *et al.*, 2002; Shoseyov *et al.*, 2006; Tomme *et al.*, 1998). The CBM of Mc-CBP-1 is C-terminally linked to a domain of 80 amino acids that shares some similarity with Fn3-like domains. The latter domains are involved in protein-protein interactions, and in addition, the 80 amino acid N-terminal domain is rich in charged amino acids which are typically found on the hydrophilic surfaces of proteins to interact with other proteins. We hypothesise that Mc-CBP-1 is involved in the facilitation of cellulose hydrolysis, either as a docking station for CBM-lacking cellulases to bind to cellulose or by modification of the cellulose surface, or both. On the other hand, recent work by Gaulin *et al.* (2006) suggests that cellulose binding domains from the oomycete *Phytophthora parasitica* may also act as an elicitor of defence responses in plants (Gaulin *et al.*, 2006). It is not yet clear if this is a general phenomenon of cellulose binding domains released by plant-pathogens and to what extent nematode cellulose binding proteins induce, either directly or indirectly through cell wall modifications, defence responses.

Since *Meloidogyne* species migrate intercellularly through the roots of a host plant, they go through the middle lamella which is rich in pectic polysaccharides. The complete degradation of pectins into monomers requires the combined action of several types of pectinolytic enzymes. Root-knot nematodes do not feed on carbohydrates released from pectin degradation and thus do not require the

complete degradation of plant cell wall pectins. In contrast, root-knot nematodes actually make use of a controlled and local degradation of the pectins in the middle lamella only and will benefit most likely from hydrolases and lyases (depolymerases), which only cleave the backbone of pectin (Tamaru and Doi, 2001), to weaken the intercellular bondings between cells. Both the pectate lyases and polygalacturonases identified from *M. chitwoodi* in this study belong to this group of pectin degrading enzymes.

Four pectate lyases, named *Mc-pel-1* to *Mc-pel-4*, were identified from *M. chitwoodi* invasive J2-s. In case of *Mc-pel-1* and *Mc-pel-2*, full-length sequences were obtained, whereas *Mc-pel-3* and *Mc-pel-4* were represented by partial cDNA sequences in the *M. chitwoodi* ppJ2 cDNA library. Both *Mc-pel-1* and *Mc-pel-2* have an N-terminal signal peptide for secretion. *In situ* hybridisation performed on *M. chitwoodi* invasive J2-s revealed a specific expression confined to the subventral oesophageal secretory glands in case of *Mc-pel-1*. Based on similarity search results, the (putative) pectate lyases are all considered to be member of pectate lyase family III, which also comprises the pectate lyases from other plant-parasitic nematodes. Vanholme *et al.* (2007) showed with RNA interference (RNAi) knock-down experiments that a moderate but significant reduction in transcripts of a pectate lyase leads to strongly reduced infectivity of cyst nematodes (Vanholme *et al.*, 2007). Cyst nematodes are not particularly focused on a stealthy invasion through the middle-lamella of host cells. Instead, they use a brute force approach during plant invasion and leave behind a trail of destructed cells. Stealthily invading root-knot nematodes are predicted to have stronger requirement for pectin degrading enzymes than cyst nematodes, however, there is no experimental data from RNAi knock-down experiments available at the moment to support this prediction.

A family of at least four pectate lyases was found in *M. chitwoodi*, while others reported similar findings in earlier studies on *M. incognita*. The pectate substrate of pectate lyase is essentially built from repeating units of galacturonic acid with varying degrees of methylation. The actual composition of the oligogalacturonate units changes during the life time of a plant, and varies between different plant species. The overall topology of pectate lyases is the same, the core consists of β-strands forming a right-handed parallel β-helix, which suggests a conserved mode of action (D'Ovidio *et al.*, 2004). In an attempt to explain multiple pectate lyase gene families in plant-pathogens, Herron *et al.* (2000) proposed that plant-pathogens with broad host ranges have multiple isozymes of pectate lyases with similar catalytic properties, but that recognize differently composed and decorated oligogalacturonate units (Herron *et al.*, 2000). This proposal now seems to find support in the repertoire of pectate lyases of the polyphagous root-knot nematodes.

Root-knot nematodes have evidently explored other venues, besides the production of pectate lyases, to achieve degradation of the pectin component in host cell walls. In total, two polygalacturonases, named *Mc-pg-1* and *Mc-pg-2*, were identified from invasive J2-s of *M. chitwoodi*. The full-length sequence of *Mc-pg-1* was obtained and with *in situ* hybridisation performed on *M. chitwoodi* invasive J2-s, a specific expression confined to the subventral oesophageal secretory glands was revealed. Based on similarities with other exo-acting polygalaturonases, we believe that Mc-PG-1 codes for an *exo*-polygalacturonase. The polygalacturonase MI-PG-1, which was cloned from *M. incognita* is also classified as an exo-acting enzyme (Jaubert *et al.*, 2002). Exo- and endo-polygalacturonases have different impacts on the cell wall integrity of plants. Endo-polygalacturonases are known to play a role in plant defence responses, e.g. through the release of elicitor-active oligogalacturonides (Cervone *et al.*, 1989; D'Ovidio *et al.*, 2004; Favaron *et al.*, 1988). Biotrophic pathogens, like *Meloidogyne* species, must avoid inducing host defence responses and therefore it seems likely that they will profit from a cell wall degrading enzyme repertoire that does not release oligogalacturonides. Experimental biochemical evidence for the classification of polygalacturonases from *M. chitwoodi* and *M. incognita* as exo-acting enzymes is still lacking. It will be interesting to investigate the enzymatic properties of these nematode polygalacturonases in order to establish to which class of polygalacturonases they belong.

Most of the research on the role of cell wall degrading enzymes in plant-pathogen interactions is done on pectin degrading enzymes in pathogens of dicotyledons. Pectins, however, represent only a minor fraction of the cell wall components in monocots, which mainly include hemicelluloses in their cell wall matrices (Carpita, 1996). A major part of the hemicellulose fraction of monocotyledonous cell walls consists of substituted xylan polymers. It is therefore expected that xylan-degrading enzymes have an important role in promoting virulence of pathogens of monocots; possibly equivalent to the role of pectinolytic enzymes for pathogens of dicots (Beliën *et al.*, 2006 and references herein).

We identified two novel xylan-degrading enzymes in two root-knot nematode species that are virulent pathogens of both dicotyledons and monocotyledons, i.e. *M. chitwoodi* and *M. incognita* (described in **Chapter 3** and **5**). These xylan-degrading enzymes are β-1,4-endoxylanases, which are capable of hydrolysing substituted xylan polymers into fragments of random size. The β-1,4-endoxylanase from *M. incognita*, named *Mi-xyl1*, is the first functional β-1,4-endoxylanase of animal origin. Both *Mi-xyl1* and the β-1,4-endoxylanase from *M. chitwoodi*, *Mc-xyl-1*, were found to be expressed in the subventral oesophageal gland cells of the nematode and the encoded proteins were predicted to have an N-terminal signal peptide for secretion. Based on similarity search results we consider MI-XYL1 and Mc-XYL-1 enzymes that

belong to GHF5. These two β-1,4-endoxylanases share 39% identity in their catalytic domain. Further comparison between the two xylanases showed that Mc-XYL-1 has an ancillary stretch of 76 amino acids at its C-terminus which is, based on sequence similarity, a putative xylan binding module (XBM).

So far, results of Southern blot analysis and EST database screenings did not show evidence of endoxylanases in nematode species that have specialized on dicots (*M. hapla*, *G. rostochiensis* and *G. pallida*), in spite of thousands of ESTs in public sequence databases. Based on our findings, we hypothesise that the production of β-1,4-endoxylanase in plant parasitic nematodes is correlated with parasitism on monocots. RNA interference studies as mentioned above may provide information about the requirement of β-1,4-endoxylanase for infectivity on monocots. In addition, it would be interesting to investigate if the specialist nematodes of monocots indeed have a bias towards xylan degradation in their repertoire of cell wall modifying proteins. Unfortunately, nematode specialists on graminaceous monocots, such as the cereal cyst nematode *Heterodera avenea*, have not been included in the scope of molecular nematology so far. Therefore, there is currently no sequence information available of such nematode species.

Chapter 4 deals with the expansin-like proteins identified in *Meloidogyne* spp., with the emphasis on those from *M. chitwoodi*. Expansins form a diverse protein superfamily in plants and play a role in various biological processes in which re-arrangement of plant cell wall polysaccharides is involved (Cosgrove, 2000a). Plant expansins lack hydrolytic activity and are proposed to weaken non-covalent interactions between cellulose and hemicellulose polymers (McQueen-Mason and Cosgrove, 1994). They seem to act synergistically with cellulases by making the plant cell wall polysaccharides more accessible to enzymatic attack (Cosgrove, 2000a). A similar synergistic action of expansins and cell wall degrading enzymes secreted by plant-pathogens may facilitate their invasion of the host.

The distribution of expansins was believed to be restricted to land plants (Cosgrove, 2000a), but recently, a small number of expansin-like sequences have been identified from other organisms (Darley *et al.*, 2003; Laine *et al.*, 2000; Saloheimo *et al.*, 2002). In addition, a functional β-expansin, Gr-EXPB1, was found to be secreted by the plant-parasitic nematode *Globodera rostochiensis* (Kudla *et al.*, 2005; Qin *et al.*, 2004). This latter sequence was used to query assembled ESTs from *M. chitwoodi* and found in total four expansin-like sequences, named *Mc-EXP1* to *Mc-EXP4*. These *M. chitwoodi* sequences harbour most of the signature motifs of α- and β-expansins, but since we were not able to produce active recombinant protein to test expansin activity of the proteins on plant tissues, the sequences are designated as expansin-like proteins.

Expansin-like *Mc-EXP1* is represented by 23 ESTs in the EST dataset from *M. chitwoodi* ppJ2-s and belongs to the top 20 of most abundantly expressed members of the secretome of this developmental stage of *M. chitwoodi* (Chapter 2 of this thesis). Both Mc-EXP1 and Mc-EXP2 are multi-domain proteins. They consist of an N-terminal signal peptide for secretion, followed by either a CBM (Mc-EXP1) or a Lysin-motif domain (LysM; Mc-EXP2) both linked to a C-terminal expansin-like domain. Compared to the domain structure of plant expansins, the expansin-like domain and the polysaccharide-binding domain (CBM) are in reverse orientation. Thus, the domain structure of Mc-EXP1 is similar to that of Gr-EXPB1. The sequence identity between the C-termini of Mc-EXP1 and Mc-EXP2 and plant expansins ranged from 30% to 35%. Notably, for both expansin-like proteins highest sequence similarity is with α-expansins from plants. Conclusive evidence on the type of activity of Mc-EXP1 and Mc-EXP2, i.e. do they resemble more the α- or β-expansins, requires the production of active recombinant protein and subsequent cell wall extension assays on the different types of cell walls. Unfortunately, we have not been able to achieve heterologous expression of Mc-EXPs in plants, which would have allowed us to study the type of expansin activity of the proteins.

Probes designed on *Mc-EXP1* localised the gene transcription in the subventral oesophageal secretory glands of ppJ2-s of *M. chitwoodi*. The same localisation was found for the Mc-EXP1 and Mc-EXP2 proteins, which was investigated by immunofluorescence microscopy with specific antisera to Mc-EXP1 and Mc-EXP2. A strong fluorescence, with a granular pattern, was observed in the subventral oesophageal gland extensions and ampullae and to a lesser extent in the gland lobes. The presence of a signal peptide for secretion and the localisation of the transcripts and protein in the subventral oesophageal secretory glands of *M. chitwoodi* ppJ2-s strongly suggest that expansin-like Mc-EXP1 is secreted by the nematode. The current data on the developmental expression and production of expansin-like Mc-EXP2 are less straightforward. *Mc-EXP2* was identified in a library made from nematode eggs and could not be amplified from the cDNA library made from *M. chitwoodi* ppJ2-s. In addition, no *in situ* hybridisation signal was observed in *M. chitwoodi* ppJ2-s with an antisense probe spanning the N-terminal putative LysM domain. On the other hand, in immunofluorescence microscopy experiments, the Mc-EXP2 protein was localised specifically in the subventral oesophageal glands of *M. chitwoodi* ppJ2-s. Therefore, a developmental expression study needs to be performed on *Mc-EXP2* in order to find out if the results obtained at the protein level are either confirmed or contradicted by results obtained at the transcriptional level.

The other two expansin-like proteins from *M. chitwoodi*, Mc-EXP3 and Mc-EXP4, only contain an expansin-like domain with an N-terminal signal peptide for secretion. *Mc-EXP4* is represented by 55 ESTs from the ppJ2 stage of *M. chitwoodi* and

herewith it belongs to the top 10 of most abundantly represented transcripts in the *M. chitwoodi* EST dataset (see Chapter 2 of this thesis). Despite its high abundance, no hybridisation signal was detected for *Mc-EXP4* by *in situ* hybridisation in ppJ2-s of *M. chitwoodi*. Specific staining of the subventral oesophageal secretory glands in this developmental stage was observed in case of *Mc-EXP3*.

In addition to the expansin-like genes in *M. chitwoodi*, evidence was also found for the presence of expansin-like genes in five other root-knot nematode species, a root-lesion nematode species, and plant pathogenic oomycete and fungal species. It remains to be shown whether these organisms produce functional expansins, but our findings make us point at two things. Firstly, it appears that expansins do not occur in only a small number of organisms outside the plant kingdom, but instead, are widespread and likely to be involved in many plant-pathogen interactions and plant-microbe interactions in general. Secondly, the proposed and currently adopted nomenclature of the expansin superfamily might need revision. For, the designations 'expansin' (expansin A and expansin B) and 'expansin-like' (expansin-like family A and B) in this nomenclature are kept exclusively for plant proteins. Proteins from other organisms that share structural similarity with both domains of plant expansins are grouped in a separate 'catch-all' category and designated 'expansin-related', solely based on the fact that they do not originate from plants. When the expansin-like proteins identified in this study are indeed functional expansins, there would be no biological rationale behind the 'catch-all' group of expansin-related proteins.

A series of conserved cysteines (C) and the HFD motif around amino acid position 110 are used as the key signatures of the plant expansin family (Cosgrove, 2000b). A series of conserved cysteines is also present along the backbone of the root-knot nematode expansin-like domains. In case of the HFD motif, only the H and D residues are conserved in the expansin-like proteins of root-knot nematodes. A similar conservation was found in the β-expansin Gr-EXPB1 from *G. rostochiensis*, for which cell wall expansion activity on type II primary cell walls was found (Kudla *et al.*, 2005; Qin *et al.*, 2004). These findings suggest that the functional significance of the HFD motif is obscure and can be clarified by the biochemical characterization of root-knot nematode expansin-like proteins, e.g. by site directed mutagenesis followed by activity assays.

It seems likely that the CWMP repertoire of *M. chitwoodi* mirrors its wide host plant range since the latter represents a broad diversity of cell wall polysaccharides. The CWMPs identified from *M. chitwoodi* enable the migrating nematode to cleave the backbone of all major types of plant cell wall polysaccharides and to modify their interactions. The subventral oesophageal gland specific expression and presence of predicted secretion signal peptides suggest that the CWMPs are secreted from the

nematode. Hereby, this enzyme complex facilitates the intercellular migration of *M. chitwoodi* in the host plant. Post-transcriptional gene silencing by soaking nematode juveniles in double-stranded RNA (RNA interference) was successfully applied to assess the importance of β-1,4-endoglucanases and pectate lyases for infection of plant roots by cyst nematodes (Chen *et al.*, 2005; Vanholme *et al.*, 2007). Similar RNA interference experiments targeting CWMP encoding genes from *M. chitwoodi*, followed by infection tests on monocots and dicots, may reveal the influence of the individual CWMPs on the host range of *M. chitwoodi*.

On the other hand, our findings raise the questions how it is possible that a nematode can secrete such a variety of CWMPs without i) inflicting any detectable damage to the cells along the migratory track and ii) being noticed by the host plant? An explanation for the fact that plant cells remain intact might be that the cell wall modifying enzymes that are found to be produced by the nematode only cleave the backbone of plant cell wall polysaccharides. As mentioned above, root-knot nematodes do not feed on carbohydrates released from plant cell wall degradation and thus do not require the complete degradation of plant cell wall polysaccharides. In contrast, root-knot nematodes only seem to weaken the intercellular bondings between cells by local secretion of the CWMPs through the stylet into the middle lamella resulting in controlled degradation. This controlled degradation might also be a way to prevent the production of elicitor-active oligogalacturonides involved in plant defence responses. Alternatively or in addition, the nematode might constantly repress the host defence mechanisms during intercellular migration through the roots.

In relation to these questions it is interesting to study whether the nematode can 'sense', either inside or outside the plant root system or both, with which type of plant (cell wall) it is dealing. Following the perception of cell wall type signals, the nematode could adapt its repertoire of CWMPs to it. This would enable the nematode to avoid the production of CWMPs that do not act on the available substrate, thus saving valuable energy and reducing the chance of being detected by the host. Potato root exudates are for example known to induce the secretion of proteins by cyst nematodes (Smant *et al.*, 1997). One way to test whether host plant root exudates already influence the production of CWMPs at the transcriptional level would be to treat eggs or hatched nematodes with different root exudates followed by real-time quantitative PCR analysis. One could for example hatch nematodes in water and in root exudates obtained from a monocot and a dicot and compare the transcriptional levels of a certain CWMP.

7

Samenvatting

Nematoden, ook wel aaltjes genoemd, vormen een diverse groep rondwormen. Ze worden in praktisch ieder grondmonster gevonden, vaak met een hoge soortendiversiteit en in groten getale. Om een idee te geven: het aantal nematoden in de grond ligt in het algmeen ergens tussen één en tien miljoen per vierkante meter (!). Nematoden hebben zeer uiteenlopende voedingsstrategieën: sommige soorten voeden zich met bacteriën of schimmels terwijl anderen omnivoor zijn ('alles-eters') of op planten(wortels) of dieren parasiteren. Naast hun uiteenlopende voedingsstrategieën komen nematoden ook in zeer uiteenlopende habitats voor en worden ze in de meest extreme milieus op aarde gevonden. Voorbeelden hiervan zijn hete en koude woestijnen, die gekarakteriseerd worden door extreme temperaturen, beperkte beschikbaarheid van water, hoge bodem-pH en zoutgehaltes. Nematoden zijn bijvoorbeeld op verschillende plaatsen op Antarctica, een koude woestijn, in de bodem gevonden. Plantparasitaire nematoden zijn, zelfs op een diepte van 11 tot 12 meter, gevonden in de Chihuahuan woestijn in New Mexico, Verenigde Staten.

Dit proefschrift beschrijft het onderzoek dat gedaan is aan nematodensoorten die behoren tot de groep die op het wortelstelsel van planten parasiteert. De nematoden waar het in dit proefschrift in hoofdzaak om gaat zijn de wortelknobbelaaltjes, in het Latijn *Meloidogyne* genaamd, in het bijzonder de soorten *M. chitwoodi* en *M. incognita*. Deze wortelknobbelaaltjes hebben een zeer brede waardplantenreeks, ze kunnen bijvoorbeeld bijna alle gecultiveerde plantensoorten infecteren en leveren daardoor in de landbouw een enorme economische schade op. Deze nematoden zijn zogenaamde 'sedentaire endoparasieten', hetgeen wil zeggen dat ze een deel van hun levenscyclus op een vaste plek in de wortels van de plant doorbrengen. Volwassen vrouwtjes leggen honderden eitjes in een gelatine-achtige matrix op het oppervlak van de plantenwortel. Na een eerste vervelling in het ei komen mobiele juvenielen vrij in de bodem die vervolgens een plantenwortel binnendringen. In de wortel migreert de juveniele nematode tussen de plantencellen door (intercellulair) naar een vaste plaats in de wortel. Op deze plek induceren de nematoden de vorming van voedingscellen, de zogenaamde 'reuzencellen'. Deze reuzencellen zijn plantencellen die de nematoden van een constante toevoer van voedingsstoffen voorzien. Gedurende het voeden ondergaat de nematode een aantal vervellingen en ontwikkelt zich zo tot het volwassen stadium. De cellen rondom de zich ontwikkelende nematode vermenigvuldigen zich en worden groter waardoor er ter plaatse een knobbel ontstaat, vandaar de naam 'wortelknobbelaaltjes'. De gehele levenscyclus kan, afhankelijk van allerlei omgevingfactoren, één tot twee maanden in beslag nemen.

Gedurende de migratiefase door de plantenwortel en bij de inductie en instandhouding van de voedingscellen spelen door de nematode uitgescheiden (gesecreteerde) eiwitten een grote rol. Een belangrijk deel van de secretie-eiwitten

wordt aangemaakt in drie grote kliercellen. Deze eiwitten worden uitgescheiden via een mondstylet, een holle en uitsteekbare speer, die ook gebruikt wordt om de plantencelwand te penetreren en om voedingsstoffen uit de reuzencellen op te nemen.

Het complete repertoire aan secretie-eiwitten van een organisme wordt in de moleculaire biologie aangeduid met 'secretoom'. Identificatie en screening van secretie-eiwitten kan leiden naar de groep eiwitten die de moleculaire basis vormen van de interactie tussen plantparasitaire nematode en waardplant. Uit eerdere studies is bijvoorbeeld naar voren gekomen dat nematoden celwandmodificerende eiwitten uitscheiden. Deze eiwitten, voornamelijk enzymen, werken in op de plantencelwand en vergemakkelijken het binnendringen van en de intercellulaire migratie van nematoden door de plantenwortels. Het feit dat nematoden deze enzymen zelf produceren is heel bijzonder. Ze zijn de eerste dieren waarvoor is aangetoond dat ze hier zelf toe in staat zijn, zonder de hulp van micro-organismen als bacteriën en schimmels.

Zoals eerder aangegeven hebben de wortelknobbelaaltjes *M. chitwoodi* en *M. incognita* een zeer brede waardplantenreeks. De plantensoorten in deze reeks verschillen, onder andere, in de opbouw van hun celwanden. Omdat de celwanden de barrières vormen bij het binnendringen van de plant, betekent dit dat de nematode goed uitgerust moet zijn om deze verschillende barrières te overwinnen. Onze hypothese luidde dan ook dat *M. chitwoodi* en *M. incognita* over een breed repertoire aan celwandmodificerende eiwitten moeten beschikken om succesvol de duizenden verschillende plantensoorten binnen te kunnen dringen.

Aangezien er aan het begin van dit project op moleculair gebied weinig tot niets bekend was over *M. chitwoodi*, was het eerste onderzoek gericht op de identificatie van het secretoom van deze nematodensoort. Vervolgens is, om onder andere antwoord te krijgen op voornoemde hypothese, de aandacht voornamelijk gericht op de identificatie en karakterisatie van celwandmodificerende eiwitten van beide nematodensoorten. In dit proefschrift worden de resultaten van dit onderzoek beschreven.

In **Hoofdstuk 1** wordt een algemene introductie gegeven aangaande het onderwerp van dit proefschrift. Het doel van de studie beschreven in **Hoofdstuk 2** was het identificeren van secretie-eiwitten (het secretoom) van *M. chitwoodi*, gevolgd door het aanwijzen van kandidaat parasitismegenen. De parasitismegenen coderen voor eiwitten die uitgescheiden worden door de nematode naar het raakvlak tussen plant en nematode. Tezamen maken deze producten het mogelijk om succesvol op een waardplant te leven. Als startpunt van deze studie werden zogenaamde *expressed sequence tags* (ESTs) gebruikt. Dit zijn korte DNA-sequenties die op een relatief eenvoudige, goedkope en snelle manier geproduceerd kunnen worden en een ruw

beeld geven van de genen die actief zijn (tot expressie komen) in bijvoorbeeld een bepaald ontwikkelingsstadium van een organisme. In deze studie zijn 12.218 ESTs gebruikt uit drie ontwikkelingsstadia van *M. chitwoodi*: eieren, pre-parasitaire juvenielen en volwassen vrouwtjes. Secretie-eiwitten worden gekenmerkt door een korte sequentie die voorafgaat aan het eigenlijke eiwit, het zogenaamde signaalpeptide. Dit signaalpeptide zorgt ervoor dat het eiwit in de secretieroute van een cel terechtkomt en uiteindelijk buiten de cel wordt uitgescheiden. Door sequenties dus te analyseren op de aan- of afwezigheid van zo'n signaalpeptide kunnen de secretie-eiwitten als zodanig herkend worden. Deze analyse is uitgevoerd op de ESTs van *M. chitwoodi* met behulp van verschillende computeralgoritmes (bioinformatica). Dit resulteerde in de identificatie van 398 potentieel gesecreteerde eiwitten. Opvallend is dat tweederde van deze 398 sequenties geen significante overeenkomst heeft met reeds eerder gevonden sequenties en dus zogenaamde pioniersequenties zijn. De 398 sequenties tezamen vormen het eerste overzicht van het secretoom van *M. chitwoodi*.

Door te achterhalen op welke plek in de nematode een gen coderend voor een secretie-eiwit tot expressie komt krijgt men een indicatie of het eiwit een belangrijke rol speelt in de moleculaire interactie tussen nematode en waardplant. Dit laatste is bijvoorbeeld het geval wanneer een eiwit in één van de drie grote kliercellen geproduceerd wordt, aangezien bekend is dat eiwitten uit deze klieren uitgescheiden worden naar het raakvlak tussen nematode en waardplant. Een techniek die gebruikt wordt om de plaats van expressie zichtbaar te maken is *in situ* hybridisatie. Met behulp van deze techniek is van elf genen coderend voor secretie-eiwitten de locatie van expressie achterhaald. Voor acht eiwitten is gevonden dat ze tot expressie komen in de kliercellen van de nematode. Met deze resultaten hebben we dus ook, naast het eerste overzicht van het secretoom van *M. chitwoodi*, de eerste kandidaat parasitismegenen van deze nematodensoort geïdentificeerd. Deze vondsten vormen een solide basis voor verder (functioneel) onderzoek.

Hoofdstuk 3 en **Hoofdstuk 4** richten zich specifiek op de identificatie van celwandmodificerende eiwitten uit het juveniele stadium van *M. chitwoodi*, wederom uitgaande van ESTs. Deze celwandmodificerende eiwitten spelen zoals gezegd een belangrijke rol bij het binnendringen van de plantenwortel en tijdens de migratie tussen de plantencellen door. De ESTs, 4.067 in totaal, zijn gescreend voor sequenties die coderen voor celwandmodificerende eiwitten, hetgeen resulteerde in de identificatie van het meest brede repertoire aan dergelijke eiwitten dat in één enkele nematodensoort tot op heden is gevonden. Het bevat β-1,4-endoglucanases (cellulases), een β-1,4-endoxylanase, pectaatlyases, polygalacturonases, een cellulosebindend eiwit (alle beschreven in **Hoofdstuk 3**) en expansine-achtige eiwitten (**Hoofdstuk 4**). Deze eiwitten tezamen werken in op alle hoofdcomponenten

van de plantencelwand, zijnde cellulose, hemi-cellulose en pectine. Voor de verschillende celwandmodificerende eiwitten zijn signaalpeptides voor secretie gevonden en aangetoond kon worden dat deze eiwitten in de kliercellen van de nematode tot expressie komen. Deze resultaten vormen een sterke aanwijzing dat de eiwitten daadwerkelijk door de nematode uitgescheiden worden.

Cellulose vormt de ruggegraat van de celwand, bij elkaar gehouden door hemi-celluloses. Dit cellulose/hemi-cellulosenetwerk ligt ingebed in een matrix van pectine, hetgeen als een soort lijm tussen de verschillende cellen fungeert. De cellulases en β-1,4-endoxylanase zijn enzymen die cellulose en hemi-cellulosecomponenten kunnen afbreken. Van de cellulases zijn vier unieke sequenties in onze dataset gevonden. Alle behoren tot dezelfde cellulasefamilie als de cellulases die in andere plantparasitaire nematodensoorten gevonden zijn.

Het endo-1,4-β-xylanase dat in *M. chitwoodi* is geïdentificeerd, is het tweede xylanase dat in een plantparasitaire nematode is gevonden. Het eerste is een endo-1,4-β-xylanase van *M. incognita*, dat reeds eerder was gevonden en beschreven wordt in **Hoofdstuk 5**. Vergeleken met het endo-1,4-β-xylanase uit *M. incognita* heeft dat uit *M. chitwoodi* een extra domein dat mogelijk betrokken is bij de binding van het enzym aan zijn substraat.

Het cellulosebindend eiwit bestaat uit twee domeinen waarvan er één de binding aan cellulose kan verzorgen en/of de afbraak van het cellulosenetwerk kan vergemakkelijken. Dit deel zit gekoppeld aan een domein dat enige gelijkenis vertoont met domeinen die betrokken zijn bij eiwit-eiwit interacties. Aan de hand van de gelijkenis met andere sequenties verwachten wij dus dat het cellulosebindend eiwit uit *M. chitwoodi* direct of indirect betrokken zal zijn bij de afbraak van het cellulosenetwerk.

De pectaatlyases en polygalacturonases die in onze dataset van *M. chitwoodi* gevonden zijn, behoren tot de groep enzymen die de hoofdketen van pectine verbreken. In totaal werden sequenties coderend voor vier verschillende pectaatlyases en twee polygalacturonases gevonden. Op basis van gelijkenis met andere sequenties worden alle vier de pectaatlyases gerekend tot familie III van deze groep enzymen, de familie waartoe ook de pectaatlyases van andere plantparasitaire nematoden behoren. Eveneens gebaseerd op overeenkomst met andere sequenties nemen we aan dat het polygalacturonase uit *M. chitwoodi* een *exo*-polygalacturonase is. Exo- en endo-polygalacturonases hebben een verschillende uitwerking op de structuur van de celwand en het is dan ook interessant om met behulp van biochemische analyses te achterhalen tot welk type het polygalacturonase uit *M. chitwoodi* behoort. Uit onze resultaten kunnen we dus concluderen dat het repertoire aan celwandmodificerende eiwitten van *M. chitwoodi* inderdaad de grote variëteit

aan celwandcomponenten weerspiegelt die de nematode in zijn brede waardplantenreeks tegenkomt.

In **Hoofdstuk 4** wordt specifiek ingegaan op één groep van celwandmodificerende eiwitten, namelijk expansines of expansine-achtige eiwitten, die in deze studie in *M. chitwoodi* gevonden zijn. Expansines vormen een diverse eiwitfamilie in planten, waar ze een rol spelen bij uiteenlopende biologische processen waarbij herschikking van de plantencelwand van belang is, bijvoorbeeld bij de groei van plantencellen en het rijpen van fruit. Expansines bezitten zelf geen hydrolytische activiteit, maar lijken samen te werken met andere celwandafbrekende enzymen door de celwandcomponenten toegankelijker te maken voor enzymatische aanvallen. Expansines leken uniek voor het plantenrijk, maar recent werd door ons een functioneel expansine, Gr-EXPB1, gevonden dat gesecreteerd werd door het aardappelcystenaaltje *Globodera rostochiensis*. Naar aanleiding van deze vondst vroegen wij ons af of de secretie van expansines, in combinatie met celwandmodificerende enzymen, uniek is voor cystenaaltjes of dat vergelijkbare combinaties van celwandmodificerende eiwitten ook gesecreteerd worden door andere plantparasitaire nematoden, zoals het wortelknobbelaaltje *M. chitwoodi*, of zelfs door andere, ongerelateerde plantpathogenen.

De Gr-EXPB1 sequentie is gebruikt om ESTs van *M. chitwoodi* te screenen voor expansinesequenties, hetgeen resulteerde in de identificatie van vier verschillende sequenties (Mc-EXP1 t/m Mc-EXP4). Aangezien deze sequenties uit *M. chitwoodi* de meeste signatuurmotieven van expansines bevatten, maar we niet in staat waren om de activiteit van de eiwitten te testen op plantenweefsels, worden de sequenties aangeduid met 'expansine-achtige' eiwitten. Mc-EXP1 en Mc-EXP2 zijn multidomeineiwitten: ze bevatten een signaalpeptide voor secretie, gevolgd door een cellulosebindend domein (Mc-EXP1) of een lysinemotiefdomein (Mc-EXP2), beiden gekoppeld aan een expansine-achtig domein. Resultaten van *in situ* hybridisatie experimenten lieten zien dat Mc-EXP1 in de kliercellen van de nematode tot expressie komt. Eenzelfde localisatie werd gevonden voor het Mc-EXP1 en Mc-EXP2 eiwit, hetgeen onderzocht werd met immunofluorescentiemicroscopie met specifieke antilichamen tegen Mc-EXP1 en Mc-EXP2. De aanwezigheid van een signaalpeptide voor secretie en de localisatie van de transcripten en eiwitten in de kliercellen van *M. chitwoodi* vormen een sterke indicatie dat de expansine-achtige eiwitten Mc-EXP1 en Mc-EXP2 worden uitgescheiden door de nematode. De andere twee expansine-achtige eiwitten, Mc-EXP3 en Mc-EXP4, bestaan alleen uit een expansine-achtig domein en een signaalpeptide voor secretie.

Naast expansine-achtige sequenties in *M. chitwoodi* hebben we dergelijke sequenties ook in vijf andere wortelknobbelaaltjessoorten, een wortellesieaaltje, een plantpathogene oomyceet en diverse schimmelsoorten gevonden. Het moet nog

onderzocht worden, net als voor *M. chitwoodi*, of deze organismen *functionele* expansines produceren, maar onze resultaten wijzen in ieder geval op twee punten. Ten eerste lijkt het zo te zijn dat expansines niet slechts in een kleine groep organismen buiten het plantenrijk voorkomen, maar juist wijdverspreid zijn en mogelijk een rol spelen in veel interacties tussen plant en pathogeen en in interacties tussen plant en micro-organisme in het algemeen. Ten tweede behoeft de huidige naamgeving van de expansinefamilie waarschijnlijk revisie aangezien de indicaties 'expansine' en 'expansine-achtig' op dit moment exclusief voorbehouden zijn aan eiwitten uit planten. Wanneer blijkt dat vele organismen buiten het plantenrijk functionele expansines produceren is deze naamgeving niet langer houdbaar.

In **Hoofdstuk 5** beschrijven we de identificatie van het eerste functionele endo-1,4-β-xylanase geproduceerd door een dier. Dit endo-1,4-β-xylanase is gevonden in *M. incognita* en hebben we *Mi-xyl1* genoemd. De expressie van *Mi-xyl1* is gelocaliseerd in de kliercellen van de nematode en het enzym is door ons geclassificeerd als lid van glycosidehydrolasefamilie 5, dezelfde familie waartoe ook de eerder genoemde cellulases behoren. Endo-1,4-β-xylanases kunnen celwandpolymeren afbreken die een belangrijk deel uitmaken van de hemicellulosefractie van plantencelwanden, vooral in zogeheten monocotylen waartoe bijvoorbeeld de grassen en granen behoren. Daarnaast komt uit sequentie-analyses die in ons onderzoek uitgevoerd zijn het beeld naar voren dat nematodensoorten die niet op monocotylen parasiteren ook geen xylanases produceren. Het lijkt er dus op dat xylanases alleen nodig zijn voor nematoden die op monocotyle planten leven.

Uit eerdere studies is reeds gebleken dat *M. incognita* in staat is om cellulose- en pectinebestanddelen van plantencelwanden af te breken. Met de identificatie van xylanase Mi-XYL1 uit *M. incognita* wordt, zoals mede bleek uit de studies aan *M. chitwoodi*, onze hypothese bevestigd dat wortelknobbelaaltjes met een brede waardplantenreeks ook een breed repertoire aan celwandmodificerende eiwitten produceren dat actief is op alle hoofdbestanddelen van plantencelwanden.

References

Adam, M.A.M., Phillips, M.S. and Blok, V.C. (2007) Molecular diagnostic key for identification of single juveniles of seven common and economically important species of root-knot nematode (*Meloidogyne* spp.). *Plant Pathology* **56**, 190-197.

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

Andreeva, A., Howorth, D., Brenner, S.E., Hubbard, T.J.P., Chothia, C. and Murzin, A.G. (2004) SCOP database in 2004: Refinements integrate structure and sequence family data. *Nucleic Acids Research* **32**, D226-D229.

Bailey, M.J., Biely, P. and Poutanen, K. (1992) Interlaboratory testing of methods for assay of xylanase activity. *Journal of Biotechnology* **23**, 257-270.

Baldauf, S.L., Roger, A.J., Wenk-Siefert, I. and Doolittle, W.F. (2000) A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science* **290**, 972-977.

Banik, M., Garrett, T.P.J. and Fincher, G.B. (1996) Molecular cloning of cDNAs encoding (1-4)-beta-xylan endohydrolases from the aleurone layer of germinated barley (*Hordeum vulgare*). *Plant Molecular Biology* **31**, 1163-1172.

Barras, F., van Gijsegem, F. and Chatterjee, A.K. (1994) Extracellular Enzymes and Pathogenesis of Soft-Rot Erwinia. *Annual Review of Phytopathology* **32**, 201-234.

Bateman, A. and Bycroft, M. (2000) The structure of a LysM domain from *E. coli* membrane-bound lytic murein transglycosylase D (MltD). *Journal of Molecular Biology* **299**, 1113-1119.

Beg, Q.K., Kapoor, M., Mahajan, L. and Hoondal, G.S. (2001) Microbial xylanases and their industrial applications: A review. *Applied Microbiology and Biotechnology* **56**, 326-338.

Beliën, T., Van Campenhout, S., Robben, J. and Volckaert, G. (2006) Microbial endoxylanases: Effective weapons to breach the plant cell-wall barrier or, rather, triggers of plant defense systems? *Molecular Plant-Microbe Interactions* **19**, 1072-1081.

Bendtsen, J.D., Nielsen, H., Von Heijne, G. and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology* **340**, 783-795.

Bih, F.Y., Wu, S.S.H., Ratnayake, C., Walling, L.L., Nothnagel, E.A. and Huang, A.H.C. (1999) The predominant protein on the surface of maize pollen is an endoxylanase synthesized by a tapetum mRNA with a long 5' leader. *Journal of Biological Chemistry* **274**, 22884-22894.

Bird, A.F. (1983) Changes in the dimensions of the esophageal glands in root-knot nematodes during the onset of parasitism. *International Journal for Parasitology* **13**, 343-348.

Bird, D.M. (1996) Manipulation of host gene expression by root-knot nematodes. *Journal of Parasitology* **82**, 881-888.

Bird, D.M. and Koltai, H. (2000) Plant Parasitic Nematodes: Habitats, Hormones, and Horizontally-Acquired Genes. *Journal of Plant Growth Regulation* **19**, 183-194.

Boraston, A.B., Bolam, D.N., Gilbert, H.J. and Davies, G.J. (2004) Carbohydrate-binding modules: Fine-tuning polysaccharide recognition. *Biochemical Journal* **382**, 769-781.

Boucher, G. and Lambshead, P.J.D. (1995) Ecological biodiversity of marine nematodes in samples from temperate, tropical, and deep-sea regions. *Conservation Biology* **9**, 1594-1604.

Boudart, G., Lafitte, C., Barthe, J.P., Frasez, D. and EsquerreTugaye, M.T. (1998) Differential elicitation of defense responses by pectic fragments in bean seedlings. *Planta* **206**, 86-94.

Bourne, Y. and Henrissat, B. (2001) Glycoside hydrolases and glycosyltransferases: families and functional modules. *Current Opinion in Structural Biology* **11**, 593-600.

Bowles, D.J. (1990) Defense-related proteins in higher plants. *Annual Review of Biochemistry* **59**, 873-907.

Brennan, Y., Callen, W.N., Christoffersen, L., Healey, S., Keller, M., Li, K., Palackal, N., Wells, S., Hazlewood, G.P., Mathur, E.J., Short, J.M., Robertson, D.E., Steer, B.A., Hernandez, M., Tamayo, G., Sittenfeld, A., Dupree, P. and Goubet, F. (2004) Unusual microbial xylanases from insect guts. *Applied and Environmental Microbiology* **70**, 3609-3617.

Campbell, I.D. and Spitzfaden, C. (1994) Building proteins with fibronectin type III modules. *Structure* **2**, 333-337.

Carpita, N. and McCann, M. (2000) The Cell Wall. In *Biochemistry and molecular biology of plants* (Buchanan, B.B., Gruissem, W. and Jones, R.L., eds.), pp. 52-108: American Society of Plant Biologists.

Carpita, N.C. (1996) Structure and biogenesis of the cell walls of grasses. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**, 445-476.

Carpita, N.C. and Gibeaut, D.M. (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant journal* **3**, 1-30.

Cazemier, A.E., Verdoes, J.C., Van Ooyen, A.J.J. and Op den Camp, H.J.M. (1999) Molecular and biochemical characterization of two xylanase-encoding genes from *Cellulomonas pachnodae*. *Applied and Environmental Microbiology* **65**, 4099-4107.

Cervone, F., Hahn, M.G., De Lorenzo, G., Darvill, A. and Albersheim, P. (1989) Host-Pathogen Interactions. XXXIII. A plant protein converts a fungal pathogenesis factor into an elicitor of plant defense responses. *Plant Physiology* **90**, 542-548.

Chen, N.J. and Paull, R.E. (2003) Endoxylanase expressed during papaya fruit ripening: Purification, cloning and characterization. *Functional Plant Biology* **30**, 433-441.

Chen, Q., Rehman, S., Smant, G. and Jones, J.T. (2005) Functional analysis of pathogenicity proteins of the potato cyst nematode *Globodera rostochiensis* using RNAi. *Molecular Plant-Microbe Interactions* **18**, 621-625.

Chiou, S.H., Huang, K.F., Chow, L.P., Tsugita, A. and Wu, S.H. (1996) Isolation of a venom factor devoid of proteolytic activity from Taiwan habu (*Trimeresurus mucrosquamatus*): N-terminal sequence homology and no functional similarity to factors IX/X-binding proteins and botrocetin. *Journal of Protein Chemistry* **15**, 667-674.

Cosgrove, D.J. (1997a) Assembly and enlargement of the primary cell wall in plants. *Annual Review of Cell and Developmental Biology* **13**, 171-201.

Cosgrove, D.J. (1997b) Relaxation in a high-stress environment: The molecular bases of extensible cell walls and cell enlargement. *Plant Cell* **9**, 1031-1041.

Cosgrove, D.J. (1999) Enzymes and other agents that enhance cell wall extensibility. *Annual review of plant physiology and plant molecular biology* **50**, 391-417.

Cosgrove, D.J. (2000a) Loosening of plant cell walls by expansins. *Nature* **407**, 321-326.

Cosgrove, D.J. (2000b) New genes and new biological roles for expansins. *Current Opinion in Plant Biology* **3**, 73-78.

Cosgrove, D.J., Bedinger, P. and Durachko, D.M. (1997) Group I allergens of grass pollen as cell wall-loosening agents. *Proceedings of the National Academy of Sciences U.S.A* **94**, 6559-6564.

Darley, C.P., Li, Y., Schaap, P. and McQueen-Mason, S.J. (2003) Expression of a family of expansin-like proteins during the development of *Dictyostelium discoideum*. *FEBS Letters* **546**, 416-418.

Dautova, M., Rosso, M.N., Abad, P., Gommers, F.J., Bakker, J. and Smant, G. (2001) Single pass cDNA sequencing - a powerful tool to analyse gene expression in preparasitic juveniles of the southern root-knot nematode *Meloidogyne incognita*. *Nematology* **3**, 129-139.

Davis, E.L., Hussey, R.S. and Baum, T.J. (2004) Getting to the roots of parasitism by nematodes. *Trends in Parasitology* **20**, 134-141.

Davis, E.L., Hussey, R.S., Baum, T.J., Bakker, J., Schots, A., Rosso, M.-N. and Abad, P. (2000) Nematode parasitism genes. *Annual Review of Phytopathology* **38**, 365-396.

De Boer, J.M., McDermott, J.P., Davis, E.L., Hussey, R.S., Popeijus, H., Smant, G. and Baum, T.J. (2002a) Cloning of a putative pectate lyase gene expressed in the subventral esophageal glands of *Heterodera glycines*. *Journal of Nematology* **34**, 9-11.

De Boer, J.M., McDermott, J.P., Wang, X., Maier, T., Qui, F., Hussey, R.S., Davis, E.L. and Baum, T.J. (2002b) The use of DNA microarrays for the developmental expression analysis of cDNAs from the oesophageal gland cell region of *Heterodera glycines*. *Molecular Plant Pathology* **3**, 261-270.

De Boer, J.M., Overmars, H.A., Pomp, R., Davis, E.L., Zilverentant, J.F., Goverse, A., Smant, G., Stokkermans, J.P.W.G., Hussey, R.S., Gommers, F.J., Bakker, J. and Schots, A. (1996a) Production and characterization of monoclonal antibodies to antigens from second stage juveniles of the potato cyst nematode, *Globodera rostochiensis*. *Fundamental and Applied Nematology* **19**, 545-554.

De Boer, J.M., Smant, G., Goverse, A., Davis, E.L., Overmars, H.A., Pomp, H., Van Gent-Pelzer, M., Zilverentant, J.F., Stokkermans, J.P.W.G., Hussey, R.S., Gommers, F.J., Bakker, J. and Schots, A. (1996b) Secretory granule proteins from the subventral esophageal glands of the potato cyst nematode identified by monoclonal antibodies to a protein fraction from second-stage juveniles. *Molecular Plant-Microbe Interactions* **9**, 39-46.

De Boer, J.M., Yan, Y., Smant, G., Davis, E.L. and Baum, T.J. (1998) In-situ hybridization to messenger RNA in *Heterodera glycines*. *Journal of Nematology* **30**, 309-312.

De Boer, J.M., Yan, Y., Wang, X., Smant, G., Hussey, R.S., Davis, E.L. and Baum, T.J. (1999) Developmental expression of secretory ß-1,4-endoglucanases in the subventral esophageal glands of *Heterodera glycines*. *Molecular Plant-Microbe Interactions* **12**, 663-669.

De Meutter, J., Tytgat, T., Van der Schueren, E., Smant, G., Schots, A., Coomans, A., Van Montagu, M. and Gheysen, G. (1998) Cloning of two endoglucanase genes from *Heterodera schachtii*. *Mededelingen Faculteit Landbouwkundige en Toegepaste Biologische Wetenschappen, Universiteit Gent* **63**, 619-623.

Den Nijs, L.J.M.F., Brinkman, H. and Van der Sommen, A.T.C. (2004) A Dutch contribution to knowledge on phytosanitary risk and host status of various crops for *Meloidogyne chitwoodi* Golden *et al.*, 1980 and *M. fallax* Karssen, 1996: an overview. *Nematology* **6**, 303-312.

Din, N., Gilkes, N.R., Tekant, B., Miller Jr, R.C., Warren, R.A.J. and Kilburn, D.G. (1991) Non-hydrolytic disruption of cellulose fibres by the binding domain of a bacterial cellulase. *Bio/Technology* **9**, 1096-1099.

Ding, X., Shields, J., Allen, R. and Hussey, R.S. (1998) A secretory cellulose-binding protein cDNA cloned from the root-knot nematode (*Meloidogyne incognita*). *Molecular Plant-Microbe Interactions* **11**, 952-959.

Ding, X., Shields, J., Allen, R. and Hussey, R.S. (2000) Molecular cloning and characterisation of a venom allergen AG5-like cDNA from *Meloidogyne incognita*. *International Journal for Parasitology* **30**, 77-81.

Dong, Y.M., Yin, Y.P., Cao, Y.Q. and He, Z.B. (2002) Purification and properties of xylanase from *Anoplophora chinensis* (Forster). *Acta Entomologica Sinica* **45**, 165-169.

D'Ovidio, R., Mattei, B., Roberti, S. and Bellincampi, D. (2004) Polygalacturonases, polygalacturonase-inhibiting proteins and pectic oligomers in plant-pathogen interactions. *Biochimica et Biophysica Acta* **1696**, 237-244.

Doyle, E.A. and Lambert, K.N. (2002) Cloning and characterization of an esophageal-gland-specific pectate lyase from the root-knot nematode *Meloidogyne javanica*. *Molecular Plant-Microbe Interactions* **15**, 549-556.

Endo, B.Y. (1987) Ultrastructure of esophageal gland secretory granules in juveniles of *Heterodera glycines*. *Journal of Nematology* **19**, 469-483.

Esquerré-Tugayé, M.T., Boudart, G. and Dumas, B. (2000) Cell wall degrading enzymes, inhibitory proteins, and oligosaccharides participate in the molecular dialogue between plants and pathogens. *Plant Physiology and Biochemistry* **38**, 157-163.

Fanelli, E., Di Vito, M., Jones, J.T. and De Giorgi, C. (2005) Analysis of chitin synthase function in a plant parasitic nematode, *Meloidogyne artiella*, using RNAi. *Gene* **349**, 87-95.

Favaron, F., Alghisi, P., Marciano, P. and Magro, P. (1988) Polygalacturonase isoenzymes and oxalic acid produced by *Sclerotinia sclerotiorum* in soybean hypocotyls as elicitors of glyceollin. *Physiological and Molecular Plant Pathology* **33**, 385-395.

Freckman, D.W. and Virginia, R.A. (1989) Plant-feeding nematodes in deep-rooting desert ecosystems. *Ecology* **70**, 1665-1678.

Fry, S.C. (1989) The structure and functions of xyloglucan. *Journal of Experimental Botany* **40**, 1-11.

Gao, B., Allen, R., Davis, E.L., Baum, T.J. and Hussey, R.S. (2004a) Molecular characterisation and developmental expression of a cellulose-binding protein gene in the soybean cyst nematode *Heterodera glycines*. *International Journal for Parasitology* **34**, 1377-1383.

Gao, B.L., Allen, R., Davis, E.L., Baum, T.J. and Hussey, R.S. (2004b) Developmental expression and biochemical properties of a beta-1,4-endoglucanase family in the soybean cyst nematode, *Heterodera glycines*. *Molecular Plant Pathology* **5**, 93-104.

Gao, B., Allen, R., Maier, T., Davis, E., Baum, T. and Hussey, R. (2003) The parasitome of the phytonematode *Heterodera glycines*. *Molecular Plant-Microbe Interactions* **16**, 720-726.

Gao, B., Allen, R., Maier, T., Davis, E.L., Baum, T.J. and Hussey, R.S. (2002) Identification of a new beta-1,4-endoglucanase gene expressed in the esophageal subventral gland cells of *Heterodera glycines*. *Journal of Nematology* **34**, 12-15.

Gao, B., Allen, R., Maier, T., Davis, E.L., Baum, T.J. and Hussey, R.S. (2001) Molecular characterisation and expression of two venom allergen-like protein genes in *Heterodera glycines*. *International Journal for Parasitology* **31**, 1617-1625.

Gaulin, E., Dramé, N., Lafitte, C., Torto-Alalibo, T., Martinez, Y., Ameline-Torregrosa, C., Khatib, M., Mazarguil, H., Villalba-Mateos, F., Kamoun, S., Mazars, C., Dumas, B., Bottin, A., Esquerré-Tugayé, M.T. and Rickauer, M. (2006) Cellulose binding domains of a *Phytophthora* cell wall protein are novel pathogen-associated molecular patterns. *The Plant Cell* **18**, 1766-1777.

Gaur, H.S., Beane, J. and Perry, R.N. (2000) The influence of root diffusate, host age and water regimes on hatching of the root-knot nematode, *Meloidogyne triticoryzae*. *Nematology* **2**, 191-199.

Gheysen, G. and Fenoll, C. (2002) Gene expression in nematode feeding sites. *Annual Review of Phytopathology* **40**, 191-219.

Gilkes, N.R., Claeyssens, M., Aebersold, R., Henrissat, B., Meinke, A., Morrison, H.D., Kilburn, D.G., Warren, R.A.J. and Miller, R.C. (1991a) Structural and functional relationships in 2 families of beta-1,4-glycanases. *European Journal of Biochemistry* **202**, 367-377.

Gilkes, N.R., Henrissat, B., Kilburn, D.G., Miller Jr, R.C. and Warren, R.A.J. (1991b) Domains in microbial ß-1,4-glycanases: Sequence conservation, function, and enzyme families. *Microbiological Reviews* **55**, 303-315.

Girard, C. and Jouanin, L. (1999) Molecular cloning of cDNAs encoding a range of digestive enzymes from a phytophagous beetle, *Phaedon cochleariae*. *Insect Biochemistry and Molecular Biology* **29**, 1129-1142.

Goellner, M., Smant, G., De Boer, J.M., Baum, T.J. and Davis, E.L. (2000) Isolation of beta-1,4-endoglucanase genes from *Globodera tabacum* and their expression during parasitism. *Journal of Nematology* **32**, 154-165.

Golden, A.M., O'Bannon, J.H., Santo, G.S. and Finley, A.M. (1980) Description and SEM observations of *Meloidogyne chitwoodi* n.sp. (Meloidogynidae), a root-knot

nematode on potato in the Pacific Northwest. *Journal of Nematology* **12**, 319-327.

Gomez Gomez, E., Ruiz Roldan, M.C., Di Pietro, A., Roncero, M.I.G. and Hera, C. (2002) Role in pathogenesis of two endo-beta-1,4-xylanase genes from the vascular wilt fungus *Fusarium oxysporum*. *Fungal Genetics and Biology* **35**, 213-222.

Hansen, E., Harper, G., McPherson, M.J. and Atkinson, H.J. (1996) Differential expression patterns of the wound-inducible transgene *wun1- uidA* in potato roots following infection with either cyst or root knot nematodes. *Physiological and Molecular Plant Pathology* **48**, 161-170.

Harcus, Y.M., Parkinson, J., Fernández, C., Daub, J., Selkirk, M.E., Blaxter, M.L. and Maizels, R.M. (2004) Signal sequence analysis of expressed sequence tags from the nematode *Nippostrongylus brasiliensis* and the evolution of secreted proteins in parasites. *Genome Biology* **5**, R39.

Hatada, Y., Saito, K., Koike, K., Yoshimatsu, T., Ozawa, T., Kobayashi, T. and Ito, S. (2000) Deduced amino-acid sequence and possible catalytic residues of a novel pectate lyase from an alkaliphilic strain of *Bacillus*. *European Journal of Biochemistry* **267**, 2268-2275.

Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal* **280**, 309-316.

Henrissat, B. and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochemical Journal* **316**, 695-696.

Herron, S.R., Benen, J.A.E., Scavetta, R.D., Visser, J. and Jurnak, F. (2000) Structure and function of pectic enzymes: virulence factors of plant pathogens. *Proceedings of the National Academy of Sciences U.S.A.* **97**, 8762-8769.

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Meg, J., Trevaskis, E., Underwood, K., Wohldman, P., Waterston, R., Wilson, R. and Marra, M. (1996) Generation and analysis of 280,000 human expressed sequence tags. *Genome Research* **6**, 807-828.

Horton, P., Park, K.-J., Obayashi, T. and Nakai, K. (2006) Protein Subcellular Localization Prediction with WoLF PSORT. In *The 4th annual Asian Pacific Bioinformatics Conference APBC06* (Jiang, T., Yang, U.-C. and Chen, Y.-P.P., eds.), pp. 39-48. Taipei,Taiwan, Imperial College Press, London.

Huang, G., Dong, R., Allen, R., Davis, E.L., Baum, T.J. and Hussey, R.S. (2006) A root-knot nematode secretory peptide functions as a ligand for a plant transcription factor. *Molecular Plant-Microbe Interactions* **19**, 463-470.

Huang, G., Dong, R., Allen, R., Davis, E.L., Baum, T.J. and Hussey, R.S. (2005) Developmental expression and molecular analysis of two *Meloidogyne incognita* pectate lyase genes. *International Journal for Parasitology* **35**, 685-692.

Huang, G., Dong, R., Maier, T., Allen, R., Davis, E.L., Baum, T.J. and Hussey, R.S. (2004a) Use of solid-phase subtractive hybridization for the identification of parasitism gene candidates from the root-knot nematode *Meloidogyne incognita*. *Molecular Plant Pathology* **5**, 217-222.

Huang, G., Gao, B., Maier, T., Allen, R., Davis, E.L., Baum, T.J. and Hussey, R.S. (2003) A profile of putative parasitism genes expressed in the esophageal gland cells of the root-knot nematode *Meloidogyne incognita*. *Molecular Plant-Microbe Interactions* **16**, 376-381.

Huang, K.F., Ko, T.P., Hung, C.C., Chu, J., Wang, A.H.J. and Chiou, S.H. (2004b) Crystal structure of a platelet-agglutinating factor isolated from the venom of Taiwan habu (*Trimeresurus mucrosquamatus*). *Biochemical Journal* **378**, 399-407.

Huang, Q. and Allen, C. (1997) An exo-poly-alfa-D-galacturonosidase, PehB, is required for wild-type virulence of *Ralstonia solanacearum*. *Journal of Bacteriology* **179**, 7369-7378.

Hubert, J., Sustr, V. and Smrz, J. (1999) Feeding of the oribatid mite *Scheloribates laevigatus* (Acari: Oribatida) in laboratory experiments. *Pedobiologia* **43**, 328-339.

Hugouvieux-Cotte-Pattat, N., Condemine, G., Nasser, W. and Reverchon, S. (1996) Regulation of pectinolysis in *Erwinia chrysanthemi*. *Annual Review of Microbiology* **50**, 213-257.

Hussey, R., Davis, E. and Ray, C. (1994) *Meloidogyne* stylet secretions. In *Advances in molecular plant nematology* (Lamberti, F., ed.), pp. 233-249. New York: Plenum Press.

Hussey, R., Davis, E.L. and Baum, T. (2002) Secrets in secretions: genes that control nematode parasitism of plants. *Brazilian journal of plant physiology* **14**, 183-194.

Hussey, R.S. (1989) Disease-inducing secretions of plant-parasitic nematodes. *Annual review of phytopathology* **27**, 123-141.

Hussey, R.S. and Barker, K.R. (1973) A comparison of methods of collecting inocula of *Meloidogyne* spp., including a new technique. *Plant Disease Reporter* **57**, 1025-1028.

Hussey, R.S. and Mims, C.W. (1990) Ultrastructure of esophageal glands and their secretory granules in the root-knot nematode *Meloidogyne incognita*. *Protoplasma* **156**, 9-18.

Ingham, R.E., Hamm, P.B., Williams, R.E. and Swanson, W.H. (2000) Control of *Meloidogyne chitwoodi* in potato with fumigant and nonfumigant nematicides. *Journal of Nematology* **32**, 556-565.

Inoue, T., Moriya, S., Ohkuma, M. and Kudo, T. (2005) Molecular cloning and characterization of a cellulase gene from a symbiotic protist of the lower termite, *Coptotermes formosanus*. *Gene* **349**, 67-75.

Jaubert, S., Laffaire, J.B., Abad, P. and Rosso, M.N. (2002) A polygalacturonase of animal origin isolated from the root-knot nematode *Meloidogyne incognita*. *FEBS Letters* **522**, 109-112.

Jenkins, J., Leggio, L.L., Harris, G. and Pickersgill, R. (1995) ß-glucosidase, ß-galactosidase, family A cellulases, family F xylanases and two barley glycanases form a superfamily of enzymes with 8-fold, beta/alfa architecture and with two conserved glutamates near the carboxy-terminal ends of ß-strands four and seven. *FEBS Letters* **362**, 281-285.

Jones, J.T., Furlanetto, C., Bakker, E., Banks, B., Blok, V., Chen, Q., Phillips, M. and Prior, A. (2003) Characterization of a chorismate mutase from the potato cyst nematode *Globodera pallida*. *Molecular Plant Pathology* **4**, 43-50.

Jones, J.T., Reavy, B., Smant, G. and Prior, A.E. (2004) Glutathione peroxidases of the potato cyst nematode *Globodera rostochiensis*. *Gene* **324**, 47-54.

Jones, M.G.K. and Dropkin, V.H. (1976) Scanning electron microscopy of nematode induced giant transfer cells. *Cytobios* **15**, 149-161.

Kamoun, S. (2006) A catalogue of the effector secretome of plant pathogenic oomycetes. *Annual Review of Phytopathology* **44**, 41-60.

Kamoun, S., Hraber, P., Sobral, B., Nuss, D. and Govers, F. (1999) Initial assessment of gene diversity for the oomycete pathogen *Phytophthora infestans* based on expressed sequences. *Fungal Genetics and Biology* **28**, 94-106.

Karssen, G. (1999) The plant-parasitic nematode genus *Meloidogyne* Göldi, 1892 (Tylenchida) in Europe. In *Biology Department*, pp. 161. Gent, University of Gent.

Karssen, G. and Den Nijs, L. (2004) *Meloidogyne chitwoodi* and *Meloidogyne fallax*. *EPPO Bulletin* **34**, 315-320.

Kataeva, I.A., Seidel III, R.D., Shah, A., West, L.T., Li, X.-L. and Ljungdahl, L.G. (2002) The fibronectin type 3-like repeat from the *Clostridium thermocellum* cellobiohydrolase CbHa promotes hydrolysis of cellulose by modifying its surface. *Applied and Environmental Microbiology* **68**, 4292-4300.

Keen, N.T., Boyd, C. and Henrissat, B. (1996) Cloning and characterization of a xylanase gene from corn strains of Erwinia chrysanthemi. *Molecular Plant-Microbe Interactions* **9**, 651-657.

Kende, H., Bradford, K.J., Brummell, D.A., Cho, H.T., Cosgrove, D.J., Fleming, A.J., Gehring, C., Lee, Y., McQueen-Mason, S., Rose, J.K.C. and Voesenek, L.A.C.J. (2004) Nomenclature for members of the expansin superfamily of genes and proteins. *Plant Molecular Biology* **55**, 311-314.

Kikuchi, T., Jones, J.T., Aikawa, T., Kosaka, H. and Ogura, N. (2004) A family of glycosyl hydrolase family 45 cellulases from the pine wood nematode *Bursaphelenchus xylophilus*. *FEBS Letters* **572**, 201-205.

Kikuchi, T., Shibuya, H., Aikawa, T. and Jones, J.T. (2006) Cloning and characterization of pectate lyases expressed in the esophageal gland of the pine wood nematode *Bursaphelenchus xylophilus*. *Molecular Plant-Microbe Interactions* **19**, 280-287.

Kim, W.Y., Ha, J.K., Han, I.K., Lee, S.S., Shin, K.J. and Kim, W.Y. (1999) The Rumen Ecosystem : As a Fountain Source of Nobel Enzymes - Review. *Asian-Australasian Journal of Animal Sciences* **12**, 988-1001.

Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology* **305**, 567-580.

Kudla, U., Qin, L., Milac, A., Kielak, A., Maissen, C., Overmars, H., Popeijus, H., Roze, E., Petrescu, A., Smant, G., Bakker, J. and Helder, J. (2005) Origin, distribution and 3D-modeling of Gr-EXPB1, an expansin from the potato cyst nematode *Globodera rostochiensis*. *FEBS Letters* **579**, 2451-2457.

Kulkarni, N., Shendye, A. and Rao, M. (1999) Molecular and biotechnological aspects of xylanases. *FEMS Microbiology Reviews* **23**, 411-456.

Laine, M.J., Haapalainen, M., Wahlroos, T., Kankare, K., Nissinen, R., Kassuwi, S. and Metzler, M.C. (2000) The cellulase encoded by the native plasmid of *Clavibacter michiganensis* ssp. *sepedonicus* plays a role in virulence and contains an expansin-like domain. *Physiological and Molecular Plant Pathology* **57**, 221-233.

Lambshead, P.J.D. (1993) Recent developments in marine benthic biodiversity research. *Oceanis* **19**, 5-24.

Larson, S.B., Day, J., de la Rosa, A.P.B., Keen, N.T. and McPherson, A. (2003) First crystallographic structure of a xylanase from glycoside hydrolase family 5: Implications for catalysis. *Biochemistry* **42**, 8411-8422.

Ledger, T.N., Jaubert, S., Bosselut, N., Abad, P. and Rosso, M.N. (2006) Characterization of a new ß-1,4-endoglucanase gene from the root-knot nematode *Meloidogyne incognita* and evolutionary scheme for phytonematode family 5 glycosyl hydrolases. *Gene* **382**, 121-128.

Lee, Y., Choi, D. and Kende, H. (2001) Expansins: ever-expanding numbers and functions. *Current Opinion in Plant Biology* **4**, 527-532.

Lemesle-Varloot, L., Henrissat, B., Gaboriaud, C., Bissery, V., Morgat, A. and Mornon, J.P. (1990) Hydrophobic cluster analysis: procedures to derive structural and functional information from 2-D-representation of protein sequences. *Biochimie* **72**, 555-574.

Levy, I., Shani, Z. and Shoseyov, O. (2002) Modification of polysaccharides and plant cell wall by endo-1,4-ß-glucanase and cellulose-binding domains. *Biomolecular Engineering* **19**, 17-30.

Li, Y., Darley, C.P., Ongaro, V., Fleming, A., Schipper, O., Baldauf, S.L. and McQueen-Mason, S.J. (2002) Plant expansins are a complex multigene family with an ancient evolutionary origin. *Plant Physiology* **128**, 854-864.

Li, Y., Jones, L. and McQueen-Mason, S. (2003) Expansins and cell growth. *Current Opinion in Plant Biology* **6**, 603-610.

Liao, C.H., Revear, L., Hotchkiss, A. and Savary, B. (1999) Genetic and biochemical characterization of an exopolygalacturonase and a pectate lyase from *Yersinia enterocolitica*. *Canadian Journal of Microbiology* **45**, 396-403.

Little, E., Bork, P. and Doolittle, R.F. (1994) Tracing the spread of fibronectin type III domains in bacterial glycohydrolases. *Journal of Molecular Evolution* **39**, 631-643.

Marchler-Bauer, A. and Bryant, S.H. (2004) CD-Search: Protein domain annotations on the fly. *Nucleic Acids Research* **32**, W327-331.

Marchler-Bauer, A., Anderson, J.B., Cherukuri, P.F., DeWeese-Scott, C., Geer, L.Y., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Marchler, G.H., Mullokandov, M., Shoemaker, B.A., Simonyan, V., Song, J.S., Thiessen, P.A., Yamashita, R.A., Yin, J.J., Zhang, D. and Bryant, S.H. (2005) CDD: A Conserved Domain Database for protein classification. *Nucleic Acids Research* **33**, D192-D196.

Marchler-Bauer, A., Anderson, J.B., Derbyshire, M.K., DeWeese-Scott, C., Gonzales, N.R., Gwadz, M., Hao, L., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Krylov, D., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Lu, S., Marchler, G.H., Mullokandov, M., Song, J.S., Thanki, N., Yamashita, R.A., Yin, J.J., Zhang, D. and Bryant, S.H. (2007) CDD: A conserved domain database for interactive domain family analysis. *Nucleic Acids Research* **35**, D237-D240.

Matoub, M. and Rouland, C. (1995) Purification and properties of the xylanases from the termite *Macrotermes bellicosus* and its symbiotic fungus *Termitomyces* spp. *Comparative Biochemistry and Physiology part B* **112**, 629-635.

McCarter, J., Abad, P., Jones, J.T. and Bird, D. (2000) Rapid gene discovery in plant parasitic nematodes *via* Expressed Sequence Tags. *Nematology* **2**, 719-731.

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. and Waterston, R.H. (2003) Analysis and functional classification of transcripts from the nematode *Meloidogyne incognita*. *Genome Biology* **4**, R26.

McCartney, L., Blake, A.W., Flint, J., Bolam, D.N., Boraston, A.B., Gilbert, H.J. and Knox, J.P. (2006) Differential recognition of plant cell walls by microbial xylan-specific carbohydrate-binding modules. *Proceedings of the National Academy of Sciences U.S.A.* **103**, 4765-4770.

McClure, M.A., Kruk, T.H. and Misaghi, I. (1973) A method for obtaining quantities of clean *Meloidogyne* eggs. *Journal of Nematology* **5**, 230.

McNeil, M., Darvill, A.G., Fry, S.C. and Albersheim, P. (1984) Structure and function of the primary cell walls of plants. *Annual Review of Biochemistry* **53**, 625-663.

McQueen-Mason, S. and Cosgrove, D.J. (1994) Disruption of hydrogen bonding between plant cell wall polymers by proteins that induce wall extension. *Proceedings of the National Academy of Sciences U.S.A.* **91**, 6574-6578.

McSorley, R. (2003) Adaptations of nematodes to environmental extremes. *Florida Entomologist* **86**, 138-142.

Menne, K.M.L., Hermjakob, H. and Apweiler, R. (2000) A comparison of signal sequence prediction methods using a test set of signal peptides. *Bioinformatics* **16**, 741-742.

Mitreva, M., Elling, A.A., Dante, M., Kloek, A.P., Kalyanaraman, A., Aluru, S., Clifton, S.W., Bird, D.M., Baum, T.J. and McCarter, J.P. (2004a) A survey of SL1-spliced transcripts from the root-lesion nematode *Pratylenchus penetrans*. *Molecular Genetics and Genomics* **272**, 138-148.

Mitreva, M., McCarter, J.P., Martin, J., Dante, M., Wylie, T., Chiapelli, B., Pape, D., Clifton, S.W., Nutman, T.B. and Waterston, R.H. (2004b) Comparative genomics of gene expression in the parasitic and free-living nematodes *Strongyloides stercoralis* and *Caenorhabditis elegans*. *Genome Research* **14**, 209-220.

Mitreva-Dautova, M., Roze, E., Overmars, H., de Graaff, L., Schots, A., Helder, J., Goverse, A., Bakker, J. and Smant, G. (2006) A symbiont-independent endo-1,4-ß-xylanase from the plant-parasitic nematode *Meloidogyne incognita*. *Molecular Plant-Microbe Interactions* **19**, 521-529.

Mouratov, S., Lahav, I., Barness, G. and Steinberger, Y. (2001) Preliminary study of the soil nematode community at Machu Picchu Station, King George Island, Antarctica. *Polar Biology* **24**, 545-548.

Murzin, A.G., Brenner, S.E., Hubbard, T. and Chothia, C. (1995) SCOP: A structural classification of proteins database for the investigation of sequences and structures. *Journal of Molecular Biology* **247**, 536-540.

Nielsen, H., Brunak, S. and Von Heijne, G. (1999) Machine learning approaches for the prediction of signal peptides and other protein sorting signals. *Protein Engineering* **12**, 3-9.

Nishitani, K. (1997) The role of endoxyloglucan transferase in the organization of plant cell walls. *International Review of Cytology* **173**, 157-206.

O'bannon, J.H. (1982) Host range of the Columbia root-knot nematode. *Plant Disease* **66**, 1045-1048.

Oostenbrink, M. (1960) Estimating nematode populations by some selected methods. In *Nematology: fundamentals and recent advances with emphasis on plant parasitic and soil forms* (Sasser, J.N. and Jenkins, W.R., eds.), pp. 85-102. Raleigh: Chapel Hill: University of North Carolina Press.

Papadopoulou, J. and Triantaphyllou, A.C. (1982) Sex differentiation in *Meloidogyne incognita* and anatomical evidence of sex reversal. *Journal of Nematology* **14**, 549-566.

Pearson, M.S., McManus, D.P., Smyth, D.J., Lewis, F.A. and Loukas, A. (2005) In vitro and in silico analysis of signal peptides from the human blood fluke, *Schistosoma mansoni*. *FEMS Immunology and Medical Microbiology* **45**, 201-211.

Perry, R.N. (1989) Dormancy and hatching of nematode eggs. *Parasitology Today* **5**, 377-383.

Pinkerton, J.N., Santo, G.S., Ponti, R.P. and Wilson, J.H. (1986) Control of *Meloidogyne chitwoodi* on commercially grown Russet Burbank potatoes. *Plant Disease* **70**, 860-863.

Pitman, A.J., Jones, E.B.G. and Oevering, P. (2003) An overview of the biology of the wharf borer beetle (*Nacerdes melanura* L., Oedemeridae) a pest of wood in marine structures. *Biofouling* **19**, 239-248.

Ponting, C.P., Aravind, L., Schultz, J., Bork, P. and Koonin, E.V. (1999) Eukaryotic signalling domain homologues in archaea and bacteria. Ancient ancestry and horizontal gene transfer. *Journal of Molecular Biology* **289**, 729-745.

Popeijus, H., Blok, V.C., Cardle, L., Bakker, E., Phillips, M.S., Helder, J., Smant, G. and Jones, J.T. (2000a) Analysis of genes expressed in second stage juveniles of the potato cyst nematodes *Globodera rostochiensis* and *G. pallida* using the expressed sequence tag approach. *Nematology* **2**, 567-574.

Popeijus, H., Overmars, H., Jones, J., Blok, V., Goverse, A., Helder, J., Schots, A., Bakker, J. and Smant, G. (2000b) Degradation of plant cell walls by a nematode. *Nature* **406**, 36-37.

Preston, G.M., Studholme, D.J. and Caldelari, I. (2005) Profiling the secretomes of plant pathogenic Proteobacteria. *FEMS Microbiology Reviews* **29**, 331-360.

Prior, A., Jones, J.T., Blok, V.C., Beauchamp, J., McDermott, L., Cooper, A. and Kennedy, M.W. (2001) A surface-associated retinol- and fatty acid-binding protein (Gp-FAR-1) from the potato cyst nematode *Globodera pallida*: lipid binding activities, structural analysis and expression pattern. *The biochemical journal* **356**, 387-394.

Py, B., Bortolil-German, I., Haiech, J., Chippaux, M. and Barras, F. (1991) Cellulase EGZ of *Erwinia chrysanthemi*: Structural organization and importance of His98 and Glu133 residues for catalysis. *Protein Engineering* **4**, 325-333.

Qin, L., Kudla, U., Roze, E.H.A., Goverse, A., Popeijus, H., Nieuwland, J., Overmars, H., Jones, J.T., Schots, A., Smant, G., Bakker, J. and Helder, J. (2004) Plant degradation: A nematode expansin acting on plants. *Nature* **427**, 30.

Qin, L., Overmars, H., Helder, J., Popeijus, H., Rouppe van der Voort, J., Groenink, W., van Koert, P., Schots, A., Bakker, J. and Smant, G. (2000) An efficient cDNA-AFLP-based strategy for the identification of putative pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*. *Molecular Plant-Microbe Interactions* **13**, 830-836.

Qutob, D., Hraber, P.T., Sobral, B.W.S. and Gijzen, M. (2000) Comparative analysis of expressed sequences in *Phytophthora sojae*. *Plant Physiology* **123**, 243-253.

Ray, S.K., Rajeshwari, R. and Sonti, R.V. (2000) Mutants of *Xanthomonas oryzae* pv. *oryzae* deficient in general secretory pathway are virulence deficient and unable to secrete xylanase. *Molecular Plant-Microbe Interactions* **13**, 394-401.

Robertson, L., Robertson, W.M. and Jones, J.T. (1999) Direct analysis of the secretions of the potato cyst nematode *Globodera rostochiensis*. *Parasitology* **119**, 167-176.

Robertson, L., Robertson, W.M., Sobczak, M., Helder, J., Tetaud, E., Ariyanayagam, M.R., Ferguson, M.A., Fairlamb, A. and Jones, J.T. (2000) Cloning, expression and functional characterisation of a peroxiredoxin from the potato cyst nematode *Globodera rostochiensis*. *Molecular and Biochemical Parasitology* **111**, 41-49.

Rosso, M.N., Dubrana, M.P., Cimbolini, N., Jaubert, S. and Abad, P. (2005) Application of RNA interference to root-knot nematode genes encoding esophageal gland proteins. *Molecular Plant-Microbe Interactions* **18**, 615.

Rosso, M.N., Favery, B., Piotte, C., Arthaud, L., De Boer, J.M., Hussey, R.S., Bakker, J., Baum, T.J. and Abad, P. (1999) Isolation of a cDNA encoding a beta-1,4-endoglucanase in the root-knot nematode *Meloidogyne incognita* and expression analysis during plant parasitism. *Molecular Plant-Microbe Interactions* **12**, 585-591.

Saloheimo, M., Paloheimo, M., Hakola, S., Pere, J., Swanson, B., Nyyssönen, E., Bhatia, A., Ward, M. and Penttilä, M. (2002) Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expansins, exhibits disruption activity on cellulosic materials. *European Journal of Biochemistry* **269**, 4202-4211.

Santo, G.S. and Wilson, J.H. (1990) Evaluation of ethoprop and cadusafos for control of *Meloidogyne chitwoodi* on potato. *Nematropica* **20**, 137-142.

Santo, G.S., O'bannon, J.H., Finley, A.M. and Golden, A.M. (1980) Occurrence and host range of a new root-knot nematode (*Meloidogyne chitwoodi*) in the Pacific Northwest. *Plant Disease* **64**, 951-952.

Schneider, G. and Fechner, U. (2004) Advances in the prediction of protein targeting signals. *Proteomics* **4**, 1571-1580.

Scrivener, A.M., Watanabe, H. and Noda, H. (1998) Properties of digestive carbohydrase activities secreted by two cockroaches, *Panesthia cribrata* and *Periplaneta americana*. *Comparative Biochemistry and Physiology part B* **119**, 273-282.

Semblat, J.P., Rosso, M.N., Hussey, R.S., Abad, P. and Castagnone-Sereno, P. (2001) Molecular cloning of a cDNA encoding an amphid-secreted putative avirulence protein from the root-knot nematode *Meloidogyne incognita*. *Molecular Plant-Microbe Interactions* **14**, 72-79.

Shevchik, V.E., Robert-Baudouy, J. and Hugouvieux-Cotte-Pattat, N. (1997) Pectate lyase PelI of *Erwinia chrysanthemi* 3937 belongs to a new family. *Journal of Bacteriology* **179**, 7321-7330.

Shoham, Y., Lamed, R. and Bayer, E.A. (1999) The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides. *Trends in Microbiology* **7**, 275-281.

Shoseyov, O., Shani, Z. and Levy, I. (2006) Carbohydrate binding modules: biochemical properties and novel applications. *Microbiology and Molecular Biology Reviews* **70**, 283-295.

Showalter, A.M. (1993) Structure and function of plant cell wall proteins. *The Plant Cell* **5**, 9-23.

Sijmons, P.C., Grundler, F.M.W., Von Mende, N., Burrows, P.R. and Wyss, U. (1991) *Arabidopsis thaliana* as a new model host for plant-parasitic nematodes. *Plant Journal* **1**, 245-254.

Skinner, W., Keon, J. and Hargreaves, J. (2001) Gene information for fungal plant pathogens from expressed sequences. *Current Opinion in Microbiology* **4**, 381-386.

Smant, G., Goverse, A., Stokkermans, J.P.W.G., De Boer, J.M., Pomp, H., Zilverentant, J.F., Overmars, H.A., Helder, J., Schots, A. and Bakker, J. (1997) Potato root diffusate-induced secretion of soluble, basic proteins originating from the subventral esophageal glands of potato cyst nematodes. *Phytopathology* **87**, 839-845.

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., Schots, A. and Bakker, J. (1998) Endogenous cellulases in animals: Isolation of b-1,4-

endoglucanase genes from two species of plant-parasitic cyst nematodes. *Proceedings of the National Academy of Sciences U.S.A.* **95**, 4906-4911.

Sohlenius, B. (1980) Abundance, biomass and contribution to energy flow by soil nematodes in terrestrial ecosystems. *Oikos* **34**, 186-194.

Sugimura, M., Watanabe, H., Lo, N. and Saito, H. (2003) Purification, characterization, cDNA cloning and nucleotide sequencing of a cellulase from the yellow-spotted longicorn beetle, *Psacothea hilaris*. *European Journal of Biochemistry* **270**, 3455-3460.

Suzuki, M., Kato, A., Nagata, N. and Komeda, Y. (2002) A xylanase, AtXyn1, is predominantly expressed in vascular bundles, and four putative xylanase genes were identified in the *Arabidopsis thaliana* genome. *Plant and Cell Physiology* **43**, 759-767.

Talbott, L.D. and Ray, P.M. (1992) Molecular size and separability features of pea cell wall polysaccharides: Implications for models of primary wall structure. *Plant Physiology* **98**, 357-368.

Tamaru, Y. and Doi, R.H. (2001) Pectate lyase A, an enzymatic subunit of the *Clostridium cellulovorans* cellulosome. *Proceedings of the National Academy of Sciences U.S.A.* **98**, 4125-4129.

Ten Have, A., Tenberge, K.B., Benen, J.A.E., Tudzynski, P., Visser, J. and van Kan, J.A.L. (2002) The contribution of cell wall degrading enzymes to pathogenesis of fungal plant pathogens. In *The Mycota XI. Agricultural Applications* (Kempken, F., ed.), pp. 341-358. Berlin Heidelberg: Springer-Verlag.

Tomme, P., Boraston, A., McLean, B., Kormos, J., Creagh, A.L., Sturch, K., Gilkes, N.R., Haynes, C.A., Warren, R.A.J. and Kilburn, D.G. (1998) Characterization and affinity applications of cellulose-binding domains. *Journal of Chromatography B* **715**, 283-296.

Torto, T.A., Li, S., Styer, A., Huitema, E., Testa, A., Gow, N.A.R., van West, P. and Kamoun, S. (2003) EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen *Phytophthora*. *Genome Research* **13**, 1675-1685.

Treonis, A.M. and Wall, D.H. (2005) Soil nematodes and desiccation survival in the extreme arid environment of the Antarctic Dry Valleys. *Integrative and Comparative Biology* **45**, 741-750.

Trudgill, D.L. (1997) Parthenogenetic root-knot nematodes (*Meloidogyne* spp.); how can these biotrophic endoparasities have such an enormous host range? *Plant Pathology* **46**, 26-32.

Trudgill, D.L. and Blok, V.C. (2001) Apomictic, polyphagous root-knot nematodes: exceptionally successful and damaging biotrophic root pathogens. *Annual Review of Phytopathology* **39**, 53-77.

Turkiewicz, M., Kalinowska, H., Zielinska, M. and Bielecki, S. (2000) Purification and characterization of two endo-1,4-beta-xylanases from Antarctic krill, *Euphausia superba* Dana. *Comparative Biochemistry and Physiology Part B* **127B**, 325-335.

Tytgat, T., Vanholme, B., De Meutter, J., Claeys, M., Couvreur, M., Vanhoutte, I., Gheysen, G., Van Criekinge, W., Borgonie, G., Coomans, A. and Gheysen, G. (2004) A new class of ubiquitin extension proteins secreted by the dorsal pharyngeal gland in plant parasitic cyst nematodes. *Molecular Plant-Microbe Interactions* **17**, 846-852.

Uehara, T., Kushida, A. and Momota, Y. (2001) PCR-based cloning of two ß-1,4-endoglucanases from the root-lesion nematode *Pratylenchus penetrans*. *Nematology* **3**, 335-341.

Urwin, P.E., Lilley, C.J. and Atkinson, H.J. (2002) Ingestion of double-stranded RNA by preparasitic juvenile cyst nematodes leads to RNA interference. *Molecular Plant-Microbe Interactions* **15**, 747-752.

Vanholme, B., De Meutter, J., Tytgat, T., Van Montagu, M., Coomans, A. and Gheysen, G. (2004) Secretions of plant-parasitic nematodes: a molecular update. *Gene* **332**, 13-27.

Vanholme, B., Mitreva, M., Van Criekinge, W., Logghe, M., Bird, D., McCarter, J.P. and Gheysen, G. (2006) Detection of putative secreted proteins in the plant-parasitic nematode *Heterodera schachtii*. *Parasitology Research* **98**, 414-424.

Vanholme, B., van Thuyne, W., Vanhouteghem, K., De Meutter, J., Cannoot, B. and Gheysen, G. (2007) Molecular characterization and functional importance of pectate lyase secreted by the cyst nematode *Heterodera schachtii*. *Molecular Plant Pathology* **8**, 267-278.

Wall, D.H. and Virginia, R.A. (1999) Controls on soil biodiversity: insights from extreme environments. *Applied Soil Ecology* **13**, 137-150.

Walton, J.D. (1994) Deconstructing the cell wall. *Plant Physiology* **104**, 1113-1118.

Wang, J., Ding, M., Li, Y.H., Chen, Q.X., Xu, G.j. and Zhao, F.k. (2003) A monovalent anion affected multi-functional cellulase EGX from the mollusca, *Ampullaria crossean*. *Protein Expression and Purification* **31**, 108-114.

Wang, X., Allen, R., Ding, X., Goellner, M., Maier, T., De Boer, J.M., Baum, T.J., Hussey, R.S. and 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., Meyers, D., Yan, Y., Baum, T., Smant, G., Hussey, R. and Davis, E. (1999) In planta localization of a beta-1,4-endoglucanase secreted by *Heterodera glycines*. *Molecular Plant-Microbe Interactions* **12**, 64-67.

Williamson, V.M. and Gleason, C.A. (2003) Plant-nematode interactions. *Current Opinion in Plant Biology* **6**, 327-333.

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

Wishart, J., Phillips, M.S. and Blok, V.C. (2002) Ribosomal intergenic spacer: A polymerase chain reaction diagnostic for *Meloidogyne chitwoodi*, *M. fallax*, and *M. hapla*. *Phytopathology* **92**, 884-892.

Wu, S.C., Ham, K.S., Darvill, A.G. and Albersheim, P. (1997) Deletion of two endo-beta-1,4-xylanase genes reveals additional isozymes secreted by the rice blast fungus. *Molecular Plant-Microbe Interactions* **10**, 700-708.

Wubah, D.A., Akin, D.E. and Borneman, W.S. (1993) Biology, fiber-degradation, and enzymology of anaerobic zoosporic fungi. *Critical Reviews in Microbiology* **19**, 99-115.

Wylie, T., Martin, J.C., Dante, M., Mitreva, M.D., Clifton, S.W., Chinwalla, A., Waterston, R.H., Wilson, R.K. and McCarter, J.P. (2004) Nematode.net: a tool for navigating sequences from parasitic and free-living nematodes. *Nucleic Acids Research* **32**, D423-D426.

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

Xu, B., Hellman, U., Ersson, B. and Janson, J.C. (2000) Purification, characterization and amino-acid sequence analysis of a thermostable, low molecular mass endo-ß-1,4-glucanase from blue mussel, *Mytilus edulis*. *European Journal of Biochemistry* **267**, 4970-4977.

Yan, Y., Smant, G. and Davis, E.L. (2001) Functional screening yields a new beta-1,4-endoglucanase gene from *Heterodera glycines* that may be the product of recent gene duplication. *Molecular Plant-Microbe Interactions* **14**, 63-71.

Yan, Y., Smant, G., Stokkermans, J., Qin, L., Helder, J., Baum, T., Schots, A. and Davis, E. (1998) Genomic organization of four beta-1,4-endoglucanase genes in plant-parasitic cyst nematodes and its evolutionary implications. *Gene* **220**, 61-70.

Yano, A., Suzuki, K., Shinshi, H. and Uchimiya, H. (1998) Induction of hypersensitive cell death by a fungal protein in cultures of tobacco cells. *Molecular Plant-Microbe Interactions* **11**, 115-123.

Yeates, G.W., Bongers, T., de Goede, R.G.M., Freckman, D.W. and Georgieva, S.S. (1993) Feeding habits in soil nematode families and genera-an outline for soil ecologists. *Journal of Nematology* **25**, 315-331.

Zhu, Y.Y., Machleder, E.M., Chenchik, A., Li, R. and Siebert, P.D. (2001) Reverse transcriptase template switching: A SMART™ approach for full-length cDNA library construction. *BioTechniques* **30**, 892-897.

Zijlstra, C. (2000) Identification of *Meloidogyne chitwoodi*, *M. fallax* and *M. hapla* based on SCAR-PCR: A powerful way of enabling reliable identification of populations or individuals that share common traits. *European Journal of Plant Pathology* **106**, 283-290.

Curriculum vitae

Op 4 november 1977 werd Erwin Hubertus Andreas Roze geboren in Oirschot. Na drie jaar vertrok het gezin naar Asten, waar hij in 1990 naar het College Asten-Someren, nu het Varendonck College geheten, ging. De middelbare-schoolperiode sloot hij in 1996 af met een atheneumdiploma. In de zomer van datzelfde jaar vertrok hij naar Wageningen om aldaar Biologie te gaan studeren aan Wageningen Universiteit. Als specialisatierichting koos hij celbiologie, waarbinnen hij drie afstudeerprojecten deed. Het eerste project, getiteld 'Formiaatdehydrogenases en hydrogenases van *Syntrophobacter fumaroxidans*', liep van maart tot en met november 1999 en vond plaats aan het Laboratorium voor Microbiologie bij de werkgroep microbiële fysiologie. Aansluitend, van december 1999 tot juli 2000, hield hij zich bij het Laboratorium voor Nematologie bezig met de identificatie van pathogeniteitsfactoren van het aardappelcystenaaltje *Globodera rostochiensis* met behulp van cDNA-AFLP. Hierna haalde hij een Socratesbeurs binnen voor een stage bij het departement microbiële ecologie aan Lunds Universitet in Lund, Zweden. Dit afstudeerproject liep van augustus 2000 tot en met december 2000 en had als onderwerp '*Morphogenesis of the nematode-trapping fungus Arthrobotrys oligospora*'. In januari 2001 behaalde hij zijn diploma, hetgeen met lof verleend is. In februari 2001 startte hij zijn promotieonderzoek naar het secretoom van wortelknobbelaaltjes aan Wageningen Universiteit, Laboratorium voor Nematologie. Dit onderzoek vond plaats in het kader van het EU project '*Durable resistance against Meloidogyne*' (DREAM). De resultaten van het onderzoek staan beschreven in dit proefschrift. Van oktober 2006 tot januari 2007 was hij docent scheikunde in de klassen 4 en 5 VWO aan het Pantarijn te Wageningen. Op deze middelbare school zal hij vanaf november 2007 verdergaan als docent scheikunde en algemene natuurwetenschappen.

List of publications

Articles

Roze, E., Hanse, B., Mitreva, M., Vanholme, B., Bakker, J., Smant, G. (2008) 'Mining the secretome of the root-knot nematode *Meloidogyne chitwoodi* for candidate parasitism genes' *Molecular Plant Pathology 9 (1)*, p.1-10

Kudla, U., Milac, A., Qin, L., Overmars, H.A., Roze, E.H.A., Holterman, M.H.M., Petrescu, A., Goverse, A., Bakker, J., Helder, J., Smant, G. (2007) 'Structural and functional characterization of a novel, host penetration-related pectate lyase from the potato cyst nematode *Globodera rostochiensis*' *Molecular Plant Pathology 8 (3)*, p.293-305

Mitreva-Dautova, M., Roze, E., Overmars, H.A., Graaff, L.H. de, Schots, A., Helder, J., Goverse, A., Bakker, J., Smant, G. (2006) 'A symbiont-independent endo-1,4-beta-xylanase from the plant-parasitic nematode *Meloidogyne incognita*' *Molecular Plant-Microbe Interactions 19 (5)*, p.521-529

Kudla, U., Qin, L., Milac, A., Kielak, A., Maissen, C., Overmars, H.A., Popeijus, H.E., Roze, E.H.A., Petrescu, A., Smant, G., Bakker, J., Helder, J. (2005) 'Origin, distribution and 3D-modeling of Gr-EXPB1, an expansin from the potato cyst nematode *Globodera rostochiensis*' *FEBS Letters 579 (11)*, p.2451-2457

Roze, E.H.A., Jupovicz, J., Goverse, A., Helder, J., Bakker, J., Smant, G. (2004) 'Het wonder van de wortelknobbelaaltjes: de unieke moleculaire interacties tussen een obligate parasiet en haar waardplanten' *Gewasbescherming 35 (5)*, p.281-283

Qin, L., Kudla, U., Roze, E.H.A., Goverse, A., Popeijus, H., Nieuwland, J., Overmars, H.A., Jones, J.T., Schots, A., Smant, G., Bakker, J., Helder, J. (2004) 'Plant degradation: a nematode expansin acting on plants' *Nature 427*, p.30

Popeijus, H., Roze, E., Benen, J., Qin, L., Jones, J., Goverse, A., Helder, H., Schots, A., Bakker, J., Smant, G. (2002) 'Characterisation of a novel pectate lyase from *Globodera rostochiensis*' *In: The identification of cell wall degrading enzymes in Globodera rostochiensis*, thesis Wageningen UR, p.33-52

de Bok, F.A.M., Roze, E.H.A., Stams, A.J.M. (2002) 'Hydrogenases and formate dehydrogenases of *Syntrophobacter fumaroxidans*' *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology 81 (1-4)*, p.283-291

Congress proceedings and meetings

Roze, E.H.A., Mitreva-Dautova, M., McCarter, J., Stojanova, B., Overmars, H.A., Bakker, J., Smant, G. (2005) 'The repertoire of cell wall modifying proteins of the root-knot nematode *Meloidogyne chitwoodi*', *In: XII International Congress on Molecular Plant-Microbe Interactions*, Cancun, Mexico, 17-22 July 2005, p.160

Roze, E.H.A., Mitreva, M., Stojanova, B., Overmars, H.A., Qin, L., Bakker, J., Smant, G. (2005) 'The repertoire of cell wall modifying proteins of the Columbia root-knot nematode *Meloidogyne chitwoodi*', *Meloidogyne werkgroep*, Bergschenhoek, The Netherlands, 15 February 2005

Roze, E.H.A., Dautova, M., McCarter, J., Stojanova, B., Overmars, H.A., Bakker, J., Smant, G. (2004) 'The repertoire of cell wall modifying proteins of the root-knot nematode *Meloidogyne chitwoodi*' *In: XXVII European Society of Nematologists (ESN) International Symposium*, Rome, Italy, 14-18 June 2004, p.27

Roze, E.H.A., Mitreva, M., Stojanova, B., Overmars, H.A., Qin, L., Bakker, J., Smant, G. (2004) 'Cell wall modifying protein repertoire of the root-knot nematode *Meloidogyne chitwoodi*', *ALW Discussieplatform EPW*, Lunteren, The Netherlands, 6 April 2004

Helder, J., Qin, L., Pierzgalska, M., Roze, E., Goverse, A., Bakker, J., Smant, G. (2002) 'Yes, it is possible to find needles in a haystack : strategies for the identification of pathogenicity factors from cyst and root knot nematodes' *In: Disease Resistance in Plants: Autumn school 2002, Disease Resistance in Plants*, Wageningen, The Netherlands, 14-16 October 2002, p.75-76

Roze, E., Qin, L., Bakker, J., Smant, G. (2002) 'Identification of parasitism genes from root-knot nematode *Meloidogyne incognita* using cDNA-AFLP, EST-analysis and GenEST' *In: Nematology 4(2), International Journal of Fundamental and Applied Nematological Research: Fourth International Congress of Nematology*, Tenerife, Spain, 8-13 June 2002, p.224

Nawoord

Naast het bedanken van mensen is het nawoord bij uitstek geschikt om nog wat gras voor de voeten weg te maaien van collegae die zich mogelijk bezighouden met de voorbereiding van zogenaamde 'stukjes'. Zo is daar de vergelijking tussen het tot stand komen van het proefschrift en de snelle bevalling van onze dochters. Deze laatste zijn op een vroegtijdige en snelle wijze ter wereld gekomen. De bevalling van mijn proefschrift vergde echter heel wat meer tijd en gepuf. De meeste resultaten van het werk dat ik in de eerste pakweg twee en een half jaar van mijn aioschap uitgevoerd heb, hebben uiteindelijk nooit het levenslicht gezien, in die zin dat ze niet in dit proefschrift terecht zijn gekomen. Het vergde uiteindelijk heel wat energie en doorzettingsvermogen om de steven te keren en verder te gaan, maar het is uiteindelijk gelukt, getuige dit proefschrift.

De promotieperiode was natuurlijk niet allemaal kommer en kwel, maar heeft ook geresulteerd in interessante resultaten en kende ook veel leuke, mooie en leerzame momenten, waarvoor ik een aantal mensen wil bedanken. Allereerst Geert, mijn 'dagelijkse' begeleider. Ik heb op verschillende gebieden veel van je geleerd. Daarnaast vond en vind ik het geweldig dat er zoveel mogelijkheden geboden worden aan aio's om naar bijeenkomsten in het buitenland te gaan. Naast het opdoen van kennis en contacten in de wetenschappelijke wereld, heb ik ook delen van de wereld gezien waar ik anders waarschijnlijk nooit gekomen was. Mijn horizon is hierdoor dus op verschillende gebieden verbreed. Hiervoor wil ik uiteraard ook Jaap bedanken. Dat brengt mij bij de jaarlijkse vogelgeluidenwandeling, deze werd door Jaap leven ingeblazen en vond plaats onder mijn bezielende begeleiding. Ik bewaar mooie herinneringen aan deze wandelingen, die voor dag en dauw plaatsvonden en steevast eindigden met een versnapering bij Nol in het Bosch. Hiervoor wil ik ook de rest van de harde kern, bestaande uit Lisette, Erik, Martijn H., Jan R. en Rikus en consorten, en overige deelnemers bedanken.

Het werk werd ook geregeld opgeleukt door pauzes, vakgroepuitjes, lunches en diners. Ik heb veel gelachen tijdens deze intermezzo's en omdat er gedurende de jaren vele collega's zijn gekomen, gegaan en gebleven en ik niet het risico wil lopen ook maar iemand te vergeten: eenieder bedankt hiervoor!! Een aantal mensen wil ik nog in het bijzonder de revue laten passeren. Hein, Rikus, Jan R. en Liesbeth dank voor jullie hulp en ondersteuning bij tal van praktische zaken op het lab. Lisette, de uiteindelijke spin in het nematologieweb, dank voor je luisterend oor en je hulp bij allerlei papieren zaken. Je hebt me zelfs bijna over mijn formulierenangst heen

gekregen. Liesbeth en Nikkie, jullie waren geweldige buurvrouwen. Bram, een student waar iedere aio op hoopt, dank voor je hulp en tomeloze inzet en energie. Bartel, je ondersteuning bij het tot stand komen van hoofdstuk twee heb ik zeer gewaardeerd. In addition, I like to thank Makedonka for her input and advise during my research, I also really enjoyed my stay in St. Louis. Sven en Joost, jullie zijn goud waard, geweldig dat jullie mij als paranimf willen vergezellen.

Nu ga ik me op glad ijs bevinden, maar dat is het me wel waard. Pjotr, ik heb je leren kennen tijdens onze reis door de binnenlanden van Mexico, met een alles verwoestende en dodelijke orkaan op onze hielen. Uiteindelijk bracht dit ons samen op een rustige kamer op de tweede verdieping. Ik heb veel van je geleerd en je hebt me op verschillende manieren vooruitgeholpen en ondersteund, dank hiervoor!

Naast collega's hebben vrienden en familie natuurlijk ook een belangrijke rol gespeeld tijdens mijn promotieperiode, hetzij door me met raad en daad bij te staan hetzij door voor het broodnodige vertier en de ontspanning te zorgen. Erik S., Erik S., Nienke, Mark, Christien, Kristof en Richard, ik bewaar goede herinneringen aan de diners, Heerenlunches, wijnproefavonden, feestjes, wandelingen etc. etc.. Een deel van jullie is reeds naar het buitenland vertrokken, maar ik hoop dat er nog vele van deze samenscholingen zullen volgen. Pap en mam, Karin en Tim en Judiths familie, jullie hebben me altijd met interesse gevolgd, een luisterend oor geboden en me op allerlei manieren geholpen. De eerste twee vonden het zelfs niet erg om Myrthe af en toe onder hun hoede te nemen, opdat ik mij op het schrijven kon richten.

En dan tot slot nog de dames van de buitencategorie. Judith, ik heb veel van je gevraagd, maar je was er altijd voor me. Je hebt geluisterd, geholpen, gesteund, af en toe diep gezucht, maar bent bovenal in me blijven geloven. Ik kon mij niet meer wensen, je bent een kanjer! Zoals bij de opening reeds vermeld, hebben we er ondertussen ook twee fantastische dochters, Myrthe en Wende, bij gekregen. Ook zij hebben, voor het merendeel onbewust, bijgedragen aan de afronding van dit proefschrift. Ze hebben me geleerd meer te relativeren, door me veel te laten lachen en te laten zien dat niet alles te voorspellen valt, soms moet je gewoon vangen, handelen en doorgaan. Vooral Wende is daar bijzonder in geslaagd; op de Drentse keukenvloer werd niet alleen zij maar ook het inzicht geboren dat het uiteindelijk niet louter uitmaakt *hoe* je bevalt, *als* je maar bevalt.

Erwin.

Education Statement of the Graduate School
Experimental Plant Sciences

Issued to: Erwin Roze
Date: 15 February 2008
Group: Nematology, Wageningen University and Research Centre

1) Start-up phase	date
► **First presentation of your project**	
Identification of oesophageal secretions of *Meloidogyne* spp. using cDNA-AFLP and EST analysis	May 30, 2001
► **Writing or rewriting a project proposal**	
► **Writing a review or book chapter**	
► **MSc courses**	
► **Laboratory use of isotopes**	
Radiation course 'Safe handling with radioactive materials and sources' Level 5B	Mar 19-22, 2002
Subtotal Start-up Phase	*3.0 credits**

2) Scientific Exposure	date
► **EPS PhD student days**	
EPS PhD student day Wageningen	Jan 24, 2002
EPS PhD student day Utrecht	Mar 27, 2003
► **EPS theme symposia**	
Theme 2 Interactions between plants and biotic agents Symposium, Leiden	Dec 17, 2001
Theme 2 Interactions between plants and biotic agents Symposium, Amsterdam	Jan 10, 2003
Theme 2 Interactions between plants and biotic agents Symposium, Wageningen	Dec 12, 2003
Theme 2 Interactions between plants and biotic agents Symposium, Utrecht	Sep 17, 2004
Theme 2 Interactions between plants and biotic agents Symposium	Jun 23, 2005
► **NWO Lunteren days and other National Platforms**	
Meloidogyne werkgroep, Wageningen, PRI	Nov 18, 2003
Meloidogyne werkgroep, Lelystad, PPO-AGV	Mar 11, 2004
Meloidogyne werkgroep, Bergschenhoek, De Ruiter Seeds	Feb 15, 2005
NWO-ALW Lunteren meeting	Mar 26-27, 2001
NWO-ALW Lunteren meeting	Apr 07-08, 2003
NWO-ALW Lunteren meeting	Apr 05-06, 2004
NWO-ALW Lunteren meeting	Apr 04-05, 2005
► **Seminars (series), workshops and symposia**	
Flying seminars 2001	2001
Flying seminars 2002	2002
Frontiers in plant development/science	2002
Flying seminars 2003	2003
Frontiers in plant development/science	2003
Frontiers in research on interactions between plants and biotic agents	2004
Symposium on Systems Biology in honor of Prof. dr. Pierre de Wit	Nov 04, 2004
► **Seminar plus**	
► **International symposia and congresses**	
International SPIT meeting, Wageningen, The Netherlands	May 21-22, 2001
EU-DREAM annual meeting, Warsaw, Poland	Oct 25-26, 2001
Fourth International Congress of Nematology, Tenerife, Spain	Jun 08-13, 2002
EU-DREAM annual meeting, Dundee, Scotland	Oct 24-25, 2002
International SPIT meeting, Athens, University of Georgia, USA	Jun 01-04, 2003
EU-DREAM final meeting, Brussels, Belgium	Jan 16, 2004
XXVII International Symposium of European Society of Nematologists, Rome, Italy	Jun 14-18, 2004
XII International Congress on Molecular Plant-Microbe Interactions, Cancun, Mexico	Jul 17-20, 2005
► **Presentations**	
International SPIT meeting, Wageningen, The Netherlands, oral presentation	May 2001
EU-DREAM annual meeting, Warsaw, Poland, oral presentation	Oct 2001
Fourth International Congress of Nematology, Tenerife, Spain, poster presentation	Jun 2002
EU-DREAM annual meeting, Dundee, Scotland, oral presentation	Oct 2002
International SPIT meeting, Athens, University of Georgia, USA, oral presentation	Jun 2003
NWO-ALW Lunteren meeting 2004, oral presentation	Apr 06, 2004
XXVII International Symposium of European Society of Nematologists, Rome, Italy, oral presentation	Jun 14, 2004
Meloidogyne werkgroep, Bergschenhoek, De Ruiter Seeds, oral presentation	Feb 15, 2005
XII International Congress on Molecular Plant-Microbe Interactions, Cancun, Mexico, poster presentation	Jul, 2005
► **IAB interview**	Jun 04, 2004
► **Excursions**	
Subtotal Scientific Exposure	*22.3 credits**

3) In-Depth Studies	date
► **EPS courses or other PhD courses**	
AFLP markers in plant systematics and breeding	May 16-18, 2001
EPS Summerschool Functional genomics, Utrecht University	Aug 25-28, 2003
Guide to digital scientific artwork	Apr 20-21, 2006
► **Journal club**	
PhD student discussion group/Literature discussion group Nematology	Mar 2001- May 2006
► **Individual research training**	
Subtotal In-Depth Studies	*5.7 credits**

4) Personal development	date
► **Skill training courses**	
Begeleiden Inleiding Omgevingswetenschappen	Jul 17-18, 2002
Scientific Writing	Sep 22-Nov 17, 2003
Afstudeervak begeleiden en organiseren	Jul 12-13, 2005
Career Orientation	Nov 2006
► **Organisation of PhD students day, course or conference**	
► **Membership of Board, Committee or PhD council**	
Member of PhD student council	2003-2004
Subtotal Personal Development	*6.1 credits**

TOTAL NUMBER OF CREDIT POINTS*	**37.1**

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 credits

** A credit represents a normative study load of 28 hours of study*

This research was supported by European Union grant DREAM (QLRT-1999-1462)

Cover picture made by Erik Slootweg

Drukker: Grafisch bedrijf Ponsen & Looijen b.v., Wageningen