Molecular genetic analysis of the pathogenicity of the potato cyst nematode *Globodera rostochiensis*

Ling Qin

Promotor:
Prof. dr. ir. J. Bakker
Hoogleraar in de Nematologie,
i.h.b. Fysiologie en Moleculaire Ecologie van Nematoden
Wageningen Universiteit

Co-promotoren:
Dr. ir. J. Helder
Universitair docent bij het Laboratorium voor Nematologie
Wageningen Universiteit

Dr. ir. G. Smant
Universitair docent bij het Laboratorium voor Nematologie
Wageningen Universiteit

Promotie Commissie:
Prof. dr. A.H.J. Bisseling, Wageningen Universiteit
Prof. dr. C. Mariani, Katholieke Universiteit Nijmegen
Prof. dr. G. Gheysen, Universiteit Gent
Dr. ir. F.P.M. Govers, Wageningen Universiteit

Ling Qin

Molecular genetic analysis of the pathogenicity of the potato cyst nematode *Globodera rostochiensis*

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. dr. ir. L. Speelman,
in het openbaar te verdedigen
op vrijdag 5 October 2001
des morgens om 11.00 uur in de Aula.

Qin, Ling

Molecular genetic analysis of the pathogenicity of the potato cyst nematode *Globodera rostochiensis*

CIP-DATA Koninklijke Bibliotheek, Den Haag

Thesis Wageningen – with references
ISBN 90-5808-452-3
Subject headings: cyst nematodes/cDNA-AFLP/stylet secretions/plant-nematode interaction/expansin/RanBPM/functional genomics/nematode transformation/bioinformatics

Printed by: Ponsen & Looijen BV, Wageningen
Cover: processed image of a X-ray film exposed to an cDNA-AFLP gel - Ling Qin

NN08201, 3044

Stellingen/Propositions

1. Expansins, a class of proteins that unlock the network of plant cell wall polymers, have not evolved specifically in the land plant lineage. The plant parasitic nematode *Globodera rostochiensis* and - most likely - numerous other plant pathogens as well have discovered this 'trick' during evolution.
   - this thesis

2. The family of RanBPM (Ran-Binding Protein in microtubule organizing center) homologues identified from the dorsal gland of the potato cyst nematode might be co-responsible for shunting the M-phase during cycles of DNA endo-reduplication, an essential element in syncytium formation.
   - this thesis

3. No less than the entire genome sequence of at least one plant parasitic nematode is needed to understand the evolution of plant parasitism by nematodes, including the intriguing hypothesis of horizontal gene transfer of pathogenicity factors from bacteria to nematodes.

4. In the phylogenetic tree of molecular tools, cDNA-AFLP and SAGE (Serial Analysis of Gene Expression) would cluster closely together because both exploit the efficiency and robustness of three revolutionary molecular biology innovations: restriction enzyme cutting, PCR (polymerase chain reaction) and DNA sequencing.

5. Plant molecular biology and agricultural biotechnology industry would still be in their infancies if there hadn't been a common belief among granting agencies more than 30 years ago that research on the crown gall-inducing bacterium *Agrobacterium tumefaciens* would shed light on animal tumor pathogenesis.

6. There is a remarkable resemblance between Mandarin-Chinese and the computer language C++ (overload makes Chinese grammar easier and the function of C++ easier to use).

7. From the strikingly similar phenotype that most of the people in the Netherlands and China can't be separated from their bicycles, it can be safely concluded that Dutch and Chinese are highly related people.

8. 知之者不如　好者，好之者不如乐之者 --- 孔　子(Confucius)

9. No matter how esoteric a matter might be today, somebody, somewhere has spent years getting a doctorate in it.

10. Horizontal language transfer is much more common and far easier to prove than horizontal gene transfer.

11. Alle Chinezen kunnen goed ping-pongen.
    - de Volkskrant

12. Earth is the only heaven.

Stelling behorend bij het proefschrift getiteld: "Molecular genetic analysis of the pathogenicity of the potato cyst nematode *Globodera rostochiensis*", door Ling Qin.

Wageningen, 5 October 2001

For my family

CONTENTS

Chapter I

General introduction

Section 1.01 Intimate host-pathogen interactions

Evolution of communities of organisms has resulted in many kinds of interactions varying from peaceful, separate co-existence (= non-interaction) to symbiotic or parasitic relationships. Some pathogens have developed sophisticated mechanisms to interact with their hosts and start a durable relationship. Durable parasitic relationships are fascinating and scientifically interesting as they illustrate the maximal impact of co-evolution. Often these pathogens have evolved the ability to recognize individual compounds unintentionally released by their hosts. These signals will then trigger the production and / or the release of a rich array of pathogenicity factors from the parasites. The comprehension of the molecular events underlying the resulting intimate interactions will lead to the better understandings of the basic biology as well as development of new approaches to elegantly and specifically control harmful pathogens.

This Ph.D. thesis is about the potato cyst nematode, an obligatory parasite of a small range of solanaceous plants including potato, tomato and eggplant. This parasite has a major economical impact. The identification of signal molecules that are used by the nematode to parasitize its host (= pathogenicity factors) would give us handles to subtly block this interaction. In 1997, at the time this research was started, very little was known about the mechanisms underlying the interaction between the potato cyst nematode and its host. Hypothesizing that this interaction will probably involve a unique combination of common themes, two examples will be given of intimate pathogen - host interactions. These examples illustrate how such interactions can look like.

A remarkable example of the interaction between a microbe and an animal is the sophisticated adaptation of *Salmonella* bacteria to its host cells. As soon as *Salmonella* bacteria are in contact with host cells, the specialized Type III protein secretion system (a 'molecular syringe') and translocation machinery is activated. It is thought that the contact with host activates the whole process. The genes encoding this type III secretion system were probably acquired by *Salmonella* during evolution by horizontal gene transfer, which allowed this bacterium to parasitize eukaryotic hosts. Various bacterial effector proteins that do not contain a signal peptide for secretion are then delivered into host cells through this 'molecular syringe'. These effector proteins interact with host proteins and modify the host cellular function to the benefit of the bacteria. As a result the actin cytoskeleton of the infected cells is dramatically rearranged to promote the internalization of the bacteria. In addition, host cell transcription factors that stimulate the production of pro-inflammatory cytokines are activated, leading to disease symptoms [1;2].

A well-characterized example of co-evolution of plants and microbes is the gall formation induced on plants by the crown gall bacterium *Agrobacterium tumefaciens*.

* An adapted version of this chapter and chapter VII is published in " **Biology of Plant-Microbe Interactions, Volume 3**"- 2001 IC-MPMI Congress Proceedings. The identity and function of cyst nematode-secreted proteins in pathogenesis. **Qin, L.**, Smant,G., Bakker,J. and Helder,J.

During the long-standing association with its host, *Agrobacterium* has developed sophisticated ways to manipulate its host plant. Once the *Agrobacterium* senses acetosyringones, which are released from the wounding tissues, it will attach to the site of wounding. Subsequently, a large panel of effector proteins is activated by acetosyringones. As a result, single-stranded DNA is synthesized from the Ti (Tumor-inducing)-plasmid, transported into the host nucleus and integrated into the host genome. The integrated genes encode proteins which mediate the production of opines (opines can only be metabolized by *Agrobacterium spp*) and the plant hormones auxin and cytokinin. In this way, the host metabolic machinery is drastically modified to meet the needs of the bacteria, finally resulting in the crown gall formation [3].

Section 1.02 The life history of a cyst nematode – *Globodera rostochiensis*

(a) Introduction

Nematodes, members of the phylum Nematoda, are present in almost all the imaginable habitats on earth. Most nematodes are bacteriophagous, mycetophagous, or live as saprophytes on dead organic material. Relatively few nematode species developed into parasites of plants or animals. Plant parasitic nematodes have different feeding strategies. Most of them are either ectoparasitic or migratory endoparasitic, which means they will move around to find new food sources and generally do not engage in an intimate relationship with their host plants. On the other hand, sedentary endoparasitic nematodes have evolved complex relationships with their host plants and they complete the whole life cycle by extracting nutrients solely from a fixed feeding site inside the host plant root. The family *Heteroderidae* (order *Tylenchida*) includes cyst (*Heterodera* and *Globodera* spp.) and root knot (*Meloidogyne* spp.) nematodes. Because of the damage it causes to a wide range of crops, this family of sedentary endoparasitic nematodes is highly relevant [4].

This thesis focuses on the potato cyst nematode *Globodera rostochiensis*, an obligatory endoparasite of a small range of solanaceous plants. The potato cyst nematode originates from the Andean region in South America [5] and was probably introduced into Western Europe in the second half of the 19th century. The species name '*rostochiensis*' is derived from the name of a city in Germany, Rostock. This name could falsely suggest that this potato cyst nematode originates from Germany: the name only reflects its first discovery in North Western Europe. During the long-standing association with its solanaceous hosts, the potato cyst nematode has developed the ability to use host plant signals to synchronize its own life cycle with that of its host. Without suitable host plants in its surroundings, the potato cyst nematode can remain viable and stay in dormancy in the soil for years. The almost complete dependence on root diffusate of host plants for hatching is remarkable.

(b) Life cycle of the potato cyst nematode

The potato cyst nematode has five different developmental stages and – except of the adult stage – the end of each of these stages is marked by a molt. The first-stage juvenile (J1) molts within the egg resulting in a second-stage juvenile (J2). Initially, this J2 is in diapause, and this dormant J2 will become susceptible to hatching stimuli only after 8-12 months incubation at relatively low temperatures (≈ 4°C). After diapause, hatching of J2 is greatly stimulated by exudates from the roots of host plants. These exudates (or diffusates) are complex mixtures of compounds released by

host plant roots and among the active components are the so-called eclepins. In case of the soybean cyst nematode, these substances stimulate at low concentrations, 10^{-11} to 10^{-12} g ml^{-1}, the hatching of J2 from the eggs [6].

The freshly hatched J2's invade the root just behind the apex, in the elongation or the differentiation zone. After root penetration, they migrate intracellularly through the cortex in the direction of the vascular cylinder. Pre-parasitic J2s possess a protrusible, hollow spear called a stylet with which the nematode pierces cell walls. The migration is not only facilitated by stylet thrusts but also by cell wall-degrading enzymes that are secreted by the invading nematode. These enzymes are produced in the two subventral glands. Upon arrival in the inner cortex, the nematode selects a cortical cell, the initial syncytial cell (ISC). This cell will be the starting point of the nematode feeding site. Once the wall of the ISC has been perforated, the stylet tip stays protruded for 6-8 hours, which is called the preparation time The cortical cell subsequently re-differentiates from a highly vacuolated, differentiated cell into a metabolically highly active cell. Probably, this de- and re-differentiation process is triggered by a complex mixture of secretions produced and released by the nematode. These secretions are thought to be produced mainly by the dorsal gland. Syncytium initiation is followed by a combination of feeding and syncytium proliferation. Feeding occurs in repeated cycles, each consisting of three distinct phases [7;8]:

1. nutrients are withdrawn from the ISC through a feeding tube, by a continuous rapid pumping of the metacorpal bulb
2. stylet retraction and reinsertion into the syncytium through a feeding plug
3. a continuous forward movement of secretory granules occurs, especially from the dorsal gland; the stylet tip stays inserted in the ISC; a new feeding tube is formed

Syncytium formation coincides with an increase in the density of the cytoplasm, the amount of endoplasmic reticulum and the volume of nucleus. On the other hand, the volume of the central vacuole decreases and new small cytoplasmic vacuoles are formed. Furthermore, the ISC expands toward the vascular bundle by fusing with neighboring cells through partial cell wall dissolution. The developing syncytium extends longitudinally along the vascular cylinder and cell wall ingrowths are formed adjacent the xylem elements. The ingrowths facilitate the nutrient uptake from the xylem into the syncytium [9;10].

As a result of feeding, the nematode changes. Its body wall muscles degenerate and the nematodes become sedentary[11]. Sex differentiation takes place at the parasitic J2 stage. The availability of food is the trigger that determines the sex of a developing potato cyst nematode: a shortage of food will result in the formation of males whereas favorable nutritional conditions will lead to the appearance of females [12]. J2 males and females develop into J3s, J4s and finally they reach adulthood. In case of females, growth coincides with a gradual swelling of her body. Male juveniles initially become sausage-shaped. After the J3 stage, they stop feeding, and adult males are vermiform and mobile again. They leave the root and search for adult females to mate. The bodies of females swell extensively and become globular in shape. At some stage, they burst out of the root epidermis, and become visible as white spherical objects. Potato cyst nematodes are obligatory sexual and after fertilization eggs develop within the female body. At some point, the female dies and the remaining cuticle forms a protective layer enclosing up to several hundred eggs. At the end of the growing season, cysts mainly contain J2's in primary diapause. These cysts remain viable for

many years in the soil, and normally after one winter period, they will be able to hatch when host plants are available [13] (Fig. 1, the life cycle of the potato cyst nematode).

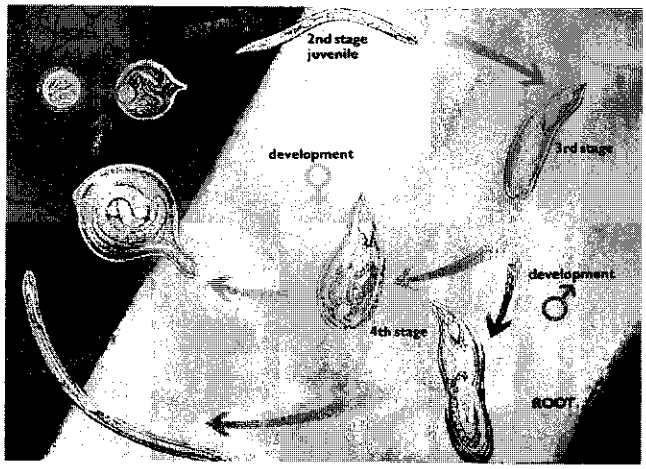


Fig. 1. The life cycle of the potato cyst nematode

Section 1.03 Nematode secretions from the esophageal glands

(a) The nematode esophageal glands

Nematode secretions originate from a range of body structures including the cuticle, the amphids and the esophageal gland cells. Secretions present on the cuticle surface may serve to mask the nematode from its host or to anchor the nematode at its feeding site [14]. Amphidial secretions may form a feeding plug that surrounds the point at which the nematode pierces the syncytium with its stylet [15].

As compared to nematodes that feed on bacteria, plant parasitic nematodes have relatively enormous esophageal glands. Cyst nematodes have one dorsal and two subventral esophageal secretory glands. (Fig. 2, [16]). Each gland is a single cell with a cytoplasmic extension that terminates in an ampulla that serves as a reservoir for secretory granules. The ampullae are connected to the lumen of the esophagus, in this way the gland secretions can be released into the esophageal lumen. The dorsal gland ends just behind the stylet whereas the subventral gland cells are connected to the lumen of the esophagus in the median bulb just behind the pump chamber. The secretions are eventually injected into plant tissue through the stylet.

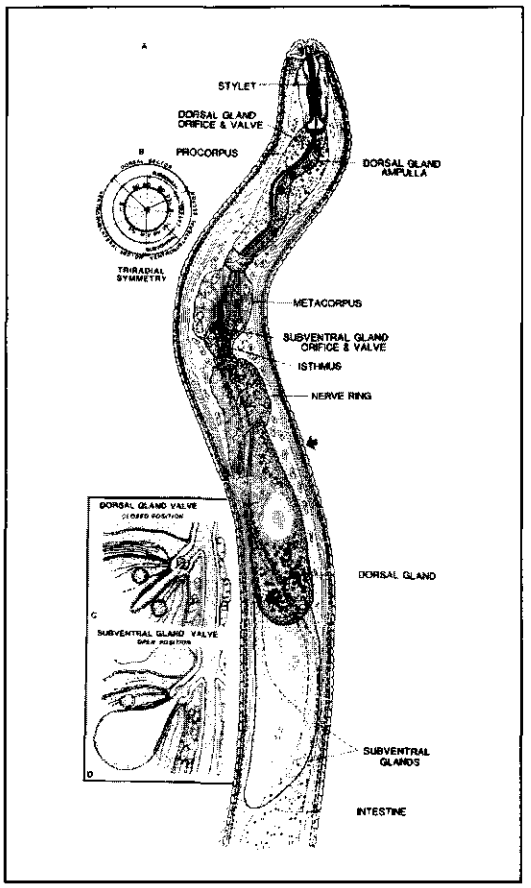


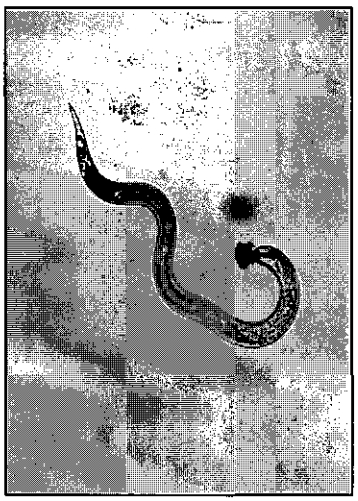
Fig. 2. The anterior part of a second stage infective juvenile of a cyst nematode (From Endo, 1984, reference 16).

(b) Different functions of the subventral and dorsal esophageal glands

The morphologies of the subventral and dorsal esophageal glands are highly different during the parasitic process, which could point at different functions. The effect of potato root diffusate (PRD) on the activities of the esophageal glands is different. In case of *G. rostochiensis,* it was observed that secretory granule accumulation in subventral glands was solely a response to hydration. Moreover, hydration induced an increase in size of the gland cell nucleolus as well as an accumulation of granules in dorsal glands. Contrary to the subventral gland, the dorsal gland was further activated by exposure to PRD [17;18]. During migration through the plant root, the two subventral secretory glands are large and packed with secretory granules. The onset of parasitism, feeding-site initiation, is accompanied by a rapid decrease in size of the two subventral glands, whereas the dorsal gland shows a remarkable enlargement.

These changes were observed for the interactions *Heterodera glycines* – soybean [19], *Meloidogyne incognita* - tomato [20] and *Heterodera schachtii* - *Brassica rapa* [21]. These morphological changes suggest an important role of subventral gland secretions during migration. To date, a small number of subventral gland proteins have been identified: β-1,4-endoglucanases from cyst and root knot nematodes [22], a pectate lyase from the potato cyst nematode [23] and a cellulose-binding protein from the root knot nematode *M. incognita* [24]. The nature of these proteins suggests that the subventral gland proteins indeed play a role during migration through the plant root. Feeding site initiation is accompanied by an activation of the dorsal gland and, as such, dorsal gland proteins could be functional during the onset of parasitism and are likely to be important for the induction and maintenance of the syncytium. Despite considerable efforts, no dorsal gland products have been identified yet.

(c) Collection of the nematode secretions



Two chemical reagents are commonly used to collect secretions from pre-parasitic juveniles *in vitro*: a serotonin analogue called 5-methyoxy-N, N-dimethyltryptamine-hydrogen-oxalate (DMT) and resorcinol. The former is used for cyst nematode [25], the latter for root knot nematodes [26]. The biological relevance of secretions produced upon exposure to these chemicals is difficult to assess. If available, a natural secretion-inducing substance would be preferred. Potato root diffusate (PRD) has long been used as the hatching stimulus of the potato cyst nematode [27]. Recently, it was shown that exposure to PRD induced the release of secretory proteins by freshly-hatched J2 [28]. PRD-induced secretions probably closely resemble the set of proteins that is released during the onset of plant parasitism.

Fig. 3. Nematode secretions were visible at the tip of the stylet of a second stage infective juvenile of the potato cyst nematode *G. rostochiensis*.

Section 1.04 What kind of pathogenicity factors can be expected from the potato cyst nematode

Since there are many similarities among different plant-microbe interaction systems, it is very informative to consider in parallel the strategies used by other pathogens. During colonization and migration in plants, the cell wall constitutes a major physical barrier to plant pathogens and parasites. To reach or to access a feeding site, bacteria, fungi and nematodes secrete cell wall-degrading enzymes. A second class of genes expressed in pathogens and parasites limits the lethal effects of anti-microbial compounds released by the plants. For instance, pathogens may produce ATP-binding-cassette-transporters to defend themselves against anti-microbial compounds produced by the host [29], or peptide methionine sulfoxide reductase to repair oxidative damage

caused to its proteins by active oxygen released by the host plant [30]. Once an interaction has been established, pathogens and parasites may allocate plant nutrients by inducing changes in the translocation of water and nutrients in the plant. However, little is known of the signals that mediate the changes induced in plants as an essential part of the compatible plant-nematode interaction.

In case of the potato cyst nematode, these three different types of pathogenicity-related proteins are likely to be important in plant parasitism as well. This can be illustrated by the fact that two kinds of cell wall-degrading enzymes, β-1,4-endoglucanases and pectate lyase, are produced in the subventral gland of *G. rostochiensis*. These enzymes most probably facilitate the migration of the pre-parasitic juvenile through the plant root [22;23]. Secondly, the potato cyst nematode was shown to secrete superoxide dismutase upon exposure to DMT [31]. Thirdly, to finally extract nutrients from its host, sedentary nematodes secrete molecules that force plant cells to re-differentiate into a feeding site. This process is marked by a reactivation of the cell cycle. DNA synthesis and progression through the G2 phase are essential for syncytium establishment [32;33]. Typically, cells in a feeding site have large nuclei, and show signs of enhanced metabolic activity. The nematode secretions might contain inhibitors that specifically block the activity of M-phase promoting factors and/or induce unknown S-phase-related protein kinases. Recently, a protein fraction in nematode secretions smaller than 3 kDa was found to contain mitogenic activity on plant protoplasts and mammalian T-cell lymphocytes [34]. In contrast to the first two aspects of nematode pathogenicity, at present, no genes have been identified that might encode the putative factor(s) for feeding site induction and maintenance.

Section 1.05 Strategies to clone nematode secretory genes

To identify the nematode secretory genes, the most straightforward approach would be to collect the nematode secretions and analyze the components directly. Unfortunately, the collection of pure nematode secretions essentially free from other microbial contaminants for protein analysis is a formidable challenge. Infective juveniles are very small and - because of their obligatory plant parasitic nature and their relatively long life cycle - they can not be reared on a large scale in monoxenic cultures. Several alternative approaches can be used to circumvent this problem. The first approach presented here is based on the identification of the protein. On the basis of partial amino acid sequences, the corresponding gene can be identified. The other approaches (2-5) in principle directly result in the cloning of the gene of interest. Approaches 4 and 5 differ from the other approaches in that they depend on homologies.

Approach 1. The use of monoclonal antibodies to localize and purify esophageal gland proteins.
Assumption I: *pathogenicity factors are synthesized in the esophageal glands*

The inability to collect sufficient nematode secretions for direct protein analysis prompted a number of groups to raise monoclonal antibodies (mAb's) against nematode secretions [35]. Actually, this approach resulted in the identification the first nematode gene for parasitism. MAb's were used to immunopurify proteins for sequence analysis and the first subventral gland secretory proteins, β-1,4-

endoglucanases, were identified in this way [22;36]. In attempts to identify dorsal gland proteins from cyst or root knot nematodes, several monoclonal antibodies (mAbs) have been raised that specifically recognized the dorsal glands of cyst and root knot nematodes. In 1988, a mAb specific to the dorsal gland and duct of second stage juveniles from the soybean cyst nematode *H. glycines* was identified [37]. Later, two other mAbs specifically bound to the dorsal glands of *H. glycines* and the root knot nematode *M. incognita* respectively were raised [25;38]. However, the corresponding genes could not be identified. The mAb approach is time-consuming and appeared to be biased towards immuno-dominant proteins.

Approach 2. Using the presence of a signal peptide for secretion as a criterion to select for putative pathogenicity factors.
Assumption II: *pathogenicity factors are preceded by a signal peptide for secretion.*

Pathogenicity factors are most likely to be secreted outside the nematode. Nematode cDNAs can be cloned into a specially designed yeast vector and transformed into suitable yeast strains [39]. This yeast signal peptide selection system in principle allows for an efficient selection of (nematode) cDNAs that encode a signal peptide for secretion (provided it is recognized by yeast) from a large number of yeast clones. This approach may result in hundreds of positive clones, so it should be used in combination with other selection criteria. If it is combined with *e.g.* a gland-specific library, in theory, it could be efficient. Several putative pathogenicity factors have been identified in this way from the root knot nematode [40].

Approach 3. The use of mRNA fingerprinting (cDNA-AFLP) as a means to identify infective stage-specific genes
Assumption III: *pathogenicity factors are transcribed solely or predominantly in the infective stage(s) of the pathogen.*

Biological responses and developmental processes are precisely controlled at the level of gene expression. Information on the temporal and spatial regulation of gene expression often sheds light on the potential function of a particular gene. Most pathogenicity factors are not expressed constitutively, but in a highly coordinated way. Powerful techniques, which are capable to reveal the differential gene expression patterns of a large number of genes reliably, would enable us to identify many pathogenicity-related genes. At the start of this thesis research, no such technology has been successfully been applied in plant nematology research.

cDNA-AFLP is a novel RNA fingerprinting technique to display differentially expressed genes. cDNAs are digested by two restriction enzymes and oligonucleotide adapters are ligated to the resulting restriction fragments to generate template DNA for PCR. PCR primers complementary to the adapter sequences with additional selective nucleotides at the 3' end allow for the amplification of a limited number of cDNA fragments. Unlike classical differential display methods that make use of small random primers [41], relatively high annealing temperatures can be used and, hence, cDNA-AFLP is more stringent and reproducible. Furthermore, cDNA-AFLP can distinguish between highly homologous genes from individual gene families. In this Ph.D. project, this relatively cheap and robust technique has been applied to monitor the gene expression of five distinct life stages of the potato cyst nematode in order to identify genes, which are important for the nematode parasitism.

Approach 4. Random sequencing of cDNAs as a way to identify putative pathogenicity factors
Limitation: *this approach depends on the availability of (partially) characterized homologues.*

Expressed sequence tags (EST) are single pass sequences of cDNA clones selected randomly from a library [42]. When high throughput sequencing facility is available, a large number of nematode genes can be discovered rather quickly. However, as a technology by itself, the identification of putative pathogenicity factors relies strongly on homology searches. To study the function of genes showing no homology to functionally characterized sequences in the database, it is essential to combine the EST approach with other techniques.

Approach 5. The use of degenerate PCR primers for the identification of putative pathogenicity factors.
Limitation: *this approach depends on the availability of (partially) characterized homologues.*

When putative pathogenicity factors have been identified in other plant pathogens or in other nematode species, the homologous genes in the nematode of interest in principle can be cloned by using degenerate primers based on conserved regions. This strategy was successfully used to clone cellulase genes from the root knot nematode [43]. However - by definition - this approach is limited to the identification of homologues of known genes.

Section 1.06 How to assess the relative importance of individual pathogenicity factors?

The techniques summarized above that are used to clone nematode parasitism genes have one thing in common – they are all reverse genetics approaches. Putative pathogenicity factors might be identified by these approaches, but no definitive function can be assigned to them. To evaluate the relative importance of a certain compound secreted during the infection process, mutant nematodes that are unable to produce and/or secrete this factor would be useful. Because no such mutants are available from the classical genetic approach, a protocol for transforming obligatory plant parasitic nematodes is urgently needed. One of the prerequisites for this transgenic approach being successful would be the availability of (constitutive) promoters that are properly recognized by the host organism. At present, several 5'-upstream regions of locally expressed plant-parasitic nematode genes have been identified [26;44;45]. However, none of them was shown to include a functional promoter region. Further effort is thus needed to isolate functional promoter sequences.

Section 1.07 Outline of the thesis

The goal of this PhD research is to identify pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*. The identification of such factors will enable us to engineer new control strategies to manage this parasite.

- For this purpose, a novel efficient cDNA-AFLP-based strategy was designed and successfully performed to identify pathogenicity-related genes (Chapter 2).
- Subsequently, a computer program (GenEST) was developed to link cDNA-AFLP data with cDNA sequence data (including ESTs) to discover such genes (Chapter 3).
- A functional expansin gene (Chapter 4) and a RanBPM-like gene family (Chapter 5) were identified, which were expressed specifically in the nematode secretory glands and up-regulated in the infective stage, indicating their involvement in plant parasitism.
- Furthermore, a promoter from the potato cyst nematode was cloned and was shown to be functional in *Caenorhabditis elegans*. This promoter region could be a valuable tool to drive gene expression in transgenic plant-parasitic nematodes. Probably, the importance of newly identified putative pathogenicity factors can only be assessed if we are able to generate mutants unable to produce these pathogenicity factors (Chapter 6).

Finally, the results are summarized and discussed in chapter 7.

Reference List

1. Galan JE, Bliska JB: **Cross-talk between bacterial pathogens and their host cells.** *Ann.Rev.Cell Dev.Biol.* 1996, **12:**221-255.
2. Galan JE: **Interactions of Salmonella with host cells: encounters of the closest kind.** *Proc.Natl.Acad.Sci.U.S.A* 1998, **95:**14006-14008.
3. Zupan J, Muth TR, Draper O, Zambryski P: **The transfer of DNA from agrobacterium tumefaciens into plants: a feast of fundamental insights.** *Plant J.* 2000, **23:**11-28.
4. Sijmons PC, Atkinson HJ, Wyss U: **Parasitic strategies of root nematodes and associated host cell responses.** *Annu.Rev.Phytopathol.* 1994, **32:**235-259.
5. Evans K, Stone AR: **A review of the distribution and biology of the potato cyst-nematodes Globodera rostochiensis and G. pallida.** *PANS* 1977, **23:**178-189.
6. Masamune T, Anetai M, Takasugi M, Katsui N: **Isolation of a natural hatching stimulus, glycinoeclepin A, for the soybean cyst nematode.** *Nature, UK* 1982, **297:**495-496.
7. Steinbach P: **Untersuchungen uber das Verhalten von Larven des Kartoffelzystenalchens (*Heterodera rostochiensis* Wollenweber, 1923) an und in Wurzeln der Wirtpflanze *Lycopersicon esculentum* Mill. II: Das Eindringen von Kartoffelnematoden in die Wirtswurtzel.** *Biol.Zbl.* 1972, **91:**743-756.
8. Steinbach P: **Untersuchungen uber das Verhalten von Larven des Kartoffelzystenalchens (*Heterodera rostochiensis* Wollenweber, 1923) an und in Wurzeln der Wirtspflanze *Lycopersicon esculentum* Mill. III. Die Nahrungsafname von Kartoffelnematodenlarven.** *Biol.Zbl.* 1973, **92:**563-582.
9. Jones MGK, Northcote DH: **Nematode-induced syncytium - a multinucleate transfer cell.** *J.Cell Sci.* 1972, **10:**789-809.
10. Grundler FMW, Sobczak M, Golinowski W: **Formation of wall openings in root cells of Arabidopsis thaliana following infection by the plant-parasitic nematode Heterodera schachtii.** *European Journal Of Plant Pathology* 1998, **104:**545-551.
11. Shepherd AM, Clark SA: **Cuticle structure and 'cement' formation at the anterior end of female cyst-nematodes of the genera Heterodera and Globodera (Heteroderidae: Tylenchida).** *Nematologica* 1978, **24:**201-208.

12. Mugniery D, Fayet G: **Sex determination in Globodera rostochiensis Woll. and influence of infestation levels on the penetration, development and sex of the nematode**. *Revue de Nematologie* 1984, **7:**233-238.
13. Perry RN: **Dormancy and hatching of nematode eggs**. *Parasitology Today* 1989, **5:**377-383.
14. Jones JT, Perry RN, Johnston MRL: **Changes in the ultrastructure of the cuticle of the potato cyst nematode, Globodera rostochiensis, during development and infection**. *Fundam.Appl.Nematol.* 1993, **16:**433-445.
15. Endo BY: **Feeding plug formation in soybean roots infected with the soybean cyst nematode**. *Phytopathology* 1978, **68:**1022-1031.
16. Endo BY: **Ultrastructure of the esophagus of larvae of the soybean cyst nematode, *Heterodera glycines***. *Proc.Helminth.Soc.Wash.* 1984, **51:**1-24.
17. Perry RN, Zunke U, Wyss U: **Observations on the response of the dorsal and subventral oesophageal glands of *Globodera rostochiensis* to hatching stimulation**. *Rev.Nematol.* 1989, **12:**91-96.
18. Atkinson HJ, Taylor JD, Fowler M: **Changes in the second stage juveniles of *Globodera rostochiensis* prior to hatching in response to potato root diffusate**. *Ann.Appl.Biol.* 1987, **110:**105-114.
19. Atkinson HJ, Harris PD: **Changes in nematode antigens recognized by monoclonal antibodies during early infections of soya beans with the cyst nematode *Heterodera glycines***. *Parasitology* 1989, **98:**479-487.
20. Hussey RS, Mims CW: **Ultrastructure of esophageal glands and their secretory granules in the root-knot nematode Meloidogyne incognita**. *Protoplasma* 1990, **156:**9-18.
21. Wyss U, Grundler FMW: **Feeding behaviour of sedentary plant parasitic nematodes**. *Neth.J.Plant Pathol.* 1992, **98, Supplement 2:**165-173.
22. Smant G, Stokkermans JP, Yan Y, de Boer J, Baum TJ, Wang X, Hussey RS, Gommers FJ, Henrissat B, Davis EL, Helder J, Schots A, Bakker J: **Endogenous cellulases in animals: isolation of beta-1, 4-endoglucanase genes from two species of plant-parasitic cyst nematodes**. 1998, 95:4906-4911.
23. Popeijus H, Overmars H, Jones J, Blok V, Goverse A, Helder H, Schots A, Bakker J, Smant G: **Degradation of plant cell walls by a nematode**. *Nature London* 2000, **406:**36-37.
24. Ding X, Shields J, Allen R, Hussey RS: **A secretory cellulose-binding protein cDNA cloned from the root-knot nematode (Meloidogyne incognita)**. *Molecular Plant Microbe Interactions* 1998, **11:**952-959.
25. Goverse A, Davis EL, Hussey RS: **Monoclonal antibodies to the esophageal glands and stylet secretions of *Heterodera glycines***. *J.Nematol.* 1994, **26:**251-259.
26. Ray C, Abbott AG, Hussey RS: **Trans-splicing of a Meloidogyne incognita mRNA encoding a putative esophageal gland protein**. *Molecular and Biochemical Parasitology* 1994, **68:**93-101.
27. Perry RN, Feil JJ: **Observations on a novel hatching bioassay for *Globodera rostochiensis* using fluorescence microscopy**. *Rev.Nematol.* 1986, **9:**280-282.
28. Smant G, Goverse A, Stokkermans JG, De Boer JM, Pomp H, Zilverentant JF, Overmars HA, Helder J, Schots A, Bakker J: **Potato root diffusate induced secretion of soluble, basic proteins originating from the subventral esophageal glands of potato cyst nematodes**. *Phytopathology* 1997, **87:**839-845.
29. Urban M, Bhargava T, Hamer JE: **An ATP-driven efflux pump is a novel pathogenicity factor in rice blast disease**. *EMBO J.* 1999, **18:**512-521.
30. Hassouni ME, Chambost JP, Expert D, Van GF, Barras F: **The minimal gene set member msrA, encoding peptide methionine sulfoxide reductase, is a virulence determinant of the plant pathogen Erwinia chrysanthemi**. *Proc.Natl.Acad.Sci.U.S.A.* 1999, **96:**887-892.
31. Robertson L, Robertson WM, Jones JT: **Direct analysis of the secretions of the potato cyst nematode Globodera rostochiensis**. *Parasitology* 1999, **119 Part 2:**167-176.

32. Niebel A, Engler Jd, Hemerly A, Ferreira P, Inze D, Montagu MV, Gheysen G: **Induction of cdc2a and cyc1At expression in Arabidopsis thaliana during early phases of nematode-induced feeding cell formation.** *Plant Journal* 1996, **10:**1037-1043.
33. Almeida-Engler Jd, Vleesschauwer Vd, Burssens S, Celenza JLJ, Inze D, Montagu Mv, Engler G, Gheysen G: **Molecular markers and cell cycle inhibitors show the importance of cell cycle progression in nematode-induced galls and syncytia.** *Plant Cell* 1999, **11:**793-807.
34. Goverse A, Rouppe van der Voort J, Rouppe van-der Voort C, Kavelaars A, Smant G, Schots A, Bakker J, Helder J: **Naturally induced secretions of the potato cyst nematode co-stimulate the proliferation of both tobacco leaf protoplasts and human peripheral blood mononuclear cells.** *Molecular Plant Microbe Interactions* 1999, **12:**872-881.
35. Hussey RS: **Monoclonal antibodies to secretory granules in esophageal glands of *Meloidogyne* species.** *J.Nematol.* 1989, **21:**392-398.
36. Yan Y, Smant G, Stokkermans J, Qin L, Helder J, Baum T, Schots A, Davis E: **Genomic organization of four beta-1,4-endoglucanase genes in plant-parasitic cyst nematodes and its evolutionary implications.** *Gene* 1998, **220:**61-70.
37. Atkinson HJ, Harris PD, Halk EJ, Novitski C, Leighton-Sands J, Nolan P, Fox PC, Sands JL: **Monoclonal antibodies to the soyabean cyst nematode, *Heterodera glycines*.** *Ann.Appl.Biol.* 1988, **112:**459-469.
38. Davis EL, Allen R, Hussey RS: **Developmental expression of esophageal gland antigens and their detection in stylet secretions of *Meloidogyne incognita*.** *Fundam.Appl.Nematol.* 1994, **17:**255-262.
39. Klein RD, Gu Q, Goddard A, Rosenthal A: **Selection for genes encoding secreted proteins and receptors.** *Proc Natl Acad Sci U S A* 1996, **93:**7108-7113.
40. Wang XH, Allen R, Ding XF, Goellner M, Maier T, De Boer JM, Baum TJ, Hussey RS, Davis EL: **Signal peptide-selection of cDNA cloned directly from the esophageal gland cells of the soybean cyst nematode Heterodera glycines.** *Molecular Plant Microbe Interactions* 2001, **14:**536-544.
41. Liang P, Pardee AB: **Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction [see comments].** *Science* 1992, **257:** 967-971.
42. Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olde B, Moreno RF, et a: **Complementary DNA sequencing: expressed sequence tags and human genome project.** *Science* 1991, **252:**1651-1656.
43. Rosso MN, Favery B, Piotte C, Arthaud L, de BJ, Hussey RS, Bakker J, Baum TJ, Abad P: **Isolation of a cDNA encoding a beta-1,4-endoglucanase in the root-knot nematode Meloidogyne incognita and expression analysis during plant parasitism.** *Mol.Plant Microbe.Interact.* 1999, **12:**585-591.
44. De Giorgi C, De Luca F, Di Vito M, Lamberti F: **Modulation of expression at the level of splicing of cut-1 RNA in the infective second-stage juvenile of the plant parasitic nematode Meloidogyne artiellia.** *Molecular & General Genetics* 1997, **253:**589-598.
45. Novitski CE, Brown S, Chen R, Corner AS, Atkinson HJ, McPherson MJ: **Major sperm protein genes from Globodera rostochiensis.** *J.Nematol.* 1993, **25:**548-554.

Chapter II

An efficient cDNA-AFLP-based strategy for the identification of putative pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*

Qin,L., Overmars,H., Helder,J., Popeijus,H., Rouppe Van Der Voort,J.N.A.M., Groenink,W., van Koert,P., Schots,A., Bakker,J. and Smant,G.
An efficient cDNA-AFLP-based strategy for the identification of putative pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*.
Molecular Plant-Microbe Interactions 2000.13:830-836.

An efficient cDNA-AFLP-based strategy for the identification of putative pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*

ABSTRACT:
A new strategy has been designed to identify putative pathogenicity factors from the dorsal or subventral esophageal glands of the potato cyst nematode *Globodera rostochiensis*. Three independent criteria were used for selection. Firstly, genes of interest should predominantly be expressed in infective second stage juveniles, and not or to a far lower extent in younger developmental stages. For this, gene expression profiles from five different developmental stages were generated using cDNA-AFLP. Secondly, the mRNA corresponding to such a putative pathogenicity factor should predominantly be present in the esophageal glands of pre-parasitic juveniles. This was checked by *in situ* hybridization. As a third criterion, these proteinaceous factors should be preceded by a signal peptide for secretion. Expression profiles of more than 4,000 genes were generated and three up-regulated, dorsal gland-specific proteins preceded by signal peptide for secretion were identified. No dorsal glands genes have been cloned before from plant parasitic nematodes. The partial sequence of these three factors, A4, A18 and A41, showed no significant homology to any known gene. Their presence in the dorsal glands of infective juveniles suggests that these proteins could be involved in feeding cell initiation, and not in migration in the plant root or in protection against plant defense responses. Finally, the applicability of this new strategy in other plant-microbe interactions is discussed.

Nucleotide sequence data reported in this paper are available in the EMBL, GeneBank and DDJB databases under the accession numbers AJ251756, AJ251757 and AJ251758.

INTRODUCTION:

Co-evolution of plants and microbes may result in intimate and durable interactions. A well-characterized example of co-evolution of plants and microbes is the symbiosis of legumes and Rhizobia. Symbiosis arises from an extensive exchange of molecular signals between two partners. In contrast to Rhizobia, biotrophic bacteria, fungi, and nematodes often flourish at the expense of the hosting plant and release pathogenicity factors in order to reach a feeding site, to counteract the plant defense responses and to extract food.

In case of sedentary plant parasitic nematodes, migration through plant tissues is facilitated by the secretion of β-1,4-endoglucanases (Smant et al. 1998; Rosso et al. 1999). Apart from mechanical barriers, nematodes have to withstand the defense response of a host plant. Antioxidant enzymes in animal parasitic nematodes are pivotal in the interaction with their host and recently, the potato cyst nematode *Globodera rostochiensis* was shown to secrete superoxide dismutase upon exposure to 5-methoxy-N,N-dimethyltryptamine-hydrogen-oxalate (Robertson et al. 1999). To finally extract nutrients from its host, sedentary nematodes secrete molecules that force plant cells to re-differentiate into a feeding site. This process is marked by a reactivation of the cell cycle (Niebel et al. 1996). Typically, cells in a feeding site have large nuclei, and show signs of enhanced metabolic activity. Nematode secretions have long been implicated as the causal agents of the physiological and molecular changes in feeding sites induced in plants (reviewed by Hussey and Williamson 1996). Recently, a protein fraction in nematode secretions smaller than 3 kilo Dalton was found to contain mitogenic activity on plant protoplasts and mammalian T-cell lymphocytes (Goverse et al. 1999). At present it is unclear whether this peptide fraction contains the causal agent(s) of feeding site induction.

The collection of pure nematode secretions essentially free from other microbial contaminants for protein analysis is a formidable challenge. Infective juveniles are very small and - because of their obligatory plant parasitic nature and their relatively long life cycle – it is difficult to rear them on a large scale in monoxenic cultures. The inability to collect sufficient nematode secretions for proper analysis prompted many groups in the past to develop polyclonal and monoclonal antibodies against nematode secretions. Some of the antisera were used to immunopurify proteins for sequence analysis (Ray et al. 1994; Smant et al. 1998). Although this laborious procedure has resulted in the identification of a pathogenicity factor in secretions of cyst nematodes (Smant et al. 1998; Yan et al. 1998), a more efficient method is desirable.

cDNA-AFLP is a novel RNA fingerprinting technique to display differentially expressed genes (Bachem et al. 1996). cDNAs are digested by two restriction enzymes and oligonucleotide adapters are ligated to the resulting restriction fragments to generate template DNA for polymerase chain reaction (PCR). PCR primers complementary to the adapter sequences with additional selective nucleotides at the 3' end allow for the amplification of a limited number of cDNA fragments. Unlike differential display methods that make use of small random primers (Liang and Pardee 1992), relatively high annealing temperatures can be used and, hence, cDNA-AFLP is more stringent and reproducible. In contrast to hybridization-based techniques, such as cDNA microarrays, cDNA-AFLP can distinguish between highly homologous genes from individual gene families. In addition, cDNA-AFLP does not need any pre-existing sequence information, which makes it an excellent tool to identify novel genes.

To select genes encoding pathogenicity factors that are present in nematode secretions from a pool of thousands of amplified cDNA fragments, the following criteria were applied. As a first selection criterion, we assumed that cDNA fragments encoding pathogenicity factors are up-regulated in an RNA pool isolated from freshly hatched pre-parasitic J2 just prior the invasion of a host plant as compared to younger developmental stages. Secondly, it is assumed that pathogenicity factors are produced in the dorsal or in the subventral oesophageal glands of the nematode. This latter criterion is justified by numerous microscopic observations ((Wyss and Zunke, U.1986); Wyss 1992). Thirdly, we assume that pathogenicity related mRNAs encode a protein which starts with a signal peptide for secretion. This seems reasonable because pathogenicity factors act outside the nematode. The first pathogenicity factor cloned from cyst nematodes, a β-1,4-endoglucanase, meets all of these criteria (Smant et al. 1998).

In this paper, the RNA fingerprints of approx. 4,000 genes in five distinct developmental stages from the potato cyst nematode *G. rostochiensis* were compared using cDNA-AFLP. The life cycle of potato cyst nematodes allowed for the isolation of highly synchronized, distinct phases. Amplified cDNA fragments up-regulated in infective second stage juveniles were sequenced and the 5' ends of the cDNAs were checked for the presence of an open reading frame that starts with a predicted signal sequence for secretion. If such a signal was present, the exact location of this mRNA species was established by in situ hybridization. This highly efficient selection procedure resulted in the identification of three dorsal gland specific genes. It is noted that - despite considerable efforts - to date no proteins have been identified from the dorsal glands of sedentary nematodes, and the three dorsal glands protein encoding cDNAs reported here are the first ones to be cloned.

RESULTS

Validation of the quantitative relationships of transcript-derived fragments in cDNA-AFLP

The expression patterns of more than 4000 transcript-derived fragments (TDFs) from five successive developmental stages of the potato cyst nematode *G. rostochiensis* were monitored by cDNA-AFLP analysis. The highly synchronous inoculation of potato plants allowed for a comparison between the mRNA populations of potato root diffusate (PRD)-hatched J2 and mRNA populations of younger stages: developing nematodes in young cysts (two months post-inoculation), developing nematodes in young cysts (three months post-inoculation), unhatched J2 in diapause and rehydrated unhatched J2 after diapause. In the first paper about cDNA-AFLP, the expression levels revealed by cDNA-AFLP were shown to be comparable with results obtained by Northern blot analysis (Bachem et al. 1996).

To establish whether the intensities of TDFs revealed by cDNA-AFLP correctly reflected the differences in the original mRNA populations in our system, the expression patterns of constitutively (glyceraldehyde-3-phosphate-dehydrogenase, encoding cDNA named *gpd*) and transiently expressed genes (GR-*engs*) were monitored.

A cDNA encoding the enzyme glyceraldehyde-3-phosphate-dehydrogenase (Qin et al. 1998) was taken as a reference. The crucial role of GAPDH in glycolysis and

gluconeogenesis suggests a constitutive expression of *gpd* in *G. rostochiensis*. If cDNA-AFLP correctly reflects the abundances of GAPDH-encoding mRNAs, similar band intensities would be observed in all five developmental stages. Based on sequence information, digestion of the GAPDH-encoding cDNA with *Eco*R I and *Taq* I would produce a band of 93 bp. Two selective primer extensions were used (E + GT, T+ CG, for E and T see Materials and Methods) to reduce the complexity and facilitate the identification of the proper TDF. In all stages, TDFs were observed at the predicted position and the intensities of the corresponding bands were nearly identical (Fig. 1). The 93 bp TDF was isolated from the gel, and sequencing revealed a perfect match with *gpd*.

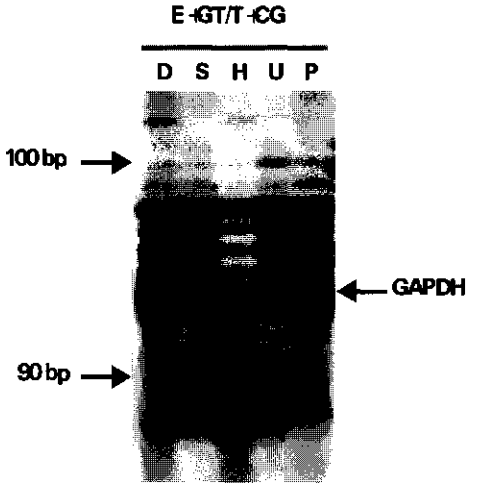


Fig. 1. The expression pattern of a GAPDH-encoding mRNA from the potato cyst nematode *G. rostochiensis* as displayed by cDNA-AFLP. The corresponding cDNA fragment was amplified using E+GT/T+CG. The band intensities were very similar in all five different developmental stages. These stages are indicated by capitals. D: unhatched J2 in diapause, S: unhatched J2 after diapause, rehydrated for 2 days in water, H: freshly hatched J2 in potato root diffusate, U: developing nematodes (J1) in gravid females two months post-inoculation, P: developing nematodes (J2) in gravid females three months post-inoculation.

β-1,4-endoglucanase shows a characteristic expression pattern in potato cyst nematodes with high levels of transcription just prior the onset of parasitism (Smant et al. 1997). To date, two highly homologous β-1,4-endoglucanases (GR-*eng*-1 and GR-*eng*-2) have been identified in *G. rostochiensis* (Smant et al. 1998). Based on available sequence information, a band of 398 bp is expected to appear on a cDNA-AFLP gel. Having an ancillary cellulose binding domain, ENG-1 is substantially larger than ENG-2, and it includes the restriction sites for a second TDF (104 bp). The primer combination E+TT and T+TG was used to amplify the 398 bp fragment. The extensions E+AA and T+GG were employed for the 104 bp fragment of ENG-1 (data not shown). The resulting AFLP patterns showed amplification products of both sizes. The highest band intensities were observed in infective J2 and in re-hydrated, unhatched J2 (Fig. 2.) which corresponds with our previous findings (Smant et al. 1997). The 398 bp TDF sequence showed 99% identity with GR-*eng*-2.

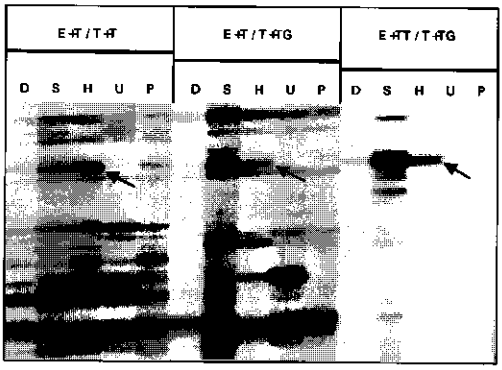


Fig. 2. The expression pattern of β-1,4-endoglucanases-encoding mRNAs from five different developmental stages of the potato cyst nematode *G. rostochiensis* as displayed by cDNA-AFLP. The arrows point to the bands produced by Gr-*eng*s. Three different primer extensions were used: E+T / T+T, E+T / T+TG and E+TT / T+TG. The capitals D, S, H, U and P refer to the different developmental stages (for explanation, see Fig. 1. legend).

To further verify the reliability of this system, a semi-quantitative PCR method was used to study the expression levels of β-1,4-endoglucanases in the five different life stages. Semi-quantitative PCR was carried out on cDNAs made from these stages. The expression levels of β-1,4-endoglucanases were normalized by the levels of *gpd* expression. A very similar expression pattern was observed as compared to cDNA-AFLP (Fig. 3).

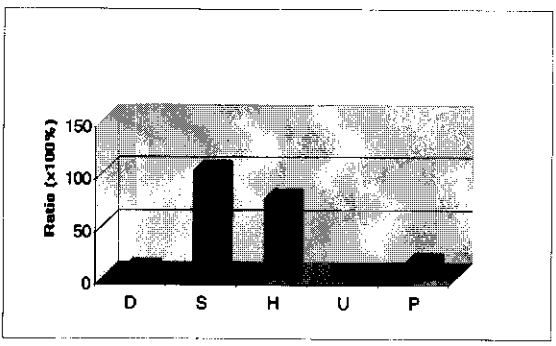


Fig. 3. Semi-quantitative PCR analysis of ß - 1, 4- endoglucanases expression in five different developmental stages of the potato cyst nematode *G. rostochiensis*. On the Y-axis the ratio is given between the net intensities of ß-1, 4-endoglucanases and GAPDH. The capitals D, S, H, U and P refer to the different developmental stages (for explanation see Fig. 1. legend).

It is concluded that cDNA-AFLP correctly reflects the expression patterns of a constitutively expressed gene, *gpd*, and the developmentally regulated genes, GR-*eng-1* and GR-*eng-2*.

The identification of TDFs with a pathogenicity factor like expression pattern

Assuming that the kinetics of expression revealed by cDNA-AFLP indeed reflect the mRNA abundances in the original mRNA populations, eight primer combinations were

used (E + A/T and T + A/C/G/T) to generate a large number of mRNA expression profiles. Each primer combination resulted in about 230 discrete bands on large polyacrylamide gels. The complexity of the resulting patterns prompted us to use one additional selective nucleotide (T+2 instead of T+1). Hence, E+C/G was used in combination with all possible T+2 extensions and, as a consequence, the average number of TDFs dropped to 85 per lane. In total 40 primer combinations were tested, resulting in the visualization of about 4500 TDFs. Approximately 100 TDFs (≈ 2%) appeared to be up-regulated in rehydrated cysts and / or in infective J2. A sample of fourteen TDFs was excised from gel, cloned and sequenced. To check whether the obtained sequences indeed corresponded with the up-regulated TDFs, one or two additional selective nucleotides were introduced complementary to the sequences at the 5' and 3' ends of the fragment. From the 14 TDFs tested, 10 TDFs still displayed the same expression pattern and only these TDFs were analyzed further. In a control experiment, the effect of additional selective nucleotides on the expression pattern of β-1,4-endoglucanases was verified. The expression patterns were nearly identical, irrespective of the number of selective nucleotides (Fig. 2). Except for one TDF - A13 - that shows high homology with a 4-hydroxyphenylpyruvate dioxygenase gene (data not shown), no significant homologies with other known genes were found among the other up-regulated TDFs.

In situ hybridization of up-regulated TDFs

Subventral and dorsal esophageal gland products are thought to be crucial in the interaction between endoparasitic nematodes and their host. This can be illustrated by β-1,4-endoglucanase, the first pathogenicity factor cloned from a plant parasitic nematode. In situ hybridization in pre-parasitic J2 of *G. rostochiensis* showed labeling in the subventral glands only (Smant et al. 1998). For 10 TDFs, sense and anti-sense DNA probes were made by asymmetric PCR using either the E+0 or the T+0 primer. We have observed that the background staining with DNA probes was always much less than with RNA probes. Anti-sense DNA probes of the TDFs A4, A18 and A41 specifically labeled the dorsal gland of infective J2. Labeling was observed around the nucleus in the gland cell, and no hybridization was detectable in the gland cell extensions or the ampullae. No labeling was observed for the sense probes. The detected transcript levels of A4, A41 and A18 were respectively high, moderate, and low (Fig. 4). These three TDFs appeared to be upregulated in the PRD-hatched J2 stage and this expression pattern was confirmed using primers with corresponding selective extensions at the 3'-end (Fig. 5). The antisense probe of A11 labeled the rectal gland (data not shown). The remaining six probes did not show any hybridization signal.

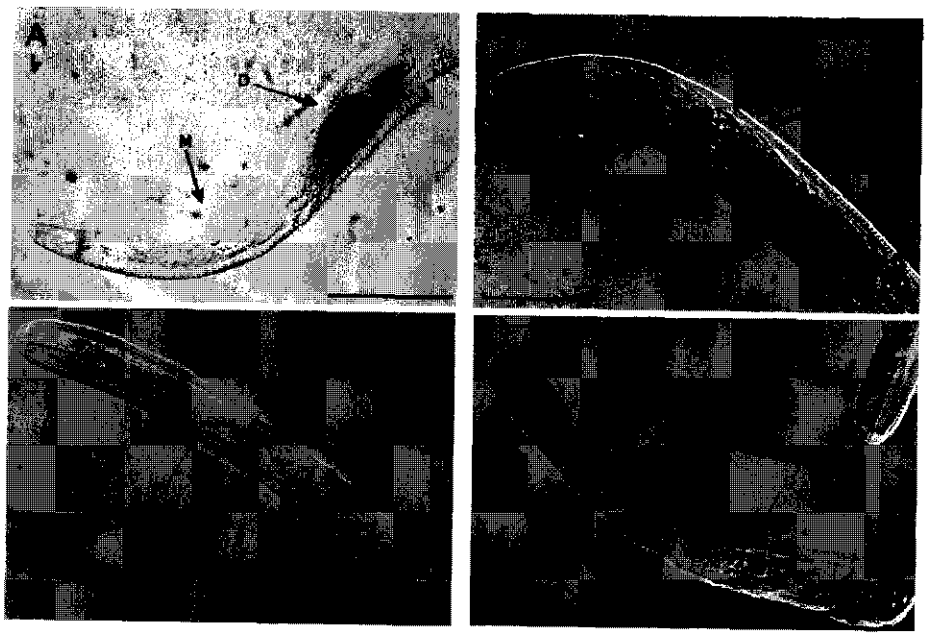


Fig. 4. *In situ* hybridization using three different TDF-derived probes from *G. rostochiensis*. Labeling was done in preparasitic J2. A: labeling pattern using an A4 antisense probe. B: labeling pattern with an A18 antisense probe. C: labeling pattern using an A41 antisense probe. D: A typical image of a sense probe labeling. Arrows point at the dorsal gland (D) and at the metacorpus (M). Scale bar = 20 μm

Analysis of 5' ends of mRNAs possibly encoding nematode secreted compounds

A dorsal gland protein could be a constituent of the nematode secretions if it is preceded by a signal peptide for secretion. To find out whether the up-regulated, dorsal gland specific TDFs A4, A18 and A41 included such a signal, the 5' end of the cDNAs was amplified from a cDNA library (Smant et al. 1998). For A4, a fragment of 162 bp was obtained showing an overlap of 59 bp with the 5' end of the corresponding TDF. Similarly, a 733 bp fragment was obtained for A18, which had 61 bp in common with the 5' end of the TDF sequence. For A41, a 470 bp fragment was isolated having 175 bp in common with the 5' end of the A41 TDF sequence. In all three cases, a putative start codon could be identified preceded by 3 to 4 adenine residues, a phenomenon frequently observed in *C. elegans* genes (Blumenthal et al. 1997) and also in potato cyst nematode genes (Qin, unpublished observations). The conceptual translation predicted that A4, A18 and A41 contained open reading frames. The computer algorithm SignalP (Nielsen et al. 1997) was used to assess whether the protein was preceded by a signal peptide for secretion. In all three cases, such a signal was predicted to be present. It is concluded that

the proteins corresponding to the TDFs A4, A18 and A41 are produced in the dorsal gland and, most likely, secreted through the stylet.

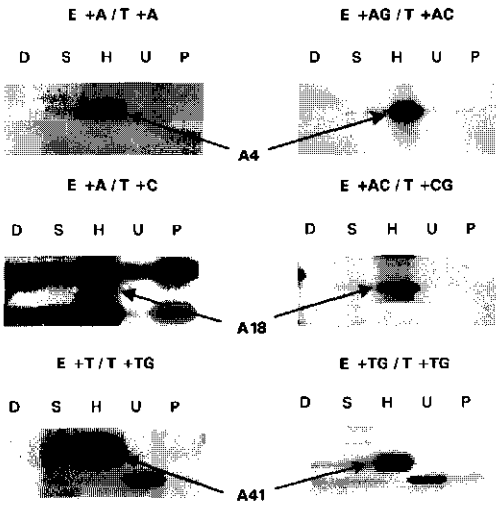


Fig. 5. The expression patterns of three TDFs, A4, A18 and A41 from the potato cyst nematode *G. rostochiensis* as revealed by cDNA-AFLP. The arrow-pointed bands from the left panel were cut out, cloned and sequenced. On the basis of the sequence results, the primers E and T were extended with one or two additional selective nucleotides. The resulting cDNA-AFLP expression patterns are shown (right panel). The capitals D, S, H, U and P refer to the different developmental stages (for explanation see Fig. 1. legend).

DISCUSSION

In this paper a novel cDNA-AFLP-based strategy is described that resulted in the cloning of three putative pathogenicity factors – named A4, A18 and A41 - from the dorsal esophageal gland of the potato cyst nematode *G. rostochiensis*. Secretions from the subventral and dorsal esophageal glands of cyst and root knot nematodes are thought to be essential for many aspects of parasitism (Williamson and Hussey, 1996). The first cloned secretory product from the subventral glands of a cyst nematode was shown to be a β-1,4-endoglucanase (Smant et al. 1998). The three dorsal gland genes presented here are the first ones to be identified from a plant parasitic nematode. None of these putative secretion compounds show significant homology to any known protein. Hence, the functions of A4, A18 and A41 in the interaction between the potato cyst nematode and its host plants remain to be established.

To identify proteinaceous pathogenicity factors present in the secretory glands of the potato cyst nematode, we developed a strategy that uses the corresponding mRNAs as a starting point. Pools of mRNAs of distinct developmental stages were compared by cDNA-AFLP. By using this PCR-based technology the problem of scarcity of starting material was circumvented. Cyst nematodes have a single infective stage - the pre-parasitic second stage juvenile - and we postulate that pathogenicity factors related to early events in the infection process will predominantly be present in this life stage and they will be functional outside the nematode. Although the results presented in this paper show the usefulness of our selection criteria, it is noted that exceptions (*e.g.* a pathogenicity factor that is expressed constitutively) will be missed.

Firstly, the reliability of the expression patterns as displayed by the cDNA-AFLP technique was verified. Contrary to differential display where random primers are used for amplification of cDNAs (Liang and Pardee 1992) , cDNA-AFLP permits a relative easy validation of the displayed profiles by using control genes with known expression patterns. The cDNA fragment corresponding to the constitutively expressed gene *gpd-1* was present in all developmental stages with similar intensities. The amplified fragment from GR-*eng-1* was up-regulated in pre-parasitic stages, and this result corresponded well with a previous study of β-1,4-endoglucanases at the protein level (Smant et al. 1997). It is noted that the expression profiles of GR-*eng*-1 and -2 (95% identical in the first 884 nucleotides) could be monitored separately. This represents a clear advantage of cDNA-AFLP technique over the hybridization-based methods where the hybridization signals from each of the highly homologous sequences will be indistinguishable. The validity of the expression patterns was further verified by using additional selective nucleotides. By doing so the number of displayed amplification products was drastically reduced and, as a consequence, the primer / target ratio was changed. This change did not affect the relative intensities of the control genes. In a previous study, gene expression during potato tuber formation as revealed by cDNA-AFLP was checked by Northern analysis and similar results were obtained (Bachem et al. 1996). It is concluded that, for the limited number of control genes available for *G. rostochiensis*, the correctness of the gene expression profiles as revealed by cDNA-AFLP could be confirmed.

In this paper, the expression levels of 4560 TDFs in five developmental stages of *G. rostochiensis* were monitored. Certainly, a small number of transcripts will be represented by more than one TDFs. In a similar study on potato tuber formation process, the authors found that about 2% out of the 200 cloned TDFs appeared to be derived from the same transcript (Bachem et al. 1996). Given the fact that this double representation could only be detected when both fragments hit upon the same gene sequence in the database during homology search, we presume this percentage is an underestimation. If we suppose that 10% of the displayed TDFs are derived from the same transcript, the expression patterns of about 4,000 genes were displayed.

To estimate the efficiency of the cDNA-AFLP approach, it is useful to make a comparison with *Caenorhabditis elegans*. The *C. elegans* genome sequencing project revealed that the 97 Mb haploid genome sequence of this free living nematode contains 19,099 predicted protein-coding genes. Exon sequences are predicted to cover 27% of the genome, giving an average size of 1.35 kb per transcript (The *C. elegans* Sequencing Consortium. 1998). The genome sizes of *G. rostochiensis* and *C. elegans* are similar (Rouppe van der Voort et al. 1999). Under the assumption that the average transcript size of *G. rostochiensis* is about 1.35 kb as well, approximately 33% of the potato cyst nematode genes will be cut at least once (1.35 kb/4^6 bp = 33%) by *Eco*R I recognizing a six bp motive. The frequent cutter *Taq* I will restrict most of the cDNA fragments generated by *Eco*R I. If these assumptions are correct, the estimated number of expressed genes in juveniles of the potato cyst nematode is about 12,000 (= 4,000 / 0.33). If a second restriction enzyme combination is used, the coverage of the genome should reach about 56% (=33%+(1-33%)×33%), a third combination about 70% (=56%+(1-56%)×33%). In fact, a survey of the expression patterns of the genes representing a significant part of the genome can be achieved with relatively small investment both in time and financial resources.

Among the 4,000 expression profiles, approximately 100 TDFs were up-regulated in rehydrated cysts and/or in infective J2. *In situ* hybridisations showed that three TDFs, A4, A18 and A41, were specifically produced in the dorsal gland in infective pre-parasitic juveniles. The presence of a predicted signal peptide for secretion indicates that these proteins are translocated in the secretory pathway and, as such, they meet the criteria set out for putative pathogenicity factors of the potato cyst nematode. The partial cDNA sequences did not reveal a significant homology with known proteins. Their presence in the dorsal esophageal gland (and their absence in the subventral glands) suggest that these proteins could be involved in processes finally resulting in food extraction from the host plant, and presumably not in migration or in protection against host plant defense reactions.

Knock-out mutants would be an ideal tool to unravel the role of these dorsal gland proteins in pathogenicity. Introduction of double-stranded RNA has been shown to specifically disrupt the activity of genes containing homologous sequences in *C. elegans* (Fire et al. 1998). Alternatively, target-binding single chain antibodies could be expressed in the cytosol of cortical root cells. Using this approach the relative importance of a putative pathogenicity factor could be assessed (Schots et al.1992). A third, more general approach is the heterologous expression of putative pathogenicity factors in plants. The resulting phenotypes could give an indication about the possible roles of these putative pathogenicity factors in parasitism.

This paper illustrates the efficiency of a novel strategy for the identification of pathogenicity factors from the potato cyst nematode *G. rostochiensis*. This relatively general approach can be applied to any plant-pathogen combination as long as distinct mRNA populations can be isolated from infectious and non-infectious life stages of the pathogen. This technique will be particularly suitable for many non-model organisms for which limited genome sequence information is available. Contrary to *e.g.* expressed sequence tags (ESTs), which relies solely on homology searches, this cDNA-AFLP-based approach allows for the identification of putative pathogenicity factors which are specifically expressed at infective stages, even if they do not show significant homology to any known gene. As such this strategy could be a valuable tool in the identification of pathogen signals mediating host plant manipulation.

MATERIALS AND METHODS

Nematode culture

G. rostochiensis pathotype Ro1-Mierenbos was cultured as described previously ((De-Boer et al.1992)). RNA was extracted from five developmental stages: (D-dormant) dehydrated unhatched second stage juveniles (J2's) in cysts (in diapause); (S-soaked) rehydrated unhatched J2's in one-year-old cysts after exposure to sterile tap water for two days; (H-hatched) pre-parasitic J2's (dry cysts were incubated in sterile tap water for one week, tap water is replaced by potato root diffusate (PRD) in the second week); (U-undifferentiated into J2's) developing nematodes (mostly J1's) in gravid females two months post-inoculation; (P-pre-diapause) developing nematodes (J2's) in gravid females three months post-inoculation.

RNA isolation

From each of the five developmental stages about 100 µg total RNA was isolated using TRIzol Reagent (Life Technologies, Breda, The Netherlands). The integrity of the total RNA was checked on a denaturing agarose gel. Dynabeads Oligo $(dT)_{25}$ (Dynal A.S, Oslo, Norway) were used to isolate about 1 µg mRNA from each of the developmental stages.

cDNA-AFLP analysis

Besides some minor modifications, cDNA-AFLP was performed as described (Bachem et al.1996)). In short, cDNA was synthesized from 1 µg of mRNA using the Superscript Choice System for cDNA synthesis-kit (Life Technologies). The first strand cDNA was reverse transcribed with a T_{12} oligonucleotide primer. Subsequently, cDNA was digested by restriction endonucleases *Eco*R I and *Taq* I and ligated to *Eco*R I and *Taq* I adapters ((Vos et al.1995)). Pre-amplification was performed in 25 cycles using primers (indicated as 'E' and 'T') corresponding to the *Eco*RI and *Taq*I adapters without extension (E+0: GAC TGC GTA CCA ATT C and T+0: GAT GAG TCC TGA CCG A) using a standard pre-amplification program. After 100 times dilution of the PCR product, the template was amplified again with one or two selective base extensions at the 3'-end of the primers E and T using a standard AFLP touchdown selective amplification program. As fragment size references on gel (50 cm in length), a 30-330 bp AFLP ladder (Life Technologies) was used in combination with a 50-500 bp SequaMark ladder (Research Genetics, Huntsville, Canada). Differentially expressed bands were cut out from the gel and re-amplified using High Fidelity polymerase (Roche Diagnostics, Mannheim, Germany). The amplified fragments were checked on a 2% agarose gel and cloned subsequently into pCR2.1-TOPO vector (Invitrogen, Leek, the Netherlands). Transformed *E. coli* colonies were checked by PCR with the same primers used in the selective amplification to confirm the presence of the expected insert. Plasmids were purified using Wizard Plus Miniprep DNA Purification System (Promega, Madison, WI) and sequenced either on an Automated Laser Fluorescent DNA sequencer (Pharmacia, Uppsala, Sweden) or on a LI-COR DNA sequencer (LI-COR , Lincoln, NE). Confirmation of the cloned cDNA-AFLP fragments was done by using one additional selective base to the 3'-ends of the E and/or T primers in selective amplifications and run on gel.

Semi-quantitative PCR

To validate the cDNA-AFLP expression patterns, semi-quantitative RT-PCR was performed on β-1,4-endoglucanases and glyceraldehyde-3-phosphate-dehydrogenase (GADPH) encoding mRNAs from five different stages. β-1,4-endoglucanases are pathogenicity factors of *G. rostochiensis* and these cell wall degrading enzymes are abundantly expressed in J2 within cyst soaked with water and PRD hatched J2 (Smant et al.1997). The primer *cel-up* (5'-TGGAACGCCGATACAGTGAAGG- 3') and *cel-down* (5'-GGAGGCGTAGAAGTGGAAGGTGTA-3') were used to amplify both endoglucanases (GR-*eng*-1 and GR-*eng*-2). cDNA was prepared as described above from 1 µg mRNA. A small fraction (0.6%) of synthesized cDNA was used as template in a PCR reaction. The primers *gpd-up* (5'-GCGCTGCGGTTGAGAAGGAC-3') and *gpd-down* (5'-CCGGTGCTGGCAGGGATGATGT-3') were used to amplify the

corresponding cDNA. PCR was performed in a 100 µl reaction volume containing 1× Supertaq PCR buffer (SphaeroQ, Breda, The Netherlands), 0.8 mM dNTPs, 0.2 U Supertaq and 0.25 µM of each primer. A touchdown PCR profile was used: 5 cycles of 20 s at 92°C, 30 s at 65°C, 60 s at 72°C , followed by 29 cycles of 20 s at 92°C, 30 s at 60°C, 60 s at 72°C. A 10 µl aliquot was taken out every three cycles and analyzed on a 2% agarose gel including 0.4 µg/ml ethidium bromide. In the log-linear phase of the amplification reaction, the PCR products were analyzed using KODAK 1D image analysis software (EASTMAN KODAK, Rochester, NY). The amount of PCR product corresponding to β-1,4-endoglucanases was normalized using GAPDH as baseline reference.

Sequence analysis

DNA sequences were analysed using the Laser gene software package (DNASTAR, Madison, WI). Database search was done using WU-Blast at Human Genome Centre, Baylor College of Medicine, Blast service at the National Centre for Biotechnology Information (NCBI) and FASTA at European Bioinformatics Institute (EBI). The computer algorithm SignalP (Nielsen et al.1997) was used to predict the presence of a signal peptide for secretion and the corresponding putative cleavage site.

***In situ* hybridization**

Except for the use of single stranded DNA (instead of RNA) probes - synthesized by linear PCR as described (Tabara et al.1996) - the nematode fixation, hybridization and detection steps were essentially performed as described ((de et al.1999)). Briefly, cDNA-AFLP fragments cloned in pCR2.1-TOPO vector (Invitrogen) were first amplified with Universal M13 (-20) forward and Universal M13 reverse primers (Invitrogen). The vector sequence from the amplified product was removed by the addition of *Eco*R I (37°C, 1 h). The digested product was checked on 1% agarose gel. The sense and antisense probes were made by asymmetric PCR in the presence of DIG-dUTP (Roche Diagnostics) using either the E+0 or T+0 primer with 20 ng of digested product as template in a 20 µl reaction volume. The DIG-labeled probes were purified through a G50 Mini Quick Spin DNA column (Roche Diagnostics) and 10 µl TE buffer was added. In each hybridization reaction, 30 µl of labeled probe was used. Alkaline phosphatase activity was detected by the addition of X-phosphate and NBT (Roche Diagnostics), and labeled infective juveniles were observed under a Leica inverted microscope (Leica, Deerfield , IL).

Cloning of the 5' end of the cDNA

Gene specific primers were used in combination with vector primers in pcDNAII to amplify the 5'-end region of each cDNA-AFLP fragment (A4, A18, A41) from a cDNA library constructed in pcDNAII as described (Smant et al.1998).

ACKNOWLEDGEMENT

The authors wish to thank Christian Bachem and Peter van Zandvoort (Wageningen University and Research Center) for very useful discussions concerning the cDNA-AFLP technique. We would like to extend our gratitude to Jan de Boer and Thomas Baum (Iowa State University, USA) for introducing us to *in situ* hybridization technique and also Tom

Tytgat and Godelieve Gheysen (University of Gent, Belgium) for useful discussion about *in situ* protocols. We are grateful to Andrew Gane (Cardiff University, UK) for sequencing some of the clones. This work was supported by EC grant BIO4-CT96-0318 (LQ).

References

1. Bachem, C.W.B., Van Der Hoeven, R. S., De Bruijn, S. M., Vreugdenhil, D., Zabeau, M., and Visser, R. G. 1996. Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: Analysis of gene expression during potato tuber development. Plant J. 9:745-753.
2. Blumenthal T, Steward, K. 1997. RNA processing and gene structure. In: Riddle DL, Blumenthal T, Meyer BJ, Priess JR, editors. *C. ELEGANS* II. Cold Spring Harbor Laboratory Press. 117-146.
3. De Boer, J.M., Overmars, H. A., Bakker, J., and Gommers, F. J. 1992. Analysis of two-dimensional protein patterns from developmental stages of the potato cyst nematode, *Globodera rostochiensis*. Parasitology 105:461-474.
4. 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 beta-1,4-endoglucanases in the subventral esophageal glands of *Heterodera glycines*. Mol. Plant Microbe Interact. 12:663-669.
5. The *C. elegans* Sequencing Consortium.1998. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science 282:2012-2018.
6. Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391:806-811.
7. Goverse, A., Rouppe van der Voort, J.N.A.M., Rouppe van der Voort, C., Kavelaars, A., Smant, G., Schots, A., Bakker, J., and Helder, J. 1999. Naturally induced secretions of the potato cyst nematode co-stimulate the proliferation of both tobacco leaf protoplasts and human peripheral blood mononuclear cells. Mol. Plant Microbe Interact. 12:872-881.
8. Liang, P. and Pardee, A. B. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257:967-971.
9. Niebel, A., Engler, J. D., Hemerly, A., Ferreira, P., Inze, D., Vanmontagu, M., and Gheysen, G. 1996. Induction of cdc2a and cyc1at expression in A*rabidopsis thaliana* during early phases of nematode induced feeding cell formation. Plant J. 10:1037-1043.
10. Nielsen, H., Engelbrecht, J., Brunak, S., and Von Heijne, G. 1997. Identification of prokaryotic and eukaryotic signal peptide and prediction of their cleavage sites. Protein Engineering 10:1-6.
11. Qin, L., Smant, G., Stokkermans, J., Bakker, J., Schots, A., and Helder, J. 1998. Cloning of a trans-spliced glyceraldehyde-3-phosphate-dehydrogenase gene from the potato cyst nematode *Globodera rostochiensis* and expression of its putative promoter region in *Caenorhabditis elegans*. Mol. Biochem. Parasitol. 96:59-67.

12. Ray, C., Abbott, A. G., and Hussey, R. S. 1994. Trans-splicing of a *Meloidogyne incognita* mRNA encoding a putative esophageal gland protein. Mol. Biochem. Parasitol. 68:93-101.
13. 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.
14. Rosso, M.N., Favery, B., Piotte, C., Arthaud, L., de Boer, J., 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. Mol. Plant Microbe Interact. 12:585-591.
15. Rouppe van der Voort, J.N.A.M., van Zandvoort. P., Overmars, H., Helder, J., and Bakker, J. 1999. Linkage analysis by genotyping of sibling populations: a genetic map for the potato cyst nematode constructed using a "pseudo-F2" mapping strategy. Mol. Gen. Genet. 261:1021-1031.
16. Schots, A., De Boer, J. M., Schouten, A., Roosien, J., Zilverentant, J. F., Pomp, H., Bouwman, L., Overmars, H., Gommers, F. J., Visser, B., Stiekema, W. J., and Bakker, J. 1992. 'Plantibodies': a flexible approach to design resistance against pathogens. Neth. J. Pl. Path. Supplement 2:183-191.
17. Smant, G., Goverse, A., Stokkermans, J. 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.
18. Smant, G., Stokkermans, J. P., Yan, Y., de Boer, J., 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 beta-1, 4-endoglucanase genes from two species of plant-parasitic cyst nematodes. Proc. Natl. Acad. Sci. USA 95: 4906-11
19. Tabara, H., Motohashi, T., and Kohara, Y. 1996. A multi-well version of *in situ* hybridization on whole mount embryos of *Caenorhabditis elegans*. Nucleic Acids Res. 24:2119-2124.
20. Vos, P., Hogers, R., Bleeker, M., Reijans, M., van-de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23:4407-4414.
21. Williamson, V.M. and Hussey, R. S. 1996. Nematode pathogenesis and resistance in plants. Plant Cell 8:1735-1745.
22. Wyss, U. and Zunke, U. 1986. Observations on the behaviour of second stage juveniles of *Heterodera schachtii* inside host roots. Rev. Nematol. 9:153-165.
23. Wyss, U. 1992. Observations on the feeding behaviour of *Heterodera schachtii* throughout development, including events during moulting. Fundam. Appl. Nematol. 15:75-89.
24. Yan, Y., Smant, G., Stokkermans, J., Qin, L., Helder, J., Baum, T., Schots, A., and Davis, E. L. 1998. Genomic organization of four beta-1,4-endoglucanase genes in plant-parasitic cyst nematodes and its evolutionary implications. Gene 220:61-70.

Chapter III

GenEST, a powerful bi-directional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP

Qin,L., Prins,P., Jones,J.T., Popeijus,H., Smant,G., Bakker,J. and Helder,J.
GenEST, a powerful bi-directional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP.
Nucleic Acids Research 2001.29:1616-1622

GenEST, a powerful bi-directional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP

ABSTRACT

The release of vast quantities of DNA sequence data by large scale genome and Expressed Sequence Tag (EST) projects underlines the necessity of the development of efficient and inexpensive ways to link sequence databases with temporal and spatial expression profiles. Here we demonstrate the power of linking cDNA sequence data (including EST sequences) with transcript profiles revealed by cDNA-AFLP, a highly reproducible differential display method based on restriction enzyme digests and selective amplification under high stringency conditions. We have developed a computer program (GenEST) that predicts the sizes of virtual transcript derived fragments (TDFs) of *in silico* digested cDNA sequences retrieved from databases. The vast majority of the resulting virtual TDFs could be traced back among the thousands of TDFs displayed on cDNA-AFLP gels. Sequencing of the corresponding bands excised from cDNA-AFLP gels revealed no inconsistencies. As a consequence, cDNA sequence databases can be screened very efficiently to identify genes with relevant expression profiles. *Vice versa*, it is possible to switch from cDNA-AFLP gels to sequences in the databases. Using the restriction enzyme recognition sites, the primer extensions and the estimated TDF size as identifiers, the DNA sequence(s) corresponding to a TDF with an interesting expression pattern can be identified. In this paper, we show examples in both directions by analyzing the plant parasitic nematode *Globodera rostochiensis*. Various novel pathogenicity factors were identified by combining ESTs from the infective stage juveniles with expression profiles of about 4,000 genes in five developmental stages produced by cDNA-AFLP.

KEYWORDS: cDNA-AFLP, EST, functional genomics, bioinformatics, potato cyst nematode

INTRODUCTION

With the advent of high throughput techniques for DNA sequencing, whole genome sequences from several organisms have become available (1,2) and many others will be available in the near future. At the same time, millions of expressed sequence tags (ESTs) - single pass sequences of cDNA clones selected randomly from a library - have been generated and deposited in public and private databases. Searching for homologous sequences in databases is usually the first step used towards understanding the functions of newly identified genes. Homology information is useful for orthologous genes, but in the case of paralogs the value of this information may be more limited. Furthermore, it is often found that a significant proportion (40% to 60%) of newly identified DNA

sequences lack homology with genes, for which the functions are known (2,3). Additional tools are therefore needed to allow functional analysis of newly identified genes.

Biological responses and developmental processes are precisely controlled at the level of gene expression. Information on the temporal and spatial regulation of gene expression often sheds light on the potential function of a particular gene. Hence, an essential aspect of functional genomics is the transcriptome, *i.e.*, the analysis of expression patterns of genes at a large scale. There are currently three high throughput techniques for large scale monitoring of gene expression: serial analysis of gene expression (SAGE) (4), hybridization-based methods (5,6) and gel-based RNA fingerprinting techniques such as differential display (7) and cDNA-AFLP (8). In principle, SAGE can provide quantitative data concerning gene expression. However, it is expensive and labor-intensive when multiple sample points are to be compared. Microarray technology is very powerful in generating a broad view of gene expression. Unlike cDNA arrays, oligonucleotide arrays are able to distinguish between highly homologous sequences. However, the design of oligonucleotide arrays requires comprehensive sequence knowledge at present only available for a small number of organisms. cDNA-AFLP is an inexpensive gel-based method for analysis of gene expression patterns and can be performed in any laboratory.

In the cDNA-AFLP procedure, cDNAs synthesized from mRNAs isolated from various sample points are digested by two restriction enzymes. Oligonucleotide adapters are then ligated to the resulting restriction fragments to generate template DNA for PCR. PCR primers complementary to the adapter sequences with additional selective nucleotides at the 3' ends allow for specific amplification of a limited number of cDNA fragments. Unlike differential display methods that make use of small random primers (7), relatively high annealing temperatures can be used and, hence, cDNA-AFLP is more stringent and reproducible. In contrast to most hybridization-based techniques, cDNA-AFLP will distinguish between highly homologous genes from gene families while (contrary to oligonucleotide arrays) no sequence foreknowledge is needed.

Since sequence information is accumulating at an unprecedented rate for a wide variety of organisms, there is an urgent need for efficient and inexpensive ways to screen these databases on genes with interesting expression profiles. Here we report on the advantages of combining ESTs with cDNA-AFLP data. The potential benefits of this combination in gene discovery and functional analysis prompted us to develop a computer program that creates restriction patterns of cDNAs *in silico* in accordance with the enzyme combinations used in cDNA-AFLP. The resulting virtual cDNA fragments are ordered according to the extensions of the amplifying primers and their sizes. These virtual fragments can then be traced back on cDNA-AFLP gels to identify the corresponding bands, with primer extensions and fragment sizes as a unique identifier. The program can also be used in the opposite direction by using the size and primer extensions of a potentially interesting band identified on a cDNA-AFLP gel as criteria to search the corresponding cDNA. This simplifies the procedure of cloning full-length genes with interesting temporal and spatial expression patterns.

In this paper, we demonstrate the utility of the program by linking EST sequence data and expression profiles of about 4,000 genes from the potato cyst nematode *Globodera rostochiensis*, which causes extensive damage to solanaceous crops. Genes potentially related to the nematode's ability to parasitize plants were identified within a pool of

hundreds of ESTs. We show that this program could be useful in any system where stage or tissue specific genes are to be selected from pools of (uncharacterized) cDNAs.

MATERIALS AND METHODS

The nucleotide sequences of the ESTs described in this study are available in the Genbank EST division (dbEST) under the accession numbers BE607308 (*GE1867*), AW506364 (*GE1782*), AW506154 (*GE1483*), AW506045 (*GE1349*), AW505895 (*GE1133*), AW506065 (*GE1373*), AW506280 (*GE1659*), AW506299 (*GE1699*), AW505736 (*GE99*), AW506094 (*GE1409*), AW505716 (*GE54*), AW505855 (*GE1084*), AW506406 (*GE1816*), BE607310 (*GE2075*) and BE607309 (*GE1156*) respectively.

Generation of expressed sequence tags (EST)

The ESTs described by Popeijus *et al.* (9) were used in this study. Briefly, total RNA was extracted from infective second stage juveniles (J2) of the potato cyst nematode *G. rostochiensis* pathotype Ro1 Mierenbos that were freshly hatched in potato root diffusate (PRD). cDNA primed with an oligo(dT) primer was directionally cloned in the pcDNA II vector (Invitrogen, Leek, The Netherlands). The resulting library contained at least 2.5 $x10^6$ recombinant plasmids. ESTs were obtained by random sequencing of the library inserts from the 5' end.

cDNA-AFLP profile

cDNA-AFLP profiles were generated as described by Qin *et al.* (10). Briefly, total RNA was extracted from five developmental stages of *G. rostochiensis*: (D) dehydrated unhatched J2s in cysts (in diapause); (S) rehydrated unhatched J2s in one-year-old cysts after exposure to sterile tap water for two days; (H) pre-parasitic J2s (dry cysts incubated in sterile tap water for one week, then PRD for a second week); (U) developing nematodes (mostly J1s) in gravid females two months post-inoculation; (P) developing nematodes (J2s) in gravid females three months post-inoculation. cDNA was synthesized with oligo$(dT)_{12\text{-}18}$ as primer. The resulting cDNAs were then digested with the restriction enzymes *Eco*R I and *Taq* I and ligated to corresponding adapters. The ligated cDNA fragments were subsequently amplified by *Eco*R I and *Taq* I primers that annealed to the adapters in PCR reactions and displayed on polyacrylamide gels.

GenEST program

GenEST was developed on Linux using the GNU C++ toolset and the Standard Template Library. It is command line-based and can be run on Unix/Linux and from the MS-DOS prompt under Microsoft Windows. GenEST (program, source code and detailed manual) can be downloaded from: http://www.spg.wau.nl/nema/nm_res-e.htm (**GenEST** hyperlink in Section I. Plant-nematode interactions).

A command file can be created with a text editor, which contains restriction enzyme recognition sites to be used as the begin- and end-tags and the marker length modifier. Multiple combinations can be defined in a command file. GenEST uses the begin tag to search for the tag sequence in the cDNA data, which are contained in the input files in FASTA format. If such a tag is found, it will continue its search for a matching end tag. This search action is executed in both directions

for all begin/end tag combinations. The marker length modifier is designed to compensate for the additional adapter sequences present in the transcript-derived fragments (TDFs) as they appear on cDNA-AFLP gel.

Furthermore, the identifiers of a band on gel (restriction enzyme recognition sequences, primer extensions and the band size) can be used as a search query to quickly identify corresponding EST(s) in an automated procedure.

RESULTS

ESTs and cDNA-AFLP-based expression profiles

A cDNA library from second stage juveniles in the H-stage of the potato cyst nematode *G. rostochiensis* was used for the sequencing of 985 cDNA clones. Starting from the 5'-end, the average read was about 600 bp (9). In parallel, cDNA-AFLP-based gene expression profiles were generated from five distinct developmental stages - D, S, H, U and P - of this nematode species. The expression profiles were highly reproducible and no significant differences were observed between independent replicates. An average number of 32 bands per lane were displayed using *Eco*R I and *Taq* I primers with two selective nucleotides (E+NN and T+NN, respectively) extending beyond the adapters into the cDNA. Approximately 8,200 TDFs were displayed using the whole set of 256 (16 × 16) primer combinations. In a previous study (10), it has been shown that genes involved in plant parasitism are usually up-regulated in the developmental stage S and H or in the stage H only. Bands showing such expression patterns were excised from gels, cloned and sequenced. Sequencing of more than 100 TDFs revealed that the marker-based size estimations corresponded well to the actual size of these TDFs (with an accuracy of ± 1 nt for bands smaller than 300 nt and ± 3 nt for bands longer than 300 bp).

Generation of virtual TDFs from ESTs by using GenEST

E+AN		E+CN		E+GN		E+TN		Total
N= A	20	N=A	17	N= A*	32*	N= A	2	
C	6	C	11	C	15	C	11	
G	12	G	14	G	19	G	14	
T	14	T	33	T	22	T	18	
Total	52		75		56*		45	**228***

Table 1. Virtual TDFs generated after the *in silico* restriction with *Eco*R I and *Taq* I of 985 ESTs randomly picked from a cDNA library from infective juveniles of the potato cyst nematode *Globodera rostochiensis* using GenEST. E+AN/CN/GN/TN are the extensions of the *Eco*R I primer (core primer = E). Each *Eco*R I primer was combined with all *Taq* I primers (T + NN). Note: E+GA will constitute both a *Taq* I and an *Eco*R I recognition sequence (**GAATTCGA**). In this case, TDFs will not be amplified and cannot be traced back on cDNA-AFLP gel. Therefore these TDFs were not included in the total count (*).

We used GenEST to generate virtual TDFs from 985 ESTs. *Eco*R I and *Taq* I recognition sites were used as begin and end tags with a length modifier of 22 nt to account for the additional adapter sequences. The following lines were included in the command file:

GAATTC TCGA 22
TCGA GAATTC 22

A total of 228 virtual TDFs derived from 159 ESTs were predicted by GenEST (Table 1.). Out of these 159 ESTs, 100 were predicted to produce a single virtual TDF, 51 were predicted to give rise to two virtual TDFs each (thereby generating 102 TDFs), 6 ESTs were predicted to result in three TDFs each (generating 18 TDFs) and two ESTs were predicted to generate four TDFs each (8 TDFs in total).

To estimate how many genes are represented by the 8200 TDFs displayed in our study, we have randomly extracted 1,000 full length cDNAs of *C. elegans* from Genbank (both the size and the average GC content of the *G. rostochiensis* genome are similar to *C. elegans* (11)). These sequences were processed by GenEST and 336 cDNAs (≈ 34%, the remaining cDNAs did not contain both restriction sites) generated 693 virtual *Eco*R I / *Taq* I TDFs. The percentage of genes which produced TDFs is approximately 48% (336/693 = 48%) of the total TDF number. Assuming that the average mRNA size and the number of genes of the potato cyst nematode do not differ substantially from *C. elegans*, the 8,200 TDFs displayed on cDNA-AFLP gels in this study would represent about 4,000 expressed genes.

From ESTs to the corresponding TDFs on cDNA-AFLP gels

The vast majority of the virtual TDFs predicted could be located on the expected position in the cDNA-AFLP gel. The cases where no matches were found between virtual and real TDFs could usually be explained by the system used. Here we describe a detailed analyses of 52 virtual TDFs that were generated *in silico* using the primers E+AN in combination with all Taq I primers (T+NN) (Table 1). Multiple TDFs that originated from a single EST sequence were all checked. Matching bands could be found on gels for 41 TDFs. Eight virtual TDFs were smaller than the exclusion limit of 50 nt used in this study. As expected, these TDFs were not displayed. Lowering the exclusion limit would allow the display of bands down to 10 nt. Among these eight ESTs, six would produce a second virtual TDF. All these TDFs were identified on gels. Within the size range analyzed, only three virtual TDFs could not traced back on the cDNA-AFLP gels (see below).

The TDF computed from EST *GE1867* could not be detected. This EST aligned almost completely with the cloned GR-*eng*-2 gene from *G. rostochiensis* (12). Careful examination of the sequence suggested that a 10 bp fragment at the 5' end of the EST, in which a *Taq* I recognition site was located, may have originated from another gene. We therefore assumed that a rare recombination event occurred during construction of the cDNA library. A second band predicted for *GE1867* - 399 nt in length with extensions E+TT/T+TG - was readily identified on gel.

For one particular EST - *GE1782* - a *Taq* I recognition sequence (bold) was found to be partially nested inside the *Eco*R I recognition site (GAATT**CG**A). Contrary to the E+GA group mentioned in Table 1, the TCGA sequence was located at the outside of the TDF. Following the cDNA-AFLP protocol, the cDNA was first digested with *Taq* I and,

as a consequence, the *Eco*R I site was lost. Hence, in this particular case, the predicted TDF was not amplified.

ESTs *GE1349* and *GE1483* were predicted to produce four TDFs. All four TDFs of *GE1349* were located on gels at the predicted size. For *GE1483*, one band was found, but the other three one were smaller than the cut-off size of the gel.

In summary, from a total of 52 TDFs predicted to be produced by E+AN, just one, from EST *GE1133*, could not be located at the predicted size and primer extensions. This minor discrepancy between the GenEST prediction and the bands displayed on gel may be caused by a PCR or sequencing error. It is concluded that predicted TDFs from ESTs can always be traced back on cDNA-AFLP gels, when taking PCR and sequencing errors into account.

Validation of virtual TDFs by sequencing the matching bands

As has been described above, sequencing of more than 100 bands excised from cDNA-AFLP gels showed that the marker-based size estimation was highly accurate. This accuracy was further confirmed by sequencing 24 bands that matched with virtual TDFs. Sequencing of these matching bands revealed no inconsistencies with the computed TDFs. It is therefore concluded that three identifiers - the restriction enzyme recognition sites, the primer extensions and the size of the band - are sufficient to find the corresponding real TDFs on cDNA-AFLP gels.

Expression patterns of virtual TDFs derived from ESTs with putative housekeeping functions

We chose several ESTs with putative housekeeping functions and investigated whether we could find TDFs from these genes on cDNA-AFLP gels at the appropriate positions and with the expected constitutive expression pattern.

EST sequences *GE1373, GE1659* and *GE1699* share high homology (BLASTX expect value < e^{-30}) with elongation factor 1-β from various organisms; *GE99* shares high homology (E-value = e^{-35}) with 40 S ribosomal protein S20 and *GE1409* is likely to be a ribosomal protein L20 homologue (E-value = e^{-41}). These proteins are essential components in protein synthesis and are constitutively expressed in most eukaryotic organisms. For all individual ESTs, GeneEST predicted the generation of at least one TDF. Examination of cDNA-AFLP gels showed discrete bands at the right positions and virtually equal band intensities were observed in the five developmental stages. From one of the developmental stages, the amplification products were cloned and sequenced. The resulting sequences were found to match perfectly with the corresponding EST sequences.

These results show that the expression profiles were in accordance with the predicted functions of the ESTs and that it is feasible to discard or select ESTs by analyzing the expression patterns of the predicted TDFs (see below).

EST to cDNA-AFLP: discarding ESTs on the basis of expression profiles

For many ESTs no function could be inferred from homology searches. About 40% of the ESTs obtained from *G. rostochiensis* were categorized as unknowns and many of these genes seemed to be nematode-specific. Proteins encoded by these nematode-specific genes are presumably important in nematode physiology and a few among them may be

related to parasitism of host plants (9). Examination of the expression profiles of the virtual TDFs provides valuable information whether such ESTs deserve further investigation or not. This is exemplified by ESTs *GE54* and *GE1084*. *GE54* was predicted to produce a TDF with extensions E+TC/T+AG and a size of 138 nt and *GE1084* with extensions E+AC/T+AC and a size of 85. Their corresponding TDFs were readily identified on cDNA-AFLP gels (Fig. 1). Both bands displayed a constitutive expression pattern throughout the five developmental stages. This argues against a direct function of the proteins encoded by these two genes in plant parasitism.

EST to cDNA-AFLP: selection of ESTs on the basis of expression profiles

GE1156 was predicted to produce three TDFs (E+CA/T+TT/65 nt; E+CT/T+GG/73 nt; E+GC/T+AT/82 nt). At each of the predicted positions, a band could be found. The bands in the hatched J2 stage showed the highest intensity. Sequence alignment revealed that *GE1156* was similar to the dorsal gland specific gene GR-*dgl*-2 from the potato cyst nematode (10). GR-*dgl*-2 was previously shown to be specifically expressed in PRD- hatched J2. *In situ* hybridization revealed specific expression of GR-*dgl*-2 in the dorsal gland of the nematode. The proteins produced by this gland may be involved in the induction of a feeding site, a so-called syncytium, in the host plant (13). The protein conceptually translated from the cDNA was predicted to be preceded by a signal peptide for secretion, indicating that this protein might be secreted by the nematode during the infection process.

GE1867 appeared to be identical to GR-*eng*-2, one of the β-1,4-endoglucanases that is secreted by cyst nematodes. *In situ* hybridization showed that GR-*eng*-2 was specifically expressed in the subventral gland (12). Unlike *GE1156*, *GE1867*-derived TDF (E+TT/T+TG/399 nt) showed high expression not only at the H- but also in the (earlier) S-stage. This points to an earlier transcription activation of subventral gland specific genes. The proteins encoded by these genes may be important in the early infection process, *viz.* penetration and migration in the plant root.

cDNA-AFLP to EST: finding (near) full-length cDNAs corresponding to TDFs with relevant expression patterns

The extensions of the *Eco*R I-primer, the extensions of the *Taq* I-primer and the size of a band on cDNA-AFLP gel constitute a unique identifier for a TDF. These parameters can be used to search in the EST database to find an EST that can produce such a TDF. In this way, TDFs with S-H or H-stage specific expression (*i.e.* gene expression just prior to invasion of the plant) were used to search the list of virtual TDFs generated from the EST database.

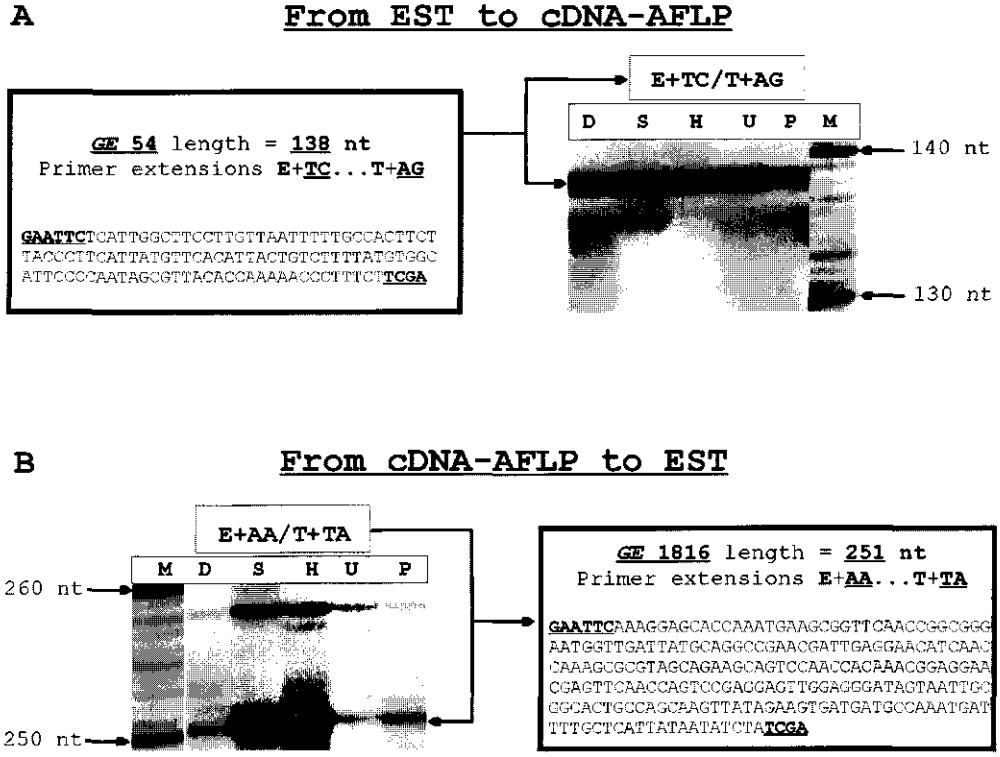


Figure 1. Overview of the bi-directional link between ESTs and cDNA-AFLP expression profiles established by GenEST. Panel A: From EST to cDNA-AFLP expression profiles. The predicted TDF of EST *GE*54 with primer extensions E+TC/T+AG and a size of 138 nt was readily identified on cDNA-AFLP gel. Panel B: From cDNA-AFLP to EST. A band on gel was amplified with primer extensions E+AA/T+TA and the size of this band was 251 nt. These identifiers were used to search the virtual TDF list generated by GenEST. The corresponding EST was identified.

M: molecular ladder; D, S, H, U, P represent five different developmental stages of the potato cyst nematode. D: unhatched J2 in diapause, S: unhatched J2 after diapause, rehydrated for 2 days in water, H: freshly hatched J2 in potato root diffusate, U: developing nematodes (J1) in gravid females two months post-inoculation, P: developing nematodes (J2) in gravid females three months post-inoculation.

One TDF specifically expressed at the H stage with extensions: E+CC/T+CT/ and 137 nt in length matched perfectly with the parameters of the predicted TDF from EST

GE2075. This band was subsequently cloned and sequenced. The sequence showed 99% match to EST *GE2075*. With the help of GenEST, the gene sequence representing this H-stage specific band was extended from just 137 bp to 477 bp (Table 2). By sequencing the original plasmid of EST *GE2075* from the 3'end, the cDNA sequence was extended to 685 bp.

Another band displaying high expression in S and H stages with extensions of E+AA, T+TA and a size of 251 nt matched perfectly to the predicted TDF from EST *GE1816* (Fig. 1). Analysis of this EST sequence revealed that the longest open reading frame contained 107 amino acids. This gene had no significant homology with existing genes in public databases. Use of the SignalP program (14) predicted that the protein had a cleavable signal sequence at its N-terminus that presumably targets the mature peptide for secretion. Hence the combination of cDNA-AFLP and EST analysis has allowed us to identify this gene as worthy of further study for its potential role in nematode parasitism of plants.

These two examples illustrate another benefit of combining cDNA-AFLP and EST, which is to facilitate the cloning of full-length cDNA sequences from which interesting TDFs are derived. The corresponding gene can be readily identified from the EST database even without cloning and sequencing the TDF displayed on gel. Once the corresponding EST is identified, obtaining (nearly) full-length sequence is relatively simple by sequencing the entire cDNA insert from which the EST was originally derived.

In Table 2, four examples of putatively interesting ESTs and their corresponding TDFs are given. In Fig. 1, an overview of the bi-directional link between ESTs and cDNA-AFLP expression profiles established by GenEST is given.

Starting point	Corresponds to	Expression pattern on gel	Homology to
E+AA/T+TA/ 251 nt	GE1816	↑in S and H	Unknown, predicted to have a signal peptide for secretion
E+CC/T+CT /137 nt	GE2075	↑in H	Nematode dorsal gland specific gene GR-*dgl-2*
GE1156	E+CA/T+TT/65 nt E+CT/T+GG/73 nt E+GC/T+AT/82 nt	↑ H (three TDFs, same pattern observed)	Nematode dorsal gland specific gene GR-*dgl-2*
GE1867	E+TT/T+TG/399 nt	↑ in S and H stages	GR-*eng*-2 from potato cyst nematode

Table 2. Genes selected on the basis of a combinatorial use of EST sequences and cDNA-AFLP data from the potato cyst nematode *Globodera rostochiensis*. In each direction, the bi-directional program GenEST allowed for the selection of putatively pathogenicity-related genes out of hundreds of EST sequences and expression profiles of thousands of genes.

DISCUSSION

In this paper, we present an efficient and bi-directional link between (partial) cDNA sequences and gene expression profiles as generated by cDNA-AFLP. A program called GenEST establishes this link. The added value of combining cDNA sequence information and cDNA-AFLP profiles is illustrated for one particular case, namely the search for putative pathogenicity factors from a plant parasitic nematode, *G. rostochiensis*. On the one hand, GenEST enabled us to find the expression profile of a given EST among the profiles of thousands of genes. The other way around, it allowed for a quick extension of TDFs by searching for the corresponding EST(s). As we have shown that the restriction enzyme recognition sites, the primer extensions and the size of the band displayed on cDNA-AFLP gel constitute a unique set of identifiers for a TDF, the corresponding (nearly) full-length cDNA can be identified even without cloning and sequencing of the TDF of interest. In this way, the bottleneck of identifying the (near) full-length cDNAs in high throughput functional genomics studies using gel-based gene expression monitoring systems can be overcome. Since database similarity searches are more robust when using longer sequence fragments, the possibility of moving directly from a short TDF to a much longer EST may be very useful in further characterizing the putative function of a gene.

Use of GenEST for the selection of putative pathogenicity factors

Selection on the basis of expression profiles of the 228 virtual TDFs that were produced by *in silico* restriction of 985 ESTs with *Eco*R I and *Taq* I revealed four putative pathogenicity related genes. One was a known gene encoding a cellulase (12). GE*2075* and GE*1156* displayed strong homology with a nematode secretory gland specific gene GR-*dgl-2*, indicating a possible role in the parasitism of host plants. GE*1816* is a novel gene. Its function will be studied further to reveal its role in nematode infection process. It is noted that this is the result of a small-scale pilot experiment only. Already at this scale, the value of GenEST that combines two high-throughput technologies is evident: four putatively pathogenicity-related genes were selected out of hundreds of EST sequences and expression profiles. The applicability of this freely available tool is broad as long as the expression of genes of interest is strictly limited, either spatially or temporarily.

Further applications of GenEST

Contrary to EST approaches, the cDNA-AFLP technique is not biased toward abundant transcripts and does not involve a selection on insert size. Moreover, there is no unwanted selection due to intolerance of *E. coli* towards a subset of the inserts. To estimate the fraction of genes not tagged by ESTs, Penn et al. (15) have spotted 10,000 predicted ORFs from the human genome on an cDNA-array and monitored the expression of these ORFs under various conditions. They concluded that potentially up to 30% of the genes in the human genome will not be discovered by an EST approach. A similar experiment could be performed by linking cDNA-AFLP and EST data with GenEST. Failure to find a good match for a TDF shown on cDNA-AFLP gel in a large-scale EST database is informative. The corresponding gene is presumably a novel gene expressed at a low level or a small gene or a gene refractory to cloning in *E. coli*. An advantage of our approach

is that ESTs and cDNA-AFLP are not linked physically as is the case for cDNA arrays. This avoids the amplification and spotting thousands of EST clones, saving huge logistic efforts.

Besides generating restriction patterns of sequences, GenEST can also be used to find other sequence motifs in a large data set, a process which is often too time-consuming to be done manually. To illustrate this application, GenEST was used to predict the occurrence of *trans*-spliced leader sequences from a database composed of approximately 1,000 ESTs of the root knot nematode *Meloidogyne incognita*. In many nematode species, up to 70% of the mature mRNAs are *trans*-spliced with a 22 nt leader sequence to the 5' end of the mRNAs (16). When *Eco*R I recognition sequence in the command file is replaced with the *trans*-spliced leader sequence, all the ESTs containing this sequence can be quickly identified using GenEST. This information can be used to estimate the fraction of full-length cDNAs present in a library and to check whether the encoded open reading frames start with a peptide signal for secretion. This latter process could be further streamlined by establishing a link between GenEST and search algorithms such as SignalP.

AFLP techniques have been used extensively in genetic mapping in various organsims and a large number of AFLP markers associated with genes of interest have been identified (17,18). Such markers combined with a fully sequenced genome (*e.g. Arabidopsis thaliana* (19)) could facilitate the efficient cloning of target genes. To this purpose, GenEST can be adapted to assist in the identification of the physical locus of an interesting gene by using the identifiers of appropriate AFLP markers.

Further improvement of the EST coverage

Only 16% (159/985 * 100%) of the 985 ESTs were digested *in silico* by *Eco*R I and *Taq* I. To increase the percentage of ESTs from which virtual TDFs are obtained, a set of alternative rare cutters including *Nco* I, *Kas* I and *Ase* I are currently being used in combination with *Taq* I. With three additional primer combinations, more than half of the EST sequences ($1- (1-0.16)^4 = 50.2\%$) will produce at least one virtual TDF, which could be identified on cDNA-AFLP gels. To further increase the coverage of the EST population, cDNA-AFLP can be performed with two frequent cutters. Alternatively, cDNAs could be digested with a frequent cutter only and ligated to the corresponding adapter. Subsequently, 3'-anchored cDNA-AFLP could be performed using an oligo(dT) primer in combination with the rare cutter adapter primer. This approach may be especially useful with organisms for which the entire genome has been sequenced or for which large scale 3'-end EST sequencing has been performed. Moreover, by increasing the fraction of full-length cDNA sequences, the chance to find at least one corresponding TDFs on gel would improve significantly.

As shown in this study, the ability to switch between sequence data and expression profiles revealed by cDNA-AFLP and *vice versa* is a very powerful approach to select genes for further research. This novel link provided by GenEST will be useful for functional genomics studies and is applicable for any organism where differentially expressed genes are of interest. The source code of GenEST program is freely downloadable. The user is free to modify the existing program according to his / her own wishes.

ACKNOWLEDGEMENTS

This work was supported by European Community grants BIO4-CT96-0318 (LQ) and QLK5-1999-01501 (LQ), FAIR5-CT97-3565 (HP) and by Scottish Executive Rural Affairs Department project number SCR/514/98 (JTJ).

REFERENCES

1. Goffeau,A., Barrell,B.G., Bussey,H., Davis,R.W., Dujon,B., Feldmann,H., Galibert,F., Hoheisel,J.D., Jacq,C., Johnston,M. *et al.* (1996) Life with 6000 genes. *Science*, **274**, 563-567.
2. The *C.elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science*, **282**, 2012-2018.
3. Blaxter,M. (1998) *Caenorhabditis elegans* is a nematode. *Science*, **282**, 2041-2046.
4. Velculescu,V.E., Zhang,L., Vogelstein,B. and Kinzler,K.W. (1995) Serial analysis of gene expression. *Science*, **270**, 484-487.
5. Schena,M. (1996) Genome analysis with gene expression microarrays. *Bioessays*, **18**, 427-431.
6. Lockhart,D.J., Dong,H., Byrne,M.C., Follettie,M.T., Gallo,M.V., Chee,M.S., Mittmann,M., Wang,C., Kobayashi,M., Horton,H. *et al.* (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat.Biotechnol.*, **14**, 1675-1680
7. Liang,P. and Pardee,A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science*, **257**, 967-971.
8. Bachem,C.W., van der Hoeven,R.S., de Bruijn,S.M., Vreugdenhil,D., Zabeau,M. and Visser,R.G. (1996) Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. *Plant J.*, **9**, 745-753.
9. Popeijus,H., Blok,V.C., Cardle,L., Bakker,E., Phillips,M.S., Helder,J., Smant,G. and Jones,J.T. (2000) 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.
10. Qin,L., Overmars,H., Helder,J., Popeijus,H., Rouppe Van Der Voort,J.N.A.M., 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*. *Mol. Plant-Microbe Interact.*, **13**, 830-836.
11. Rouppe van der Voort,J.N.A.M., van Eck,H.J., van Zandvoort,P., Overmars,H., Helder,J. and Bakker,J. (1999) Linkage analysis by genotyping of sibling populations: a genetic map for the potato cyst nematode constructed using a "pseudo-F2" mapping strategy. *Mol.Gen.Genet.*, **261**, 1021-1031.
12. Smant,G., Stokkermans,J.P., Yan,Y., de Boer,J., Baum,T.J., Wang,X., Hussey,R.S., Gommers,F.J., Henrissat,B., Davis,E.L. *et al.* (1998) Endogenous cellulases in animals: isolation of beta-1, 4-endoglucanase genes from two species of plant-parasitic cyst nematodes. *Proc.Natl.Acad.Sci.U.S.A.*, **95**, 4906-4911.

13. Williamson,V.M. and Hussey,R.S. (1996) Nematode pathogenesis and resistance in plants. *Plant Cell*, **8**, 1735-1745.
14. Nielsen,H., Engelbrecht,J., Brunak,S. and von Heijne,G. (1997) Identification of prokaryotic and eukaryotic signal peptide and prediction of their cleavage sites. *Protein Eng.*, **10**, 1-6.
15. Penn,S.G., Rank,D.R., Hanzel,D.K. and Barker,D.L. (2000) Mining the human genome using microarrays of open reading frames. *Nature Genetics*, **26**, 315-318.
16. Blaxter,M. and Liu,L. (1996) Nematode spliced leaders--ubiquity, evolution and utility. *Int.J.Parasitol.*, **26**, 1025-1033.
17. Vos,P., Hogers,R., Bleeker,M., Reijans,M., van der Lee,T., Hornes,M., Frijters,A., Pot,J., Peleman,J., Kuiper,M. *et al.* (1995) *Nucleic Acids Res.*, **23**, 4407-4414.
18. Buschges,R., Hollricher,K., Panstruga,R., Simons,G., Wolter,M., Frijters,A., van Daelen,R., van der Lee,T., Diergaarde,P., Groenendijk,J. *et al.* (1997) The barley Mlo gene: a novel control element of plant pathogen resistance. *Cell*, **88**, 695-705.
19. Kaul,S., Koo,H.L., Jenkins,J., Rizzo,M., Rooney,T., Tallon,L.J., Feldblyum,T., Nierman,W., Benito,M.I., Lin,X.Y. *et al.* (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, **408**, 796-815.

Chapter IV

Identification of an expansin-like protein from the potato cyst nematode

---

Qin,L., Kudla,U., Goverse,A., Popeijus,H., Roze,E., Overmars,H., Smant,G., Schots,A., Bakker,J. and Helder,J.
Identification of an expansin-like protein from the potato cyst nematode.
Submitted

Identification of an expansin-like protein from the potato cyst nematode

Abstract

Expansins are extracellular plant proteins that unlock networks of cell wall polymers, thereby increasing cell wall extensibility. Here, we report the identification of *Gr-ExpB1*, a homologue of plant expansins from the potato cyst nematode *Globodera rostochiensis*, an obligatory plant parasite. The C-terminal part shows significant homology to the catalytic domain of a β-expansin-like protein from *Nicotiana tabacum* (38% identity, 47% similarity). The N-terminal part of GR-EXPB1 closely resembles a number of nematode cellulose binding domains (CBDs) (40-42% identity, 53-57% similarity), previously identified as family II CBDs. The nematode expansin gene was specifically up-regulated in infective second stage juveniles and expressed only in the subventral secretory glands. Moreover, this nematode protein was predicted to be preceded by a signal peptide for secretion. Expression of the putative pathogenicity factor *Gr-ExpB1* in tobacco resulted in a significant increase in β-expansin activity and expansin activity was also detected in homogenate of infective juveniles. In combination with cell wall-degrading enzymes, this nematode expansin could facilitate intracellular migration through the plant root.

Introduction

Feature of plant cell wall

Plant cell walls are important structures in maintaining the integrity of plant cells, regulating their growth and defending the plant against invading pathogens. The backbone of the primary cell walls is a network of interwoven crystalline cellulose microfibrils. These microfibrils are embedded in a matrix of pectins and hemicelluloses. These matrix polysaccharides associate with cellulose chains mainly by non-covalent links [1].

Plant cell wall loosening

Cell wall loosening typically coincides with the differentiation of plant cells. As the extensibility of cellulose microfibrils is limited, an active loosening of the primary cell wall is a prerequisite for the expansion of small meristematic cells. Both auxin and a low pH increase the extensibility of primary cell walls, a phenomenon referred to as 'acid growth'. Active wall proteins are necessary for low pH-induced cell wall extension [2]. These proteins were identified and named expansins [3].

Expansins

Two, remarkably diverged groups of expansins can be distinguished, α- and β-expansins. Both types increase the extensibility of cell walls without hydrolyzing the wall polymers. *Arabidopsis* has approximately 25 α-expansin genes, and similarly large families have been identified in rice, tomato and tobacco [4]. β-expansins were originally identified as pollen allergens. In 1997, they were shown to possess a potent plant cell wall-loosening effect [5]. The majority of the β-expansins cloned so far was isolated from monocots. It is proposed that expansins move along the cellulose

surface and disrupt the non-covalent bonding between cellulose microfibrils and matrix glycans that are attached to the microfibrils. Expansins are extracellular proteins and consist of two distinct domains. The N-terminus is the putative catalytic domain with limited sequence similarity to family-45 endoglucanases and the C-terminus contains several aromatic residues on the surface that may function as a cellulose-binding domain [4]. They are thought to be involved in a wide range of physiological processes, including plant cell growth, pollen tube penetration of the pistil, cell wall disassembly during fruit ripening, leaf primordium initiation and abscission [6]. So far, the distribution of expansins seems to be limited to land plants.

Plant parasitic nematode

Unlike most other animals, plant cell wall degradation by the potato cyst nematode *Globodera rostochiensis* - an obligatory plant parasite - does not depend on the presence of a symbiont. β-1,4- endoglucanases and a pectate lyase are produced in the subventral esophageal glands of the infective juvenile and these enzymes are secreted into plants to facilitate the migration of the nematode in the plant root [7-9]. The subventral esophageal glands are activated upon hatching and they remain highly active during migration. Thereafter, the density of the secretory granules in the ampullae of these glands decreases. Hence, activity of the subventral glands is associated with migration of the nematode through the plant root.

In this paper, we describe the identification of expansin-encoding gene, *Gr-ExpB1*, from the potato cyst nematode. This gene is abundantly expressed in the subventral secretory glands of this nematode and, therefore, probably related to migration in the plant root. Expansin activity was shown in crude protein extracts from infective juveniles of *G. rostochiensis*. Moreover, expression of the putative expansin *Gr-ExpB1* in tobacco resulted in greatly enhanced expansin activity. Our results show that the distribution of expansins is not *per se* limited to the Embryophyta. To the best of our knowledge, the potato cyst nematode *G. rostochiensis* is the first example of an animal producing an expansin.

Material and Methods

Nematode culture

G. rostochiensis pathotype Ro1-Mierenbos was used in all experiments. Nematodes were reared and hatched as described previously [10].

EST sequence

A cDNA library was made from potato root diffusate-hatched, second stage juveniles (J2) of the potato cyst nematode *G. rostochiensis* [7]. ESTs were obtained by randomly sequencing library inserts from the 5' end [11].

cDNA-AFLP analysis

cDNA-AFLP analysis was performed essentially as described [12]. RNA was extracted from five developmental stages of *G. rostochiensis*: (D-dormant) dehydrated unhatched second stage juveniles (J2's) in cysts (in diapause); (S-soaked) rehydrated unhatched J2's in one-year-old cysts after exposure to sterile tap water for two days; (H-hatched) pre-parasitic J2's (dry cysts were incubated in sterile tap water for one week, tap water is replaced by potato root diffusate (PRD) in the second

week); (U- undifferentiated into J2's) developing nematodes in gravid females two months post-inoculation; (P-pre-diapause) developing nematodes in gravid females three months post-inoculation. In the restriction enzyme combination *Eco*RI – *Taq*I originally used by Qin *et al.* [12], the rare cutter *Eco*RI was replaced by *Kas*I. The adapters and the primers were replaced accordingly. The primers corresponding to the *Kas*I and *Taq*I adapters (without extension) are indicated as 'K' and 'T', respectively. *Kas*I/*Taq*I fragments were amplified with the following primer combinations: K+T/T+TC.

Sequence analysis

DNA sequences were analyzed using the Laser gene software package DNASTAR 4.00 (Madison, WI, USA). Database search was done using Blast service at the National Center for Biotechnology Information (NCBI) and WU-Blastp2 at European Bioinformatics Institue (EBI). The computer algorithm SignalP [13] was used to predict the presence of a signal peptide for secretion and the corresponding putative cleavage site.

Hydrophobic cluster analysis (HCA)

HCA plots were made by 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 squares, 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 [14].

Southern blot

The genomic DNA was isolated as described [15].
Primers:

gDNACBPF2 (5'-GAGCTCCTCTGAAGCAATTC-3')
gDNACBPR1 (5'-AACACTGTATAAACCTTTATGCAATT-3')

annealing to the extreme 5' and 3' ends of the cDNA were used to amplify a 2.4 kb genomic fragment. The template in this PCR reaction is genomic DNA.
Primers:

CBPfor1 (5'-ATGGTTTATTGAAAAATTCGTTG-3')
CBPrev1 (5'-CCTGTCTCGACAAAAGAGTCC-3')

were used to amplify a fragment of 409 bp (to the 3' end of CBD) from the genomic DNA of *Gr-ExpB1*. This fragment was then used as a template for DIG labeling (Roche Diagnostics, Mannheim, Germany). About 2.5 µg of genomic DNA was digested each by *Eco* RI, separated on 0.8% percent agarose gel and blotted as described [16]. Hybridization was performed at 40 °C overnight in DIG Easy Hyb solution (Roche Diagnostics). The filter was washed two times in 0.1xSSC/0.1% SDS solution before detection. Alkaline phosphatase activity was detected using DIG Luminescent Detection Kit (Roche Diagnostics) and the chemiluscent signal was exposed on X-ray films.

***In situ* hybridization**

Nematode fixation, hybridization and detection steps were performed as described [12]. For *Gr-ExpB1* probe preparation, EST clone GE1163 was amplified with the specific primers:

$WG\text{-}CBP_{for}$ (5'-ATTCTGTGTTTGTTGTGCCTTTTG)
$WG\text{-}CBP_{rev}$ (5'-GCTCCGCCACAGACATCG).

For *Gr-expB2* probe preparation, EST clone GE1032 was amplified with $poly(dT)_{16}$ and M13 primer (GTAAAACGACGGCCAG), which anneals to the pcDNAII (Invitrogen, Leek, the Netherlands) vector sequence just upstream of the 5' end of GE1032. The amplified products were checked on a 1% agarose gel. The antisense probes were made by asymmetric PCR in the presence of DIG-dUTP (Roche Diagnostics) using either the $WG\text{-}CBP_{rev}$ primer for Gr-expB1 or $poly(T)_{16}$ primer for Gr-expB2. The labeling reaction was run in a 20 µl volume with 20 ng of the amplified product from the first PCR as a template. The DIG-labeled probes were purified through a G50 Mini Quick Spin DNA column (Roche Diagnostics) and 10 µl TE buffer was added. In each hybridization reaction, 30 µl of labeled probe was used. Alkaline phosphatase activity was detected by the addition of X-phosphate and NBT (Roche Diagnostics), and labeled infective juveniles were observed under a Leica inverted microscope (Leica, Deerfield, IL, USA).

Expression of Gr-expB1 in tobacco and protein extraction

For heterologous expression in plants, a cDNA encoding the full-length protein of GR-EXPB1 was placed under control of Cauliflower Mosaic Virus 35S promoter. This construct was introduced into binary vector pBIN+ and mobilized into *Agrobacterium tumefaciens* strain pMOG 101. As a control, plants were transformed with empty vector pBIN+ [17]. Transformation of tobacco (*Nicotiana tabacum; Samsun nn*) was carried out as described by van Engelen et al [18]. Protein encoded by *Gr-ExpB1* was extracted by grinding tobacco leaves in liquid nitrogen and suspending homogenate in 50 mM sodium acetate pH 4.5. Debris was spun down and extract was used for the extensometer assay as described below.

Cell wall extension assay

Cell wall extension assay was essentially performed as described by Cosgrove et al. [5]. Briefly, a 5-cm segment of an etiolated wheat seedling was cut from the apical growing region. Subsequently, it was frozen at –20°C, thawed, abraded with carborundum slurry, heat inactivated and clamped in an extensometer loaded with 20 g constant weight. The inactivated wheat hypocotyl segment was first incubated in 1 ml 50 mM sodium acetate buffer (pH 4.5) for 45 min, then the buffer was replaced by either 1 ml nematode total homogenate solution or 1 ml tobacco leaf extract. The total nematode homogenate was obtained by grinding approximately 200,000 J2's in 1 ml 50 mM sodium acetate buffer (pH 4.5). In the control experiment, 1 ml 50 mM sodium acetate buffer (pH 4.5) was used during the measurement. The cell wall extension was recorded in a position transducer attached to the clamp. The experiments were repeated four times.

RESULTS

Two abundantly expressed transcripts from infective juveniles

Random sequencing from the 5' ends of 985 inserts from a cDNA library of infective juveniles of the potato cyst nematode *G. rostochiensis* revealed four ESTs - GE1163 (518 bp long), GE1336 (603 bp), GE1055 (615 bp) and GE2019 (414 bp) that overlapped with each other extensively. To obtain a full-length cDNA, these clones were sequenced from the 3' end. This resulted in a 1061-bp cDNA contig containing a 34-bp polyA tail. The longest open reading frame (ORF) contained 816 bp, encoding a protein of 271 aa. The molecular weight of the putative protein translated from the cDNA sequence was 28.2 kDa with a pI value of 8.65. The protein was predicted to have a signal peptide for secretion at position 1-25 [13]and the cleavage site was located between 25 and 26 (VLA↓SV). No glycosylation site was found using ScanProsite and NetOGlyc.

As many as 27 ESTs were found to share significant homology with the 3' half of the aforementioned 28.2 kDa protein. These sequences were assembled into one single contig, the full-length cDNA was shown to contain 684 bp (including a 27-bp polyA tail). At the 5' end of the assembled sequence, a motif of 'CCAAGTTTGAG' was found. This motif corresponded to the 3' half of the 22 nt SL-1 sequence which has been identified in various nematode species including potato cyst nematode [19;20]. Hence, it is most likely that this gene was *trans*-spliced by SL-1. The translated protein consisted of 165 aa and the molecular weight was predicted to be 17.8 kDa with a pI value of 8.03. This apparently abundant protein was predicted to have a signal peptide for secretion, and the cleavage site was located between 21 and 22 (NES↓CM).

Temporal and spatial expression pattern of two abundantly expressed transcripts

Putative pathogenicity factors from the potato cyst nematode are likely to be expressed solely or predominantly in the infective life stage, the pre-parasitic second stage juvenile. On the basis of primary sequence data, a *Kas*I / *Taq*I digestion of the 28,2 kDa protein-encoding cDNA should result in a transcript-derived fragment (TDF) of 137 nt. Three selective nucleotides (K+T and T+TC) were used for the cDNA-AFLP-mediated visualization of the expression pattern of this TDF. This particular cDNA was shown to be strongly and specifically up-regulated in potato root diffusate-exposed infective juveniles (Fig.1). The corresponding band was excised from the gel, cloned and sequenced. The sequence matched perfectly with the original cDNA sequence.

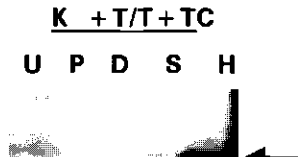


Fig. 1. The expression pattern of *Gr-expB1* from the potato cyst nematode *G. rostochiensis* as displayed by cDNA-AFLP. The corresponding cDNA fragment was amplified using K+T and T+TC. The arrow-pointed band was excised from gel, cloned and sequenced. The sequence matched perfectly with the original cDNA sequence. These stages are indicated by capitals. U: developing nematodes (J1) in gravid females two months post-inoculation, P: developing nematodes (J2) in gravid females three months post-inoculation. D: unhatched J2 in diapause, S: unhatched J2 after diapause, rehydrated for 2 days in water, H: freshly hatched J2 in potato root diffusate.

To investigate where these two genes were expressed in the nematode, *in situ* hybridization was performed on preparasitic second stage juveniles. Antisense probes of both genes hybridized specifically to the subventral gland of the nematode with strong intensities. The signal of the 17.8 kDa protein-encoding cDNA was stronger and more diffused as compared to its 28.2 kDa equivalent (Fig. 2).

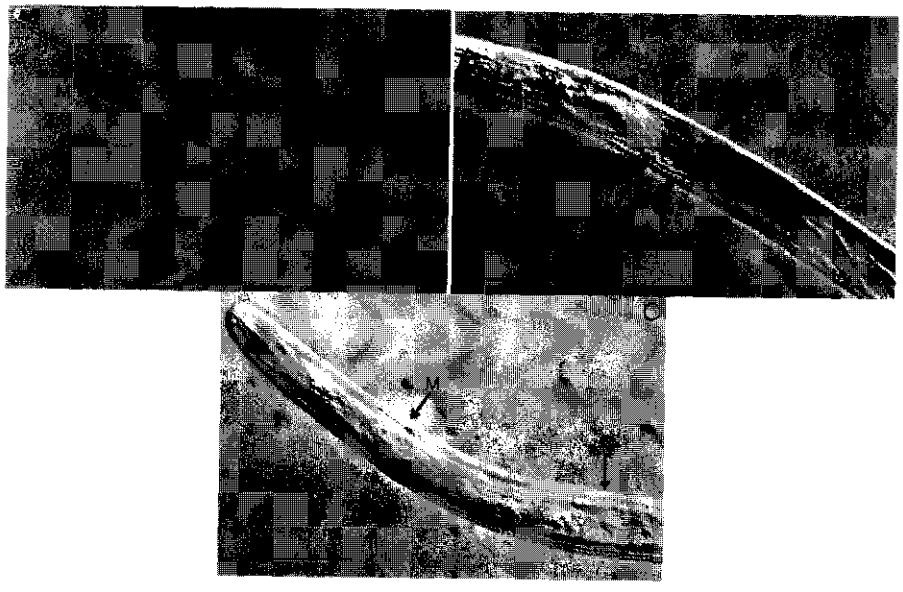


Fig. 2. *In situ* hybridization using probes derived from *Gr-expB1* and *B2* from G. rostochiensis. Labeling was done in preparasitic J2. A: labeling pattern using a *Gr-expB1* antisense probe. B: labeling pattern with a *Gr-expB2* antisense probe. C: A typical image of a sense probe labeling. Arrows point at the subventral glands (S) and the metacorpus (M) respectively. Scale bar = 20 µm

Sequence analysis of the 28.2 kDa putative pathogenicity factor

The program WU-blastp (matrix: blosum62) was used to do homology searches for the three functionally distinct regions present in the predicted mature protein:

(1) The N terminal part of the mature protein (aa residues 26-118) shared high homology with the cellulose binding domains (CBDs) of nematode cellulases. The

overall identity with CBDs of GR-ENG1 (AF056110), GT-ENG1 (AF182392) and MI-ENG1A (AF323087) from *G. rostochiensis*, *G. tabacum solanacearum* and *Meloidogyne incognita* was 42%, 41% and 40% respectively, while E values ranged from 2.0 e-12 to 6.7 e-10. These nematode CBDs all show significant similarity to type II CBDs (Prosite: PDOC00485; Pfam: PF00553). This type is previously only found in bacterial enzymes, which has a typical size of 90-108 aa [21].

(2) Interdomain linker extends from amino acids residue 119 to 149. The linker is rich in G, S and P, residues that enhance flexibility between two functional parts of the protein [22][23].

(3) The C-terminal part of the 28.2 kDa putative pathogenicity factor (aa 150- 271) showed high homology with a conceptually translated unknown gene *ORF 11* from *Streptomyces lavendulae* (AF127374) (36% identity, 50% similarity, E value = 4.6 e-11) [24]. Significant homology was also found with the C-terminal part of the *map-1* gene (AJ278663), which encoded a putative avirulence protein secreted from the amphids of the root knot nematode *Meloidogyne incognita* (35% identity, 49% similarity, E value = 2.1 e-10) [25]. Both proteins were predicted to be preceded by a signal peptide for secretion.

The C-terminal part of this protein also shared significant homology with a β-expansin-like protein (called PPAL) from *Nicotiana tabacum* (AAG52887; 38% identity, 47% similarity, E value of 2.2 e-5) and a putative β-expansin from *Arabidopsis thaliana* (O04484; 36% identity, 46% similarity, E value of 0.0006). The homology between this 28.2 kDa protein and β-expansins is restricted to the N-terminal putative catalytic domain. Comparison between the 28.2 kDa protein and expansins revealed a series of conserved cysteine residues. In addition to this, a number of amino acid motifs indicated to be typical for the catalytic domain of α- and β-expansins [4] were represented in the 28.2 kDa putative pathogenicity factor (Fig. 3). Based on the fact that this protein showed a remarkable number of expansin characteristics, it was decided to indicate this 28.8 kDa protein provisionally as 'GR-EXPB1' and the corresponding gene as '*Gr-ExpB1*'.

```
                6  12 14-15 21   26   30          56 59-60 62   67   90   104  110-112   117-118 134
GR-EXPB1:       F  TFYG  AG  RGACGLD LFK G  VDAC  ICK  CP  CT  HVDLSI  AF    G
Nt-β-exp:       F  TWYG  AG  GGACGYG LFK G  CGAC  ICK  CP  CG  HFDLSG  AFG   G
At-β-exp:       W  TGYG  SD  GGACGYG LFK G  CGAC  ACS  CP  CV  HFDLSG  AFG   G
```

Fig. 3. Comparison of GR-EXPB1 with the putative catalytic domains of two other β-expansin-like proteins: Nt-β-exp is a β-expansin-like protein from *Nicotiana tabacum* (AAG52887) and At-β-exp is a putative β-expansin from *Arabidopsis thaliana* (O04484). The underlined amino acid residues refer to the conserved residues between α- and β-expansins as proposed by Cosgrove [reference 4, Figure 1]. The numbers indicate the position of the residues in the original figure.

Sequence analysis of the 17.8 kDa putative pathogenicity factor

At the nucleotide and at protein level the 17.8 kDa encoding cDNA and *'Gr-ExpB1'* showed homologies of 84% and 75%, respectively. Therefore, this 17.8 kDa encoding cDNA was provisionally indicated as *'Gr-ExpB2'*. Although the N-terminal part of 'GR-EXPB1' corresponding to the CBD region was completely absent in 'GR-EXPB2', the signal peptide sequence was remarkably similar; 12 aa out of first 18 aa were identical. Just like 'GR-EXPB1', 'GR-EXPB2' showed significant homology to the C-terminal part of the *map-1* gene of the root knot nematode *M. incognita* (36% identity, 51% similarity, E value = 1.1 e-13) and to *ORF 11* from *S. lavendulae* (29% identity, 46% similarity, E value = 6.5e-8). The homology of 'GR-EXPB2' with β-expansins is lower as compared to 'GR-EXPB1'.

Secondary structure prediction using hydrophobic cluster analysis (HCA)

To check whether the observed homology at the protein sequence level between the nematode expansin-like proteins and plant expansins is also reflected at the protein structure level, HCA plots of 'GR-EXPB1' was compared to OS-EXPB5 (a rice putative β-expansin, accession number AF261273) and CS-EXP1 (an α-expansin from cucumber, accession number U30382). Analysis of the position and shape of the hydrophobic clusters revealed eight conserved hydrophobic clusters in all protein sequences (Fig. 4). Some of the conserved regions shown in Fig. 3 also fall into similar hydrophobic clusters. This analysis indicated that the secondary structure of 'GR-EXPB1' is similar to those of plant expansins.

Southern blot

To study the genomic organization of these genes, Southern blot was performed with probes derived from the C-terminal region of *'Gr-ExpB1'*. Using a 409-bp genomic sequence as a probe, five bands were observed on the blot, indicating there might be more related genes in the genome of *G. rostochiensis* (Fig. 5).

Cell wall extension assay

The striking number of expansin characteristics of the 28.2 kDa nematode protein, 'GR-EXPB1', prompted us to do an cell wall extension assay on homogenate of infective second-stage juveniles of this potato cyst nematode. Remarkably, addition of nematode homogenate induced rapid extension of heat-inactivated cell walls of wheat hypocotyls (Fig. 6). This activity is specific and characteristic for (plant) expansins. In control experiments with the relevant buffer (50 mM sodium acetate), the extensional rate was significantly lower (data included in Fig 6).

Subsequently, *'Gr-ExpB1'* was expressed constitutively in tobacco. Mature leaf material from four independent transformants was homogenized in 50 mM sodium acetate and the expansin activity in the soluble protein fraction was compared with the soluble protein fraction from plants transformed with the empty pBIN+ vector. Protein extracts from all four *'Gr-ExpB1'*-harbouring plants were shown to have a significantly higher expansin activity as compared to control plant extracts (Fig. 7). Based on these results it was decided to change the provisional name of the 28.2 kDa putative pathogenicity factor, *'Gr-ExpB1'*, into *Gr-ExpB1*. It remains to be established whether *'Gr-ExpB2'* is coding for a nematode expansin.

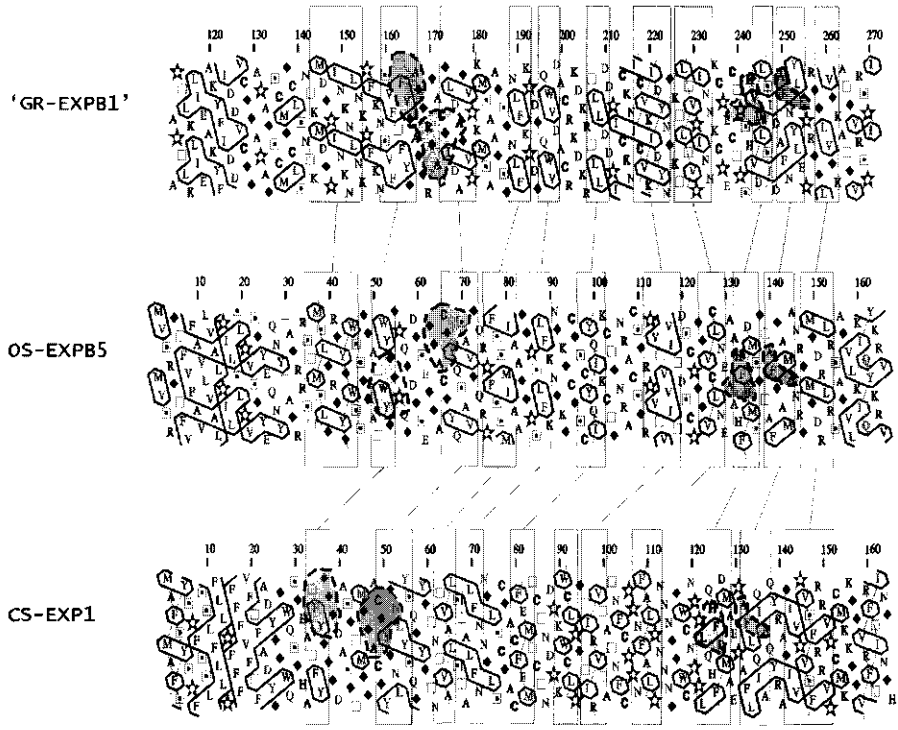


Fig. 4. Hydrophobic cluster analysis of the C-terminal part of 'GR-EXPB1', OS-EXPB5 (a rice putative β-expansin, accession number AF261273) and CS-EXP1 (an α-expansin from cucumber, accession number U30382). Analysis of the position and shape of the hydrophobic clusters revealed eight conserved hydrophobic clusters in all protein sequences. Shaded areas are the conserved blocks as shown in Fig. 3.

DISCUSSION

To date, expansins were thought to be typical plant proteins involved in plant cell wall loosening. Here, it is shown that an obligatory plant parasite, the potato cyst nematode *G. rostochiensis*, is able to produce and secrete expansin. Apart from the homologies with a β-expansin from *N. tabacum* and a putative β-expansin from *A. thaliana*, expansin activity was shown in homogenate of pre-parasitic second stage juveniles, the only infective stage of this nematode. Moreover, constitutive expression of *Gr-ExpB1* in transgenic tobacco resulted in a strongly enhanced activity in tobacco leaf

extracts. To the best of our knowledge, no expansin genes have been cloned outside land plants up to now. It is quite remarkable that a plant-feeding animal can also produce expansin.

Function of nematode expansin

Recently, potato cyst nematodes were shown to produce β-1,4- endoglucanases and a pectate lyase [7;8]. These cell wall degrading enzymes are synthesized in the activated subventral esophageal glands of second stage juveniles. The subventral esophageal gland activity is enhanced upon exposure to root diffusate of a host plant and the glands remain active during migration in the plant root. Feeding site induction coincides with a sharp decrease in the activity of these glands. *Gr-ExpB1* is specifically expressed in the subventral esophageal glands and the resulting protein possesses a predicted signal peptide for secretion. Expansins can move along the cellulose surface and disrupt the non-covalent bonding between cellulose microfibrils and matrix glycans. In this way, the cellulose microfibrils and pectate glycans would be more accessible for the nematode cell wall-degrading enzymes. Hence, it is likely that the newly identified nematode expansin act synergistically with cellulases and pectate lyase to facilitate the intracellular migration through the plant cells. The degeneration (or deactivation) of the subventral glands during feeding site formation suggests that nematode-produced expansins play - at most - a minor role in this process.

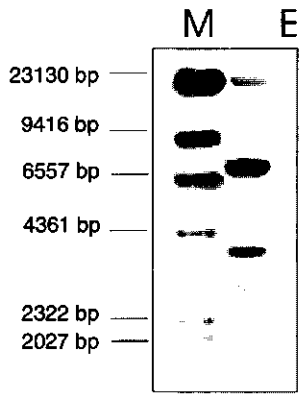


Fig. 5. Genomic Southern blot hybridized with a 409-bp genomic probe derived from *Gr-expB1*. Lane M is molecular weight marker; Lane E is loaded with 2.5 µg nematode genomic DNA digested by *Eco*R I.

The unique structure of a nematode expansin

Plant expansins generally consist of two domains, the N-terminus is the putative catalytic domain with limited sequence similarity to family-45 endoglucanases and the C-terminus contains several aromatic residues on the surface that may function as a cellulose-binding domain [4]. For the cyst nematode expansin GR-EXPB1, the putative domain structure is inverted. The C-terminal part of GR-EXPB1 shared homology with N-terminal putative catalytic domain of β-expansins from tobacco and *Arabidopsis*, whereas the N-terminal part shows resemblance to family II cellulose-binding domains. The putative catalytic domain of GR-EXPB1 resembles the catalytic domains of β-expansins from two dicotyledons, tobacco and Arabidopsis, and is only distantly related to α-expansins. This can be explained by the limited conservation between the α- and β-expansins (only 20% amino-acid identity). β-expansins appear to be highly diverse and abundant in monocots such as rice and maize. This family is underrepresented in dicotyledons. It is not known whether this divergence is related to the evolution of type I and type II primary cell walls present in all Dicotyledonae and several Monocotyledonae, and in Poaceae and some Monocotyledonae including rice and maize, respectively [1]. Potato, tomato and

eggplant - main hosts of the potato cyst nematode - have type I primary cell walls and the effect of β-expansins on the extensibility of these cell wall remains to be investigated.
At the C-terminus, plant expansins lack a clear cellulose-binding domain but structural analysis revealed a domain that could function in polysaccharide binding [4]. Contrary to plants, GR-EXPB1 includes a domain with significant homology to family II CBDs, a CBD family that is normally encountered in cellulases from bacteria. Some family II CBDs can penetrate the cellulose fibers at the surface discontinuities thereby releasing non-covalently attached fragments and uncovering new cellulose chain ends [21;26]). The simultaneous addition of family II CBD and cellulose catalytic domain to a cellulose mixture resulted in an increased hydrolytic activity compared to the catalytic domain alone [26]. The potato cyst nematode harbors a new combination of two functional domains, a family II CBD and a β-expansin catalytic domain. It would be interesting to recombine this type II CBD with the catalytic domains of plant expansins to see whether this would alter the activity and / or the substrate of plant expansins.

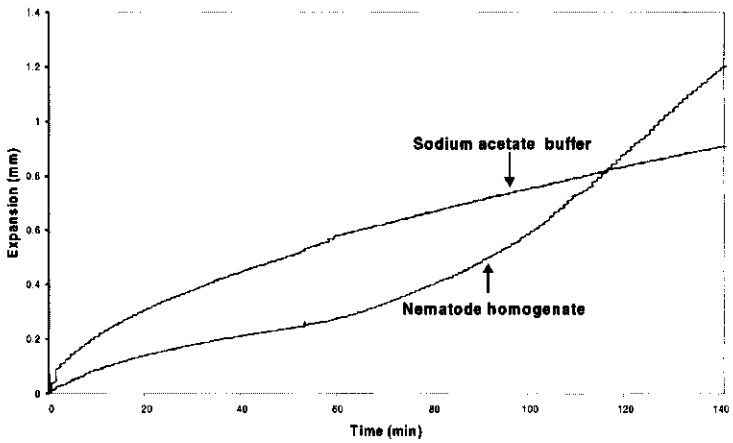


Fig. 6. The inactivated wheat hypocotyl segment was first incubated in 1 ml 50 mM sodium acetate buffer (pH 4.5) for 45 min, then the buffer was replaced by 1 ml nematode total homogenate solution. The total nematode homogenate was obtained by grinding approximately 200,000 J2's in 1 ml 50 mM sodium acetate buffer (pH 4.5). In the control experiment, 1 ml 50 mM sodium acetate buffer (pH 4.5) was used during the measurement. The cell wall extension was recorded in a position transducer attached to the clamp. The figure shows the representative result from four replicates.

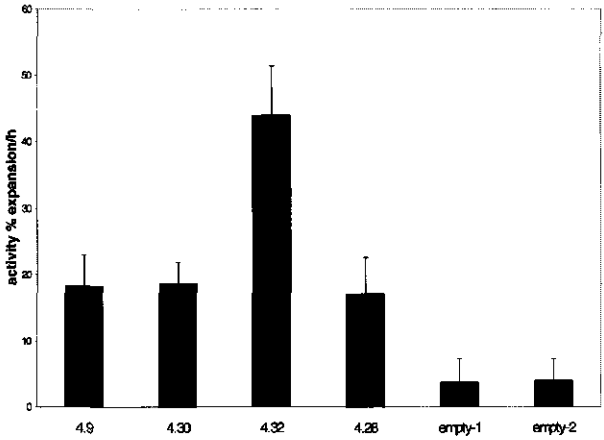


Fig. 7. The inactivated wheat hypocotyl segment was first incubated in 1 ml 50 mM sodium acetate buffer (pH 4.5) for 45 min, then the buffer was replaced by 1 ml tobacco mature leaf extract. Mature leaf material from four independent plants transformed with *Gr-expB1* (4.9, 4.30, 4.32 and 4.28) and two plants transformed with the empty pBIN+ vector (empty-1 and empty-2) were homogenized in 50 mM sodium acetate. The expansin activity in the soluble protein fraction was measured in a position transducer attached to the clamp. The activity was calculated as the difference in extension rates (% increase in length per hour) before and after addition of the extract. The results are means and SE of four replicates.

'Gr-ExpB2'

'*Gr-ExpB2*' is fairly homologous to *Gr-ExpB1* and this shorter protein does not contain the CBD present in the latter. Its homology with the putative plant β-expansins is also lower than that of *Gr-ExpB1* (for example, with the putative β-expansin from *A. thaliana* (AL138646): 29% identity, 40% similarity). Besides, several of the conserved blocks of expansins were missing in this protein. The overrepresentation of partially overlapping '*Gr-ExpB2*' clones in the EST database suggests high expression levels of the corresponding gene in the infective J2 stage. This was confirmed by the strong hybridization signal in the subventral glands observed in the *in situ* experiments. It is postulated that 'GR-EXPB2' plays a role during migration in the plant root and further experiments are needed to determine if it also has expansin activity.

Distribution

Southern blot analysis suggested that there are five *Gr-ExpB*-like genes in the genome of the potato cyst nematode. Hence, *Gr-ExpB1* and *2* are presumably representatives of a small gene family.

Screening of the EST division of Genbank with the putative catalytic domain of GR-EXPB1 revealed homologous cDNA fragments from plant parasitic nematode species such as *M. incognita, Meloidogyne hapla* and *Heterodera glycines* (Ling Qin, unpublished). So, the presence of expansin or expansin- like genes is not necessarily rare among plant parasitic nematodes.

The nematode gene most related to *Gr-ExpB1* (C- terminal part) and -2 is *map*-1 from the root knot nematode *M. incognita*. MAP-1 is an amphid-secreted putative avirulence protein and it might be involved in the early steps of recognition between the nematode and its host plant [25]. Based on the morphology of the nematode, it is unlikely that expression of *Gr-ExpB1* and '2' in the subventral glands would result in an accumulation of these two proteins in the amphids. This difference in localization may be due to the tissue specific expression of different family members. Some of them might be specifically expressed in the subventral glands of the nematode like *Gr-ExpB1* and *'2'*, whereas other genes including *map-1* might only be transcribed in the amphids. It is noted that the expansin-like ESTs from *M. incognita* were highly related but non-identical to *map*-1.

To the best of our knowledge, no animal expansins have been identified yet. Expansin activity was detected in the digestive tract of a snail, *Helix pomatia* [27], but the origin of this activity was unclear. It could have been produced by the snail itself, or, alternatively, by microorganisms present in this tract. The remarkable high homology of the C-terminal part of *Gr-Exp1* with *ORF11* from *S. lavendulae*, a soil borne bacterium and a pathogen of potato, suggests that the ability to produce expansins is not necessarily restricted to plants and plant-parasitic nematodes.

Acknowledgement

We wish to thank J. Nieuwland and Prof. C. Mariani (Nijmegen University, the Netherlands) in helping us performing the cell wall extension assay.

References

1. Carpita NC, Gibeaut DM: **Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth.** *Plant Journal* 1993, **3:**1-30.
2. McQueen MS, Durachko DM, Cosgrove DJ: **Two endogenous proteins that induce cell wall extension in plants.** *Plant Cell* 1992, **4:**1425-1433.
3. Shcherban TY, Shi J, Durachko DM, Guiltinan MJ, McQueen-Mason SJ, Shieh M, Cosgrove DJ: **Molecular cloning and sequence analysis of expansins - a highly conserved, multigene family of proteins that mediate cell wall extension in plants.** *Proc.Natl.Acad.Sci.U.S.A* 1995, **92:**9245-9249.
4. Cosgrove DJ: **New genes and new biological roles for expansins.** *Curr.Opin.Plant Biol.* 2000, **3:**73-78.
5. Cosgrove DJ, Bedinger P, Durachko DM: **Group I allergens of grass pollen as cell wall-loosening agents.** *Proc.Natl.Acad.Sci.U.S.A* 1997, **94:**6559-6564.
6. Cosgrove DJ: **Loosening of plant cell walls by expansins.** *Nature London* 2000, **407:**321-326.

7. Smant G, Stokkermans JP, Yan Y, de Boer J, Baum TJ, Wang X, Hussey RS, Gommers FJ, Henrissat B, Davis EL, Helder J, Schots A, Bakker J: **Endogenous cellulases in animals: isolation of beta-1, 4-endoglucanase genes from two species of plant-parasitic cyst nematodes.** *Proc.Natl.Acad.Sci.U.S.A* 1998, **95:**4906-4911.
8. Popeijus H, Overmars H, Jones J, Blok V, Goverse A, Helder H, Schots A, Bakker J, Smant G: **Degradation of plant cell walls by a nematode.** *Nature London* 2000, **406:**36-37.
9. Wang JF, Bown C, Young LT: **Differential display PCR reveals novel targets for the mood-stabilizing drug valproate including the molecular chaperone GRP78.** *Mol.Pharmacol.* 1999, **55:**521-527.
10. De Boer JM, Overmars HA, Bakker J, Gommers FJ: **Analysis of two-dimensional protein patterns from developmental stages of the potato cyst-nematode, *Globodera rostochiensis*.** *Parasitology* 1992, **105:**461-474.
11. Popeijus H, Blok VC, Cardle L, Bakker E, Phillips MS, Helder J, Smant G, Jones JT: **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* 2000, **2:**567-574.
12. Qin L, Overmars H, Helder J, Popeijus H, van dV, Jr., Groenink W, van Koert P, Schots A, Bakker J, Smant G: **An efficient cDNA-AFLP-based strategy for the identification of putative pathogenicity factors from the potato cyst nematode Globodera rostochiensis.** *Mol.Plant Microbe Interact.* 2000, **13:**830-836.
13. Nielsen H, Engelbrecht J, Brunak S, Von Heijne G: **Identification of prokaryotic and eukaryotic signal peptide and prediction of their cleavage sites.** *Protein Engineering* 1997, **10:**1-6.
14. Lemesle Varloot L, Henrissat B, Gaboriaud C, Bissery V, Morgat A, Mornon JP: **Hydrophobic cluster analysis: procedures to derive structural and functional information from 2-D-representation of protein sequences.** *Biochimie* 1990, **72:**555-574.
15. Curran J, Baillie DL, Webster JM: **Use of genomic DNA restriction fragment length differences to identify nematode species.** *Parasitology* 1985, **90:** 137-144.
16. Sambrook J, Fritsch EF, Maniatis T: *Molecular cloning: A laboratory manual.* New York: Cold Spring Harbor Laboratory Press; 1989.
17. Engelen Fv, Molthoff JW, Conner AJ, Nap JP, Pereira A, Stiekema WJ: **pBINPLUS: an improved plant transformation vector based on pBIN19.** *Transgenic Research* 1995, **4:**288-290.
18. Engelen Fv, Schouten A, Molthoff JW, Roosien J, Salinas J, Dirkse WG, Schots A, Bakker J, Gommers FJ, Jongsma MA, Bosch D, Stiekema WJ: **Coordinate expression of antibody subunit genes yields high levels of functional antibodies in roots of transgenic tobacco.** *Plant Molecular Biology* 1994, **26:**1701-1710.
19. Blaxter M, Liu L: **Nematode spliced leaders--ubiquity, evolution and utility.** *Int.J.Parasitol.* 1996, **26:**1025-1033.
20. Qin L, Smant G, Stokkermans J, Bakker J, Schots A, Helder J: **Cloning of a trans-spliced glyceraldehyde-3-phosphate-dehydrogenase gene from the potato cyst nematode Globodera rostochiensis and expression of its putative promoter region in Caenorhabditis elegans.** *Mol.Biochem.Parasitol.* 1998, **96:**59-67.

21. Tomme P, Driver DP, Amandoron EA, Miller-RC J, Antony R, Warren J, Kilburn DG: **Comparison of a fungal (family I) and bacterial (family II) cellulose-binding domain**. *J Bacteriol.* 1995, **177:**4356-4363.
22. Gilkes NR, Henrissat B, Kilburn DG, Miller RC, Jr., Warren RA: **Domains in microbial beta-1, 4-glycanases: sequence conservation, function, and enzyme families**. *Microbiol.Rev.* 1991, **55:**303-315.
23. Huston JS, Levinson D, Mudgett Hunter M, Tai MS, Novotny J, Margolies MN, Ridge RJ, Bruccoleri RE, Haber E, Crea R, et al : **Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli***. *Proc.Natl.Acad.Sci.U.S.A.* 1988, **85:**5879-5883.
24. Mao Y, Varoglu M, Sherman DH: **Molecular characterization and analysis of the biosynthetic gene cluster for the antitumor antibiotic mitomycin C from Streptomyces lavendulae NRRL 2564**. *Chem.Biol.* 1999, **6:**251-263.
25. Semblat JP, Rosso MN, Hussey RS, Abad P, Castagnone SP: **Molecular cloning of a cDNA encoding an amphid-secreted putative avirulence protein from the root-knot nematode Meloidogyne incognita**. *Molecular Plant Microbe Interactions* 2001, **14:**72-79.
26. Din N, Damude HG, Gilkes NR, Miller-RC J, Warren RA, Kilburn DG: **C1-Cx revisited: intramolecular synergism in a cellulase**. *Proc.Natl.Acad.Sci.U.S.A* 1994, **91:**11383-11387.
27. Cosgrove DJ, Durachko DM, Wang TL, Davies WJ, Pollock CJ: **Autolysis and extension of isolated walls from growing cucumber hypocotyls**. Journal-of-Experimental-Botany. 1994, **45:**1711-1719.

Chapter V

Identification of a RanBPM-like gene family specifically expressed in the dorsal glands of infective juveniles of the potato cyst nematode *Globodera rostochiensis*

Manuscript in preparation

Identification of a RanBPM-like gene family specifically expressed in the dorsal glands of infective juveniles of the potato cyst nematode *Globodera rostochiensis*

Abstract

Cyst nematodes induce large multinucleated feeding cells in the roots of their hosts. Feeding cell formation is accompanied by highly coordinated cell wall degradation and a reactivation of the cell cycle. Secretions from the dorsal esophageal glands are thought to be responsible for the induction and maintenance of these so-called syncytia. cDNA-AFLP was used to identify nematode genes that are up-regulated in the infective stage. This resulted in the identification of a multiple gene family with at least 8 members showing significant homologies to RanBPMs (Ran-Binding Protein in microtubule organizing center) from various organisms. Four members were chosen for *in situ* hybridization and all of them showed specific labeling in the dorsal glands. Three full-length cDNAs were determined and the conceptual proteins were preceded by a predicted signal peptide for secretion. Hence, these RanBPM-like proteins could play a role in the compatible plant-nematode interaction. Based on the conserved nature of the RanBPM-interacting proteins Ran-GTP and γ-TuRC, it is hypothesized that RanBPM-like proteins of nematode origin could change the dynamic instability of microtubules in plant cells and, as a consequence, affect nematode-induced cell cycle reactivation.

Introduction

Plant parasitic nematode

Globodera rostochiensis, the potato cyst nematode, is a highly specialized plant-parasite. Potato, tomato and eggplant are among the main hosts of this sedentary endoparasitic nematode. Apart from these crops, the potato cyst nematode-multiplication is restricted to about 90 species of the genus *Solanum* including *S. dulcamara* (bitter nightshade), *S. sarachoides* (hairy nightshade) and *S. xanti* (purple nightshade).

Feature of a nematode-induced syncytium

Secretion of cell wall-degrading enzymes including β-1,4-endoglucanases [1] and pectate lyase [2] facilitates the intracellular migration of the pre-parasitic second stage juveniles (J2) through the cortical tissue. An inner cortical cell is selected and used as a starting point for the formation of a specialized feeding structure, the syncytium. A highly coordinated breakdown of cell walls finally connects the nematode to the vascular tissues of the host. The resulting conglomerate of up to 200 fused protoplasts shows a high metabolic activity; cytoplasm becomes denser, amount of endoplasmic reticulum increases as well as the number of organelles, whereas the central vacuole disappears completely. Adjacent to the xylem, cell-wall ingrowths develop, suggesting the withdrawal of nutrients from the xylem into the syncytium.

Reactivation of the cell cycle by cyst nematodes

Syncytium formation coincides with an enlargement of both the nuclei and the nucleoli. This process is marked by a re-entry of the cell cycle as indicated by the activation of the promoters of both *cdc2a* and *cyc1*. Several lines of evidence suggest that feeding cell development involves several rounds of DNA endoreduplication shunting the M phase [3]. In addition, d'Almeida-Engler et al. [4] showed that DNA synthesis and progression through the G2 phase are essential for syncytium establishment. The nematode secretions, especially the proteins from the esophageal glands, might contain inhibitors that specifically block the activity of M phase promoting factors and/or induce unknown S-phase-related protein kinases [4]. Recently, a protein fraction in nematode secretions smaller than 3 kDa was found to contain mitogenic activity on plant protoplasts and mammalian T-cell lymphocytes [5]. At present, it is unclear whether this peptide fraction contains the causal agent(s) of feeding site induction.

The identification of RanBPM homologues

cDNA-AFLP - a robust differential mRNA fingerprinting technique - has been used to compare gene expression in infective and non-infective developmental stages of the potato cyst nematode. This approach resulted in the identification of a number of putative pathogenicity factors that might be involved in syncytium formation [6]. Here we report the identification a multiple gene family, which share significant homology with RanBPMs (Ran-Binding Protein in microtubule organizing center) from various organisms. These genes were up-regulated in the infective second-stage juvenile (J2), preceded by a predicted signal peptide for secretion, and specifically expressed in the dorsal gland of the nematode. The putative function of these genes in the induction and maintenance of cyst nematode-induced syncytia is discussed.

Material and Methods

Nematode culture

G. rostochiensis pathotype Ro1-Mierenbos was cultured as described previously [7].

cDNA-AFLP analysis

cDNA-AFLP analysis was performed essentially as described [6]. RNA was extracted from five developmental stages of *G. rostochiensis*: (D) dehydrated unhatched J2s in cysts (in diapause); (S) rehydrated unhatched J2s in one-year-old cysts after exposure to sterile tap water for two days; (H) pre-parasitic J2s (dry cysts incubated in sterile tap water for one week, then in PRD for a second week); (U) developing nematodes (mostly J1s) in gravid females two months post-inoculation; (P) developing nematodes (J2s) in gravid females three months post-inoculation. Dynabeads Oligo $(dT)_{25}$ (Dynal A.S, Oslo, Norway) were used to isolate mRNA. An oligo$(dT)_{12\text{-}18}$ primer was used to synthesize cDNA. The resulting cDNAs were then digested with the restriction enzyme combinations *Eco*RI and *Taq*I or *Nco*I and *Taq*I and ligated to corresponding adapters. The ligated cDNA fragments were subsequently amplified by *Eco*RI and *Taq*I or *Nco*I and *Taq*I primers (indicated as "E", "T" and "N"). The resulting amplification products were separated on polyacrylamide gels.

EST sequence
A cDNA library was made from potato root diffusate-hatched, second stage juveniles (J2) of the potato cyst nematode *G. rostochiensis* [1]. ESTs were obtained by randomly sequencing library inserts from the 5' end as described by Popeijus *et al* [8].

Sequence analysis
DNA and protein sequences were analyzed using the Laser gene software package (DNASTAR 4.00, Madison, WI). Database search was done using Blast service at the National Center for Biotechnology Information (NCBI) and WU-Blast at Human Genome Center, Baylor College of Medicine. The computer algorithm SignalP was used to predict the presence of a signal peptide for secretion and the corresponding putative cleavage site [9].

***In situ* hybridization**
The nematode fixation, hybridization and detection steps were essentially performed as described by Qin et al. [6]. The transcript-derived fragment (TDF) A29 cloned in pCR2.1-TOPO vector (Invitrogen, the Leek, the Netherlands) was amplified with primers E+TG (GAC TGC GTA CCA ATT CTG) and T+CG (GAT GAG TCC TGA CCG ACG). The amplified product was checked on a 1% agarose gel. The antisense probes were made by asymmetric PCR in the presence of DIG-dUTP (Roche Diagnostics, Mannheim, Germany) using T+CG as the primer and 20 ng of the amplified product as template in a 20-µl reaction volume. The DIG-labeled probes were purified through a G50 Mini Quick Spin DNA column (Roche Diagnostics) and 10-µl TE buffer was added. In each hybridization reaction, 30 µl of labeled probe was used. Alkaline phosphatase activity was detected by the addition of X-phosphate and NBT (Roche Diagnostics). The labeled infective juveniles were observed under a Leica inverted microscope (Leica, Deerfield, IL).

Southern blot
Primers E+TG (GAC TGC GTA CCA ATT CTG) and T+TG (GAT GAG TCC TGA CCG ATG) annealing to the 5' and 3' ends of the TDF A41 were used to amplify a 225-bp cDNA fragment. The template for this PCR was 2-ng pCR2.1-TOPO vector (Invitrogen) which contained TDF A41 sequence. The amplified PCR fragment was then used as a template for DIG labeling (Roche Diagnostics) to generate the probe. The genomic DNA was isolated as described [10]. About 2.5 µg of genomic DNA was digested by *Eco*RI and separated on a 0.8% agarose gel. The Southern blot was performed as described [11]. Hybridization was performed at 40 °C overnight in DIG Easy Hyb solution (Roche Diagnostics). The filter was washed two times in 0.1xSSC/0.1% SDS solution before detection. Alkaline phosphatase activity was detected using DIG Luminescent Detection Kit (Roche Diagnostics) and the chemiluscent signal was exposed on X-ray films.

Cloning of the full-length cDNAs
Gene specific primers were used in combination with vector primers in pcDNAII to amplify the 5'- and 3'-end region of each cDNA-AFLP fragment (A18, A41, E19) from a cDNA library constructed in pcDNAII as described [1].

Results

Identification of a family of infective stage-specific TDFs

Detailed comparison of mRNA pools isolated from five different developmental stages of the potato cyst nematode *G. rostochiensis* by cDNA-AFLP revealed a few dozens of transcript-derived fragments (TDFs) that were detected only in the infective stage of this plant parasite, the root diffusate-exposed, pre-parasitic second stage juvenile (stage H). Upon sequencing, a subgroup of the stage H-specific mRNAs appeared to be related. The restriction enzyme combination *Eco*RI and *Taq*I revealed five related, non-identical stage H-specific TDFs, A18, A29, A30, A41 and E19. A second restriction enzyme combination, *Nco*I and *Taq*I, resulted in one more related TDF, namely NA16. Once the TDFs of interest were sequenced, the expression patterns among the different developmental stages were checked by adding one or two additional selective nucleotides complementary to the sequences at the 5' and 3' ends. The extended primers resulted in a substantial reduction of the number of TDFs per gel, and the five related TDFs obtained from *Eco*RI and *Taq*I combination again showed stage H-specific expression patterns (Fig. 1).

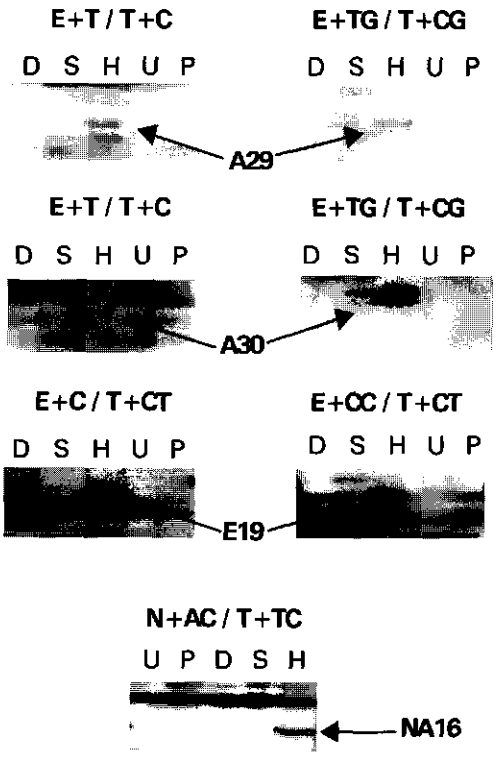


Fig. 1. The expression patterns of four TDFs, A29, A30, E19 and NA16 from the potato cyst nematode *G. rostochiensis* as revealed by cDNA AFLP. For TDFs A29, A30 and E19, the arrow-pointed bands from the left panel were cut out, cloned and sequenced. On the basis of the sequence results, the primers E and T were extended with one or two additional selective nucleotides. The resulting cDNA-AFLP expression patterns are shown (right panel). The capitals D, S, H, U and P refer to the different developmental stages. D: unhatched J2 in diapause, S: unhatched J2 after diapause, rehydrated for 2 days in water, H: freshly hatched J2 in potato root diffusate, U: developing nematodes (J1) in gravid females two months post-inoculation, P: developing nematodes (J2) in gravid females three months post-inoculation.

Using A41 as a query to search a *G. rostochiensis* EST database [12], four ESTs, GE1156, 1519, 1855, 2075, were found to share significant homology with A41. GE1855 and A30 showed a partial perfect overlap indicating both fragments were derived from the same cDNA, and the same holds true for GE2075 and E19. The

remaining ESTs, GE1156 and GE1519, are representing two additional members of the stage H-specific gene family.

A cDNA library [1] was used to identify the full size cDNAs corresponding to three TDFs, A41, A18 and E19. For A41, a full-length cDNA of 871 bp, indicated as GR-*dgl-2A*, was cloned, encoding a protein of 261 aa. The full-length cDNA corresponding to A18, GR-*dgl-2B*, included 845 bp and encoded a protein of 216 aa. The third member, E19, was 885 bp long and was indicated as GR-*dgl*-2C. GR-DGL-2C is a protein of 232 aa. GR-DGL-2B and –2C are basic, whereas the pI of GR-DGL-2A is 5.97. The presence or absence of a localization signal was tested by using SignalP [9] V2.0.b2 (including both SignalP-NN and SignalP-HMM). Each of the putative proteins was predicted to have a signal peptide for secretion.

Hence, in total eight related cDNAs were identified (Table. 1). The full size cDNAs will be indicated as GR-*dgl-2A* (corresponding to TDF A41), *-2B* (*idem* A18), *-2C* (*idem* E19), *-2D* (*idem* A29), *-2E* (*idem* A30), *-2F* (*idem* NA16), *-2G* (*idem* GE1156) and *-2H* (*idem* GE1519). The alignment is shown in Figure 2. The shared identity at the protein level among these eight proteins ranged from 28% to 57%. To confirm the existence of such a family on genome level, a Southern blot was performed with a 225-bp cDNA probe derived from TDF A41. Multiple bands were observed on the blot, showing that there are indeed multiple RanBPM-like genes present in the genome of *G. rostochiensis* (Fig. 3).

Gene name	TDF No.	EST No.	Full length cDNA (bp)	Length partial cDNA (bp)	In situ hybridization
GR-*dgl-2A*	A41	-	871	-	Dorsal gland
GR-*dgl-2B*	A18	-	845	-	Dorsal gland
GR-*dgl-2C*	E19	GE2075	885	-	n.d
GR-*dgl-2D*	A29	-	-	246	Dorsal gland
GR-*dgl-2E*	A30	GE1855	-	193	n.d
GR-*dgl-2F*	NA16	-	-	192	Dorsal gland
GR-*dgl-2G*	-	GE1156	-	470	n.d
GR-*dgl-2H*	-	GE1519	-	244	n.d

Table 1. Overview of eight cDNAs of a RanBPM-like gene family predominantly expressed in pre-parasitic second stage juveniles of the potato cyst nematode *G. rostochiensis*. n.d.: not determined.

Specific expression of stage H-specific TDFs in the dorsal gland in infective juveniles

Recently, Qin *et al.* [6] proposed a novel strategy for the identification of putative pathogenicity factors. In case of cyst nematodes, it was hypothesized that pathogenicity factors are synthesized in the dorsal or in the subventral esophageal glands. Sense and anti-sense DNA probes of the TDFs A29 and NA16 were made by asymmetric PCR using the E+0, the T+0 or the N+0 primer. Anti-sense DNA probe of the TDF A29 specifically labeled the dorsal gland of infective J2. Labeling was observed in the gland cell around the nucleus, and no signal was detectable in the gland cell extensions or in the ampullae. No labeling was observed for the sense

probes. A second *in situ* hybridization experiment with an anti-NA16 probe essentially showed a similar pattern (Fig. 4). Two other members of the gene family, A18 and A41, were recently shown to be specifically expressed in the dorsal glands of activated infective juveniles as well [6]. It remains to be determined whether the other members of the family, GR-*dgl-2C, -2E, -2G, -2H,* are specifically expressed in the dorsal esophageal gland as well. Alternatively, these RanBPM-like proteins could be functional in the nematode itself.

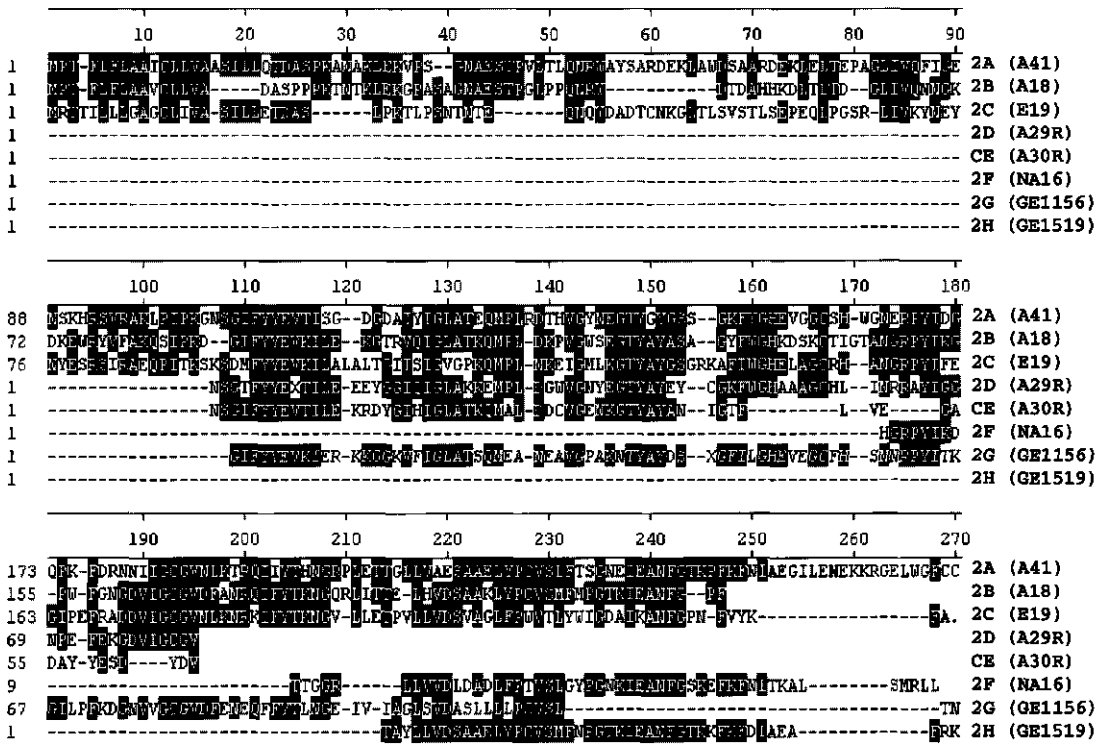
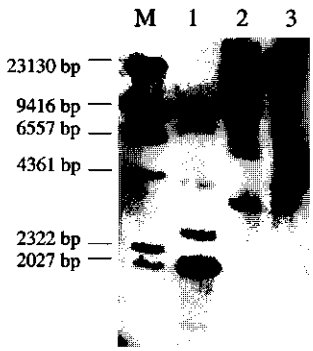


Fig. 2. The comparison of eight different RanBPM-like proteins from GR-*dgl*-2 family from *G. rostochiensis* (indicated as 2A to 2H; the original TDF and EST numbers are included between brackets). Conserved residues are shaded. The alignment was made by using Clustal-method with PAM250 residue weight table from DNASTAR 4.00.

Homologies of members of the stage H-specific gene family with known genes. Using the homology search programs BLASTP, GR-DGL-2B was shown to be homologous to the N terminal parts of a number of RanBPMs (Ran-Binding Protein in microtubule organizing center) from *Xenopus laevis* (AB018696: 38% identity, 50% similarity, and an E value of 3e-15), from fission yeast (*Schizosaccharomyces pombe*) (T40901: 31% identity, 45% similarity, and an E value of 7e-14), from *Caenorhabditis elegans* (T27155 (= hypothetical protein Y54E5A.7): 31% identity, 45% similarity, E value of 2e-12) and from *Homo sapiens* (AB008515: 40% identity; 50% similarity, E value of 5e-12). Only the function of the *Homo sapiens* RanBPM protein has been studied in some more detail. GR-DGL-2B also shared significant homology with the full-length putative RanBPM from a cryptomonad (*Guillardia theta*) (AJ010592, 280 aa: 29% identity; 43% similarity, E=3e-9) and most part of a putative *Arabidopsis thaliana* protein (AL117386, 349 aa): 25% identity; 43% similarity, E=2e-6). Other proteins from this nematode gene family also displayed similar homology with the RanBPM proteins. An alignment is shown in Fig. 5.



Comparison with the Pfam protein domain collection revealed a common SPRY motive [13] both in the full length dorsal gland proteins, GR-DGL-2A, -2B and -2C (E values between 2.9 e-18 and 1.6 e-10) and in the N-terminal parts of the RanBPMs from *X. laevis, S. pombe* and *C. elegans* (E values around e-35). The SPRY domain is thought to serve as a protein-protein interaction module [14] and/or as a domain involved in the interaction with small RNA molecules [13].

Fig. 3. Genomic Southern blot hybridized with a 225-bp cDNA probe derived from TDF A41. Lane M is molecular weight marker; Lane 1 to 3 are 2.5 µg nematode genomic DNA digested with *Eco*R I, *Bam*H I and *Kpn* I respectively.

Discussion

Comparative mRNA fingerprinting of five developmental stages of the potato cyst nematode *Globodera rostochiensis* by the cDNA-AFLP technique resulted in the identification of a RanBPM-like gene family with at least 8 members, all predominantly expressed in the infective stage of this nematode. Four members were randomly chosen for *in situ* hybridisation and all of them were shown to be expressed specifically in the dorsal esophageal gland of potato root diffusate-exposed second stage juveniles. Full-length cDNAs were identified for three family members and all of them were predicted to be preceded by a signal peptide for secretion. Probably, these RanBPM-like proteins contribute to the pathogenicity of this potato cyst nematode. Their localization in the dorsal glands suggests that these proteins are involved in feeding sites induction and not in penetration or migration through the root.

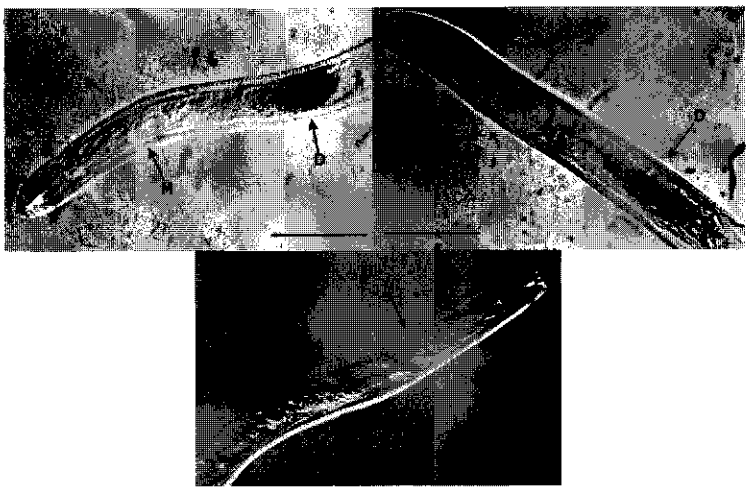


Fig. 4. In situ hybridization using TDF-derived probes of A29 and NA16 from G. rostochiensis. Labeling was done in preparasitic J2. A: labeling pattern using an A29 antisense probe. B: labeling pattern with an NA16 antisense probe. C: A typical image of a sense probe labeling. Arrows point at the dorsal gland (D) and the metacorpus (M) respectively. Scale bar = 20 μm

Ran and one of its effectors, RanBPM

Ran, a small GTPase (≈ 25 kDa) of the Ras superfamily, is very abundant and highly conserved among eukaryotes. Ran is a complex and multifunctional protein. It is involved in cell cycle regulation, microtubule spindle assembly, nuclear transport and postmitotic nuclear assembly. Ran's intrinsic rates of nucleotide exchange and hydrolysis are low, and *in vivo* activating proteins are required for these reactions. During interphase, Ran-GTP is mainly localized in the nucleus whereas Ran-GDP is largely cytosolic. Depending on the nucleotide to which it is bound, Ran associates with different cellular proteins.

RanBPM, Ran-binding protein in microtubule organizing center (MTOC), is a relatively novel Ran-binding protein which is localized in the centrosome of either interphase or mitotic cells. Human RanBPM was shown to interact specifically with Ran-GTP and not with Ran-GDP or nucleotide free Ran [15]. Moreover, RanBPM was co-fractionated with γ-tubulin ring complexes (γ-TuRCs). γ-TuRCs are cytosolic and consist of 10-14 γ-tubulin molecules and at least six additional proteins resulting in a complex of about 2 Mda. It was concluded that RanBPM can associate physically with both Ran-GTP and γ-TuRCs. Whether it directly interacts with γ-tubulin as well remains to be demonstrated [15]. Ran-binding domains have been

identified for RanBP1 [16] and importin β [17], and no similar domain was found in human RanBPM [15]. As compared to other Ran-binding proteins, the physical interaction between RanBPM and Ran-GTP is relatively weak. Overexpression of human *RanBPM* cDNA in COS cells caused ectopic microtubule nucleation, resulting in a reorganization of the microtubule network. High levels of RanBPM were suggested either to activate γ-TuRCs or to recruit γ-tubulin [15].

Could nematode RanBPM-like proteins be functional in a plant cell?
G. rostochiensis-produced RanBPM-like proteins can only be functional in plants if the proteins they normally associate with are available and recognized as such. The availability will be guaranteed if the interacting proteins are constitutively expressed and these proteins may interact with foreign RanBPM only if they are highly conserved. Ran(-GTP) is an ubiquitous, extremely conserved protein. Comparison of a small set of plant and nematode Ran species revealed that they are approximately 95% identical at amino acid level. γ-TuRC, a large protein complex, is known to physically associate with RanBPM as well. A major constituent of γ-TuRC, γ-tubulin, is thought to interact with RanBPM. Both are essentially present in every cell. Comparison of a number of nematode and plant γ-tubulins showed 40-50 % identity and 60-70% similarity at amino acid level. Though not all proteins present in γ-TuRCs have been identified yet, these complexes are thought to be highly conserved [18]. Provided that the members of the RanBPM-like gene family from the potato cyst nematode *G. rostochiensis* are indeed RanBPMs, it is likely that these cyst nematode RanBPMs can function as such in a plant environment because the interacting proteins - as far as they are known - will be available and recognizable.

Possible role of nematode RanBPM in syncytium induction.
Upon the selection of an initial syncytial cell (ISC) in the inner cortex, the pre-parasitic second stage juvenile releases secretory proteins that trigger the re-differentiation of the ISC. Dorsal esophageal gland-produced proteins are presumably injected into the cytosol of the ISC. One of the results of the presence of RanBPM-like proteins in the ISC could be a change in the microtubule stability.

Microtubules are dynamic polymers and their (in)stability is determined by the polymerization and the depolymerization rates, and by the catastrophe and the rescue frequencies. Transition from interphase to mitosis coincides with an increased catastrophe frequency and a reduced rescue frequency. As a result, the microtubules become smaller and the half-life is reduced from 5-10 minutes to 0.5-1 minutes [19]. Recently, it was shown that Ran-GTP stimulates microtubule assembly by increasing the rescue frequency three- to eightfold [20]. A release of foreign RanBPMs in the cytosol of an ISC could result in an increase in the cytosolic RanBPM-bound Ran-GTP concentration. Assuming there is no major difference between the biological activities of bound and unbound Ran-GTP, this would result in a stabilization of the microtubule network. Such a stabilization could hamper the transition from interphase to mitosis.

It was shown in syncytial cells that DNA synthesis and progression through the G2 phase are essential for their establishment, but mitosis was never observed. It was suggested that feeding cells development is accompanied by cycles of DNA endoreduplication shunting the M-phase (G1-S-G2-G1 *etc.*). The nematode secretions are thought to contain inhibitors that specifically block the activity of M-phase promoting factors and/or induce unknown S-phase-related protein kinases [4]. The RanBPM-like proteins identified in this study might be involved in the shunting of the

M-phase in nematode induced syncytium. The relevance of changes in the microtubule structure during syncytium development is further illustrated by the effects of oryzalin, a plant microtubule polymerization inhibitor. In *Arabidopsis thaliana*, low concentrations of this herbicide completely blocked syncytium initiation by *Heterodera schachtii* [4].

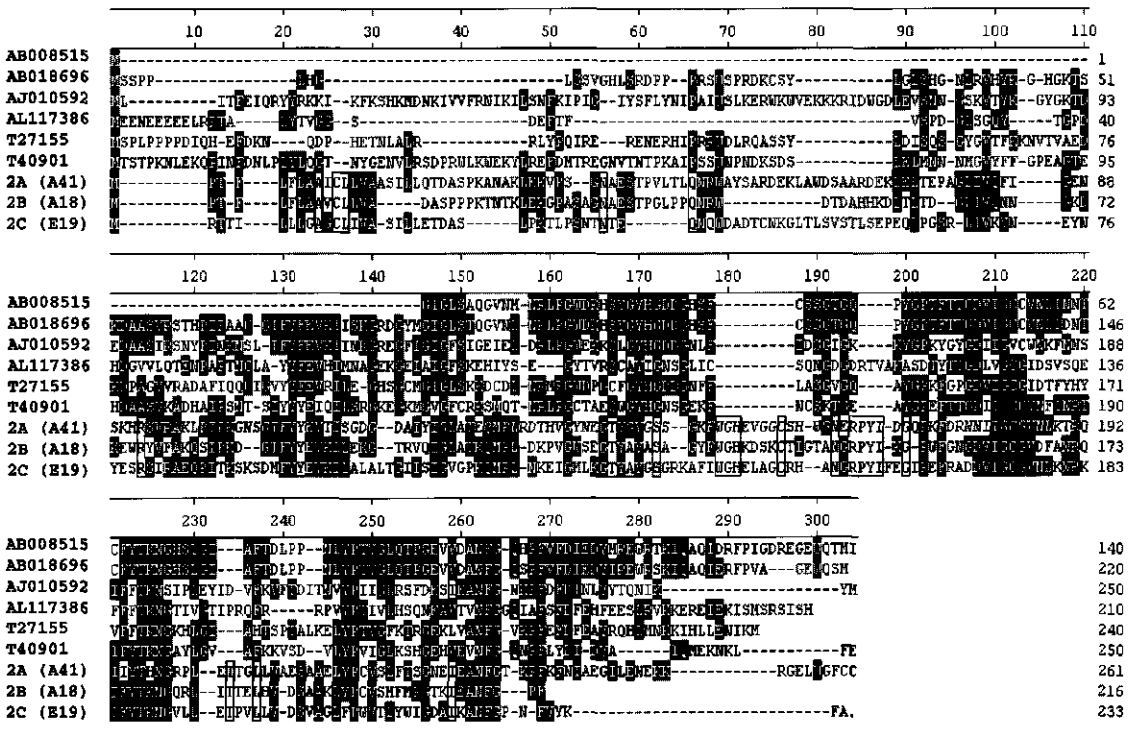


Fig. 5. The comparison of three RanBPM-like proteins (GR-dgl-2A, -2B and -2C) from *G. rostochiensis* with six other RanBPM proteins from various organisms. AB008515 from *Homo sapiens*; AB018696 from *Xenopus laevis*; AJ010592 from *Guillardia theta*; AL117386 from *Arabidopsis thaliana*; T27155 from *Caenorhabditis elegans*; T40901 from *Schizosaccharomyces pombe*. Conserved residues are shaded and the nematode specific residues are boxed. The alignment was made by using Clustal-method with PAM250 residue weight table from DNASTAR 4.00.

Microtubule manipulation by micro-organisms is not unique. This can be illustrated by the tomato mosaic virus that is hijacking host plant microtubules to target virus movement proteins to plasmodesmata [21] and by the interaction between *Rhizobium* and alfalfa. In the latter case, essential microtubule changes were shown to be induced by Nod factors [22].

In this paper, we have identified a relatively large family of RanBPM-like proteins from the potato cyst nematode *G. rostochiensis* that are probably involved in syncytium induction. This redundancy is intriguing. Whether this is a nonfunctional relic or a way to spread risks remains to be determined. As a logical next step, we will investigate the effects of over-expression of RanBPM-like proteins in plant cells both by protoplast transfection and by stable plant transformation. In this way, we will pinpoint the role(s) of this newly identified gene family in potato cyst nematode pathogenicity.

References

1. Smant G, Stokkermans JP, Yan Y, de Boer J, Baum TJ, Wang X, Hussey RS, Gommers FJ, Henrissat B, Davis EL, Helder J, Schots A, Bakker J: **Endogenous cellulases in animals: isolation of beta-1, 4-endoglucanase genes from two species of plant-parasitic cyst nematodes.** *Proc.Natl.Acad.Sci.U.S.A* 1998, **95:**4906-4911.
2. Popeijus H, Overmars H, Jones J, Blok V, Goverse A, Helder H, Schots A, Bakker J, Smant G: **Degradation of plant cell walls by a nematode.** *Nature London* 2000, **406:**36-37.
3. Niebel A, Engler Jd, Hemerly A, Ferreira P, Inze D, Montagu MV, Gheysen G: **Induction of cdc2a and cyc1At expression in Arabidopsis thaliana during early phases of nematode-induced feeding cell formation.** *Plant Journal* 1996, **10:**1037-1043.
4. Almeida-Engler Jd, Vleesschauwer Vd, Burssens S, Celenza JLJ, Inze D, Montagu Mv, Engler G, Gheysen G: **Molecular markers and cell cycle inhibitors show the importance of cell cycle progression in nematode-induced galls and syncytia.** *Plant Cell* 1999, **11:**793-807.
5. Goverse A, Rouppe van der Voort J, Rouppe van der Voort C, Kavelaars A, Smant G, Schots A, Bakker J, Helder J: **Naturally induced secretions of the potato cyst nematode co-stimulate the proliferation of both tobacco leaf protoplasts and human peripheral blood mononuclear cells.** *Mol.Plant-Microbe Interact.* 1999, **12:**872-881.
6. Qin L, Overmars H, Helder J, Popeijus H, Rouppe van der Voort, J., Groenink W, van Koert P, Schots A, Bakker J, Smant G: **An efficient cDNA-AFLP-based strategy for the identification of putative pathogenicity factors from the potato cyst nematode Globodera rostochiensis.** *Mol.Plant-Microbe Interact.* 2000, **13:**830-836.
7. De Boer JM, Overmars HA, Bakker J, Gommers FJ: **Analysis of two-dimensional protein patterns from developmental stages of the potato cyst-nematode, *Globodera rostochiensis*.** *Parasitology* 1992, **105:**461-474.
8. Popeijus H, Blok VC, Cardle L, Bakker E, Phillips MS, Helder J, Smant G, Jones JT: **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* 2000, **2:**567-574.

9. Nielsen H, Engelbrecht J, Brunak S, Von Heijne G: **Identification of prokaryotic and eukaryotic signal peptide and prediction of their cleavage sites.** *Protein Engineering* 1997, **10:**1-6.
10. Curran J, Baillie DL, Webster JM: **Use of genomic DNA restriction fragment length differences to identify nematode species.** *Parasitology* 1985, **90:** 137-144.
11. Sambrook J, Fritsch EF, Maniatis T: *Molecular cloning: A laboratory manual.* New York: Cold Spring Harbor Laboratory Press; 1989.
12. Qin L, Prins P, Jones JT, Popeijus H, Smant G, Bakker J, Helder J: **GenEST, a powerful bidirectional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP.** *Nucleic Acids Res.* 2001, **29:**1616-1622.
13. Ponting C, Schultz J, Bork P: **SPRY domains in ryanodine receptors (Ca(2+)-release channels).** *Trends Biochem.Sci.* 1997, **22:**193-194.
14. Hilton DJ, Richardson RT, Alexander WS, Viney EM, Willson TA, Sprigg NS, Starr R, Nicholson SE, Metcalf D, Nicola NA: **Twenty proteins containing a C-terminal SOCS box form five structural classes.** *Proc.Natl.Acad.Sci.U.S.A* 1998, **95:**114-119.
15. NAKAMURA M, Masuda H, Horii J, Kuma K, Yokoyama N, Ohba T, Nishitani H, Miyata T, Tanaka M, Nishimoto T: **When overexpressed, a novel centrosomal protein, RanBPM, causes ectopic microtubule nucleation similar to gamma-tubulin.** *J Cell Biol.* 1998, **143:**1041-1052.
16. Coutavas E, Ren M, Oppenheim JD, D'Eustachio P, Rush MG: **Characterization of proteins that interact with the cell-cycle regulatory protein Ran/TC4.** *Nature* 1993, **366:**585-587.
17. Gorlich D, Dabrowski M, Bischoff FR, Kutay U, Bork P, Hartmann E, Prehn S, Izaurralde E: **A novel class of RanGTP binding proteins.** *J Cell Biol.* 1997, **138:**65-80.
18. Moritz M, Agard DA: **gamma-Tubulin complexes and microtubule nucleation.** *Curr.Opin.Struct.Biol.* 2001, **11:**174-181.
19. Nishitani H, Ohtsubo M, Yamashita K, Iida H, Pines J, Yasudo H, Shibata Y, Hunter T, Nishimoto T: **Loss of RCC1, a nuclear DNA-binding protein, uncouples the completion of DNA replication from the activation of cdc2 protein kinase and mitosis.** *EMBO J* 1991, **10:**1555-1564.
20. Wilde A, Lizarraga SB, Zhang L, Wiese C, Gliksman NR, Walczak CE, Zheng Y: **Ran stimulates spindle assembly by altering microtubule dynamics and the balance of motor activities.** *Nat.Cell Biol.* 2001, **3:**221-227.
21. Heinlein M, Epel BL, Padgett HS, Beachy RN: **Interaction of tobamovirus movement proteins with the plant cytoskeleton.** *Science* 1995, **270:**1983-1985.
22. Timmers AC, Auriac MC, Truchet G: **Refined analysis of early symbiotic steps of the Rhizobium-Medicago interaction in relationship with microtubular cytoskeleton rearrangements.** *Development* 1999, **126:**3617-3628.

Chapter VI

Cloning of a *trans*-spliced glyceraldehyde-3-phosphate-dehydrogenase gene from the potato cyst nematode *Globodera rostochiensis* and expression of its putative promoter region in *Caenorhabditis elegans*

Qin,L., Smant,G., Stokkermans,J., Bakker,J., Schots,A., and Helder,J.
Cloning of a *trans*-spliced glyceraldehyde-3-phosphate-dehydrogenase gene from the potato cyst nematode Globodera rostochiensis and expression of its putative promoter region in *Caenorhabditis elegans*.
Molecular and Biochemical Parasitology 1998. 96:59-67.

Cloning of a *trans*-spliced glyceraldehyde-3-phosphate-dehydrogenase gene from the potato cyst nematode *Globodera rostochiensis* and expression of its putative promoter region in *Caenorhabditis elegans*

Abstract:

Reverse genetics to determine the relative importance of individual pathogenicity factors of the potato cyst nematode *Globodera rostochiensis* depends, apart from an efficient transformation protocol for this obligatory plant parasite, on the availability of an efficient promoter. PCR-based cloning was used to isolate a cDNA encoding glyceraldehyde-3-phosphate-dehydrogenase (GAPDH, a crucial enzyme in glycolysis and gluconeogenesis; this gene was designated *gpd*) and its 5'-flanking region. The cDNA includes 1047 nucleotides encoding an open reading frame that shows high homology with GAPDHs from *Caenorhabditis elegans* and other species. Analysis of the 745 bp 5'-flanking region of the *gpd* gene showed no homology with a similar region in *C. elegans*. In this region several eukaryotic promoter elements are present. 5' Rapid amplification of cDNA ends revealed this gene was *trans*-spliced with a SL1 spliced leader. The 5'-flanking region of the *gpd* gene was fused to green fluorescent protein reporter gene and microinjected into the gonads of C. *elegans*. Green fluorescent protein expression, under the transcriptional control of the 5'-flanking region of *gpd*, was mainly observed in body wall muscles of transgenic animals. This putative promoter region of GAPDH could be a valuable tool to drive gene expression in transgenic *G. rostochiensis* and other related plant-parasitic nematode species.

Keywords: *Globodera rostochiensis*, glyceraldehyde-3-phosphate dehydrogenase, putative promoter region, spliced leader, gene cloning

List of abbreviations: PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends; UTR, untranslated region; GFP, green fluorescent protein; ORF, open reading frame

Note: Nucleotide sequence data reported in this paper is available in the EMBL, GenBank™ and DDJB data bases under the accession number AF004522.

1. Introduction

Globodera rostochiensis, the potato cyst nematode, is a highly specialized plant-parasite. The main host of this sedentary endoparasite is potato and, as such, it may cause substantial losses in potato crops. Upon arrival at the root surface, preparasitic second stage juveniles (J2) secrete a complex mixture of compounds [1]. This complex mixture includes cell wall-degrading enzymes [2]. The secretions facilitate the penetration of the rhizodermis and subsequent cell layers. After intracellular migration through the cortical tissue, a feeding structure - a so-called syncytium - is induced within the central cylinder.

If we want to know more about the relative importance of compounds secreted during the infection process, mutant nematodes that are unable to produce and/or secrete this compound would be useful. Because no such mutants are available from the classical genetic approach, a protocol for transforming obligatory plant parasitic nematodes is urgently needed. One of the prerequisites for this transgenic approach being successful would be the availability of (constitutive) promoters that are properly recognized by the host organism. At present several 5'-upstream regions of locally expressed plant-parasitic nematode genes have been identified [3, 4, 5]. However, none of them were shown to be functional promoter regions.

Both in pro- and eukaryotes, glyceraldehyde-3-phosphate dehydrogenase (GAPDH, EC 1.2.1.12) is an abundant and crucial enzyme in glycolysis and gluconeogenesis. It catalyzes reversibly the oxidation and phosphorylation of D-glyceraldehyde-3-phosphate to 1,3-diphosphoglycerate. Apart from being a glycolytic enzyme, GAPDH is involved in the translational control of gene expression; in the nucleus, it functions in nuclear tRNA export, in DNA replication, and in DNA repair [6]. Genes encoding GAPDH have been isolated from many organisms revealing strong homologies in the NAD-binding domain, and, even to a higher extent, in the catalytic domain [7, 8]. For constitutive expression in cyst nematodes, the promoter of GAPDH could be a good candidate.

The members of the phylum Nematoda can, on the basis of their feeding behavior, be divided into free-living species, animal- and plant-parasites. *gpd-1*, *gpd-4* and their putative promoter regions have been isolated - not characterized - from the free-living species *Caenorhabditis elegans*, whereas *gpd-2* and *gpd-3* are located downstream of another gene in a single operon [9]. GAPDH-encoding cDNA sequences are available from the animal parasitic nematodes *Onchocerca volvulus* (U96177) and *Brugia malayi* (U18137). Nothing is known about their promoters.

If heterologous promoters (*e.g.* from *C. elegans*) are used for the transformation of plant parasitic nematodes, a negative result could either be due to an inappropriate transformation procedure or the unsuitability of the heterologous promoter. Here we reported the molecular cloning of GAPDH-encoding cDNA from *G. rostochiensis* and its putative promoter sequence, which is the first reported from a plant parasitic nematode species. By using the green fluorescent protein (GFP) marker gene from the jellyfish *Aequorea victoria*, we showed that this putative promoter region is functional in *C. elegans*. This endogenous promoter of an abundantly expressed GAPDH-encoding gene could be very useful in developing a transformation method for *G. rostochiensis* and other plant parasitic species.

2. Material and methods

2.1. Nematode culture and isolation

Preparasitic second stage juveniles (J2) of *G. rostochiensis* Ro_1 were cultured on plant roots and extracted as described previously [10]. Wild-type *C. elegans* strain N2 was grown as described by Brenner [11].

2.2. cDNA cloning of *gpd*

Approximately 1.0 ml of *G. rostochiensis* preparasitic second-stage juveniles [J2, $\approx 2x10^6$ J2 ml^{-1}] were homogenized in 10 mM Tris-Cl pH 7.4, including 3 mM Pefabloc (Boehringer Mannheim Gmbh, Mannheim, Germany) and subjected to preparative SDS-PAGE (BioRad, Hercules, USA) on 10% denaturing polyacrylamide gel [10]. In our laboratory, a monoclonal antibody MGR48 was produced which could recognize a 39 kD protein from the subventral esophageal gland of *G. rostochiensis* [10]. In order to identify this protein, those fractions with an apparent molecular mass of approximately 39 kD were pooled and subsequently separated on an MGR48 immunoaffinity chromatography column [12,2]. For N-terminal amino acid sequencing a 39 kD protein was subjected to western-blotting using a Tris-tricine buffer system [13]. The single protein band on the blot was stained with Coomassie Brilliant Blue R-250 and cut out for N-terminal protein sequencing (ARIAD Pharmaceuticals, Cambridge, USA). Two N-terminal amino acid sequences were obtained from this band. One was identified as endoglucanase [2], the other showed high degree of homology with GAPDHs from different species in a database search. A degenerate gene-specific forward primer gpd-f (AARGTIGGIATHAAYGGNTTYGG) was designed from the N-terminal amino acid sequence. A second reverse primer gpd-r (TAICCRAAYTCYTACCA) was designed from a conserved amino acid region at the C-terminus in *gpd-1* through *gpd-4* of *C. elegans*. The primers gpd-f and -r were used to amplify a 1 kb fragment from a *G. rostochiensis* cDNA library [2]. Polymerase chain reaction (PCR) conditions: 30 cycles of 1 min at 93°C, 1 min at 60°C, and 3 min at 68°C. A 1 kb PCR fragment was ligated into pCR2.1 and used for transformation of *E. coli* INVαF' (Invitrogen, Leek, The Netherlands). Inserts were sequenced (Amersham Life Science, Buckinghamshire, UK) on an Automated Laser Fluorescent DNA sequencer (Pharmacia, Uppsala, Sweden).

2.3. 5' rapid amplification of cDNA ends (5' RACE)

The 5' end of cDNA was captured using GIBCOBRL 5' RACE kit (Life technologies) according to manufacturer's manual. Total RNA was isolated from a frozen pellet of about 100,000 J2 using TRIzol Reagent (Life technologies, Breda, The Netherlands). A *gpd*-specific primer gpd-up3 (CGGCCATGGGTCGAGTCGTAGTTG) was used in first strand cDNA synthesis from about 5 µg of total RNA. Two other *gpd*-specific primers, which anneal closer to the 5' end of the structural gene,
gpd-up2 (CCACCACCTGAACGGTGTCCTTC) and
gpd-up1 (TAATGCCGACCTTCGGTTTCACCA) were used in subsequent nested PCR steps performed in a MJ thermal cycler. The amplification conditions were as follows: 90s at 94°C, followed by 35 cycles of 60 s at 94°C, 30 s at 63°C, 90 s at 72°C, last step 5 min at 72°C. Amplification products were cloned into pCR2.1. Five clones were sequenced from both directions. Resulting sequences were analyzed using the Laser gene

(DNASTAR Inc, Madison, USA) software package.

2.4. Cloning of 5'- flanking region

In order to clone the 5'-flanking region of *gpd*, a genomic library of *G. rostochiensis* was constructed in the plasmid vector pZErO-2 (Invitrogen). DNA was extracted from preparasitic J2 as described by Curran [14]. Genomic DNA was partially digested by *Sau*3A. Fragments ranging from 500 bp to 4 kb were gel purified and ligated to *Bam*HI-digested pZErO-2 at 16°C overnight. The ligation product was transformed into TOP10 *E.coli* electrocompetent cells (Invitrogen) by electroporation using a Bio-Rad Gene Pulser (Bio-Rad). This resulted in 10^6 primary recombinants containing an average insert of about 1 kb. Library was plated out and collected in LB medium. Plasmid preparation was done using Wizard Plus Midiprep (Promega, Madison, USA). In PCR reactions, nested primers gpd-up1, -up2 were used in combination with either of the two pZErO-2 vector-specific primers:

PCDNA-f: (GACGGCCAGTGAATTGTAATACGACTCACT)

PCDNA-r: (GGTGACACTATAGAATACTCAAGCTATGCA)

using plasmid preparation of the genomic library as template. Expand Long Template PCR system (Boehringer Mannheim Gmbh) was used to minimize possible errors. The amplification conditions were as follows: 2 min at 94°C, followed by 35 cycles of 60 s at 94°C, 60 s at 65°C, 2 min at 68°C, last step 5 min at 68°C. Amplified products were cloned into pCR2.1. Two clones amplified from two different PCR reactions were sequenced from both directions.

2.5. GFP fusion construct

The GFP reporter construct gpd95.75 was generated by joining the 5'- flanking region of *gpd* to the *gfp* gene in the C. *elegans* promoter-less GFP vector pPD95.75 (kindly provided by Dr. Andy Fire, Carnegie Institution of Washington, USA) using PCR. To this end, the 745 bp 5'-flanking region of *gpd* together with a DNA sequence encoding the first 8 predicted amino acids of GAPDH was amplified using the primers up1- *Bam*HI (CGCGGATCCTTAATGCCGACCTTCGGTTTCACCA, underlined nucleotides indicating the introduced *Bam*HI site) and PCDNA-r. The resulting fragment were double digested with *Bam*HI and *Hind*III, and inserted into the *Bam*HI and *Hind*III sites of pPD95.75. To check whether the 5' coding region of *gpd* and the *gfp* gene were in the same reading frame, the junction was sequenced.

2.6. Microinjection and analysis of transgenic *C. elegans*

The GFP expression construct gpd95.75 (53µg ml^{-1}) and *rol-6* containing plasmid pRF4 (100µg ml^{-1}) were injected into the distal parts of the gonads of C. *elegans* N2 adult hermaphrodites [15]. Transgenic F1 animals were recognized by their rolling behavior. These animals were used to establish transmitting lines of transgenic animals expressing both plasmids in an extrachromosomal array. About 40 transgenic animals from four different lines were checked using epifluorescence microscopy (Leica. filter set: BP450-490nm excitation filter, 510nm dichoric and LP520 emission filter) for expression of GFP.

A.

```
          1              15 16             30 31             45 46             60 61             75 76             90
1 CELEG2 MPKPNVGINGFGRIG RLVLRAAVE------ ---------------- ---------KDSVNV DAVNDPFISIDYMVY LFQYDSTHGRFEGTV      60
2 CBRIG2 MSKPTVGINGFGRIG RLVLRAAVE------ ---------------- ---------KDSVNV VAVNDPFISIDYMVY LFQYDSTHGRFEGTV      60
3 BMALA  MSKPKVGINGFGRIG RLVLRAAVE------ ---------------- ---------KDTVDV VAVNDPFINIDYMVY MFKYDSTHGRFKGSV      60
4 OVOLVU MSKPKIGINGFGRIG RLVLRAAVE------ ---------------- ---------KDTVEV VAVNDPFINIDYMVY MFKYDSTHGRFKGHV      60
5 GROST  MVKPKVGINGFGRIG RLALRAAVE------ ---------------- ---------KDTVQV VAINDPFIELDYMVY MFNYDSTHGRFNGKI      60
6 JORIEN --KVKVGVNGFGRIG RLVTRAAFNSGKVDI VAINDPFIDLNYMVY MFKYDSTHGKFKVDI VAINDPFIDLNYMVY MFKYDSTHGKFKGTV      88

          91             105 106            120 121            135 136            150 151            165 166            180
1 CELEG2 AHEGDYLLVAKEGKS QHKIKVYNSRDPAEI QWGASGADYVVESTG VFTTIEKANAHLKGG AKKVIISAPSADAPM FVVGVNHEKYDHAND     150
2 CBRIG2 KHEGDYLIVANEGKS QHKIKVYNSKDPAEI QWGAAGADYVVESTG VFTTIEKANAHLKGG AKKVIISAPSADAPM FVVGVNHEKYDHAND     150
3 BMALA  SAEGGKLIVTN-GKT THEISVHNSKDPAEI PWGVDGAEYVVESTG VFTTTDKASAHLKGG AKKVIISAPSADAPM FVMGVNNDMYDKANN     149
4 OVOLVU SAEGGKLIVTN-GKT THQIAVHNSKDPAEI PWGVEGAEYVVESTG VFTHTEKASAHLKGG AKKVIISAPSADAPM FVMGVNNDKYDKANN     149
5 GROST  STSAGNLVVEKEGKA THTIKVFNLKDPAEI KWAEVGAEYVIESTG VFTTIEKASAHLKGG AKKVVISAPSADAPM YVMGVNEDKYDPAKD     150
6 JORIEN KAENGKLVING---- -HAITIPQERDPSKI KWGDAGAEYVVESTG VFTTMEKAGAHLKGG AKRVIISAPSRDAPM FVMGVNHEKYDKS-L     172

          181            195 196            210 211            225 226            240 241            255 256            270
1 CELEG2 HIISNASCTTNCLAP LAKVINDNFGIIEGL MTTVHAVTATQKTVD GPSGKLWRDGRGAGQ NIIPASTGAAKAVGK VIPELNGKLTGMAFR     240
2 CBRIG2 HIISNASCTTNCLAP LAKVINDNFGIIEGL MTTVHAVTATQKTVD GPSGKLWRDGRGAGQ NIIPASTGAAKAVGK VIPELNGKLTGMAFR     240
3 BMALA  HIISNASCTTNCLAP LAKVIHDKFGIIEGL MTTVHATTATQKTVD GPSGKLWRDGRGAGQ NIIPASTGAAKAVGK VIPGLNGKLTGMANR     239
4 OVOLVU HIISNASCTTNCLAP LAKVIHDKFGIIEGL MTTVHATTATQKTVD GPSGKLWRDGRGAGQ NIIPASTGAAKAVGK VIPDLNGKLTGMASR     239
5 GROST  NVISNASCTTNCLAP LAKVINDEFGIIEGL MTTVHAVTATQKTVD GPNGKQWRDGRGAAQ NIIPASTGAAKAVGK VIPELNGKLTGMAFR     240
6 JORIEN KIVSNASCTTNCLAP LAKVIHDNFGIVEGL MTTVHAITATQRTVD GPSAKLWRDGAGAAQ NIIPASTGAAKAVGK VIPELNGKLTGMAFR     262

          271            285 286            300 301            315 316            330 331            345 346            360
1 CELEG2 VPTPDVSVVDLTARL EKPASLDDIKKVIKA AADGPMKGILAYTED QVVSTDFVSDTNSSI FDAGASISLNPHFVK LVSWYDNEFGYSNRV     330
2 CBRIG2 VPTPDVSVVDLTARL EKPASLDDIKRVIKA AAEGPLKGVLAYTED QVVSTDFVSDTHSSI FDAGASIILNPNFVR LISWYDNEFGYSNRV     330
3 BMALA  VPTPDVSVVDLTCRL QKGATMDEIKAAVKE AAMGPMKGILEYTED QVVSTDFTGDTHSSI FDSLACISLNPNFVK LIAWYDNEYGYSNRV     329
4 OVOLVU VPTPDVSVVDLTCRL QKGASMDEIKAAVKE AAAGPMKGILEYTED QVVSSDFVGDPHSSI FDALACISLNPNFVK LIAWYDNEYGYSNRV     329
5 GROST  VPTPNVSVVDLTARL EKPASLDAIKAAVKK AAEGNLKGILGYTED QVVSTDFLGDSRSSI FDAGACISLNPHFVK LVSWYDNEF------     324
6 JORIEN VPTANVSVVDL/TCRL EKPAKYDDIKRVVKQ ACDGPLKGMLGYTEH QVVSSDFNGDSHSST FDAGAGIALNDHFVK LVSWYDNEFGYSNRV     352

          361        375
1 CELEG2 VDLISYIATKA    341
2 CBRIG2 VDLISYIATKA    341
3 BMALA  VDLISYIASR-    339
4 OVOLVU VDLISYNASK-    339
5 GROST  -----------    324
6 JORIEN VDLMVHMASKE    363
```

B.

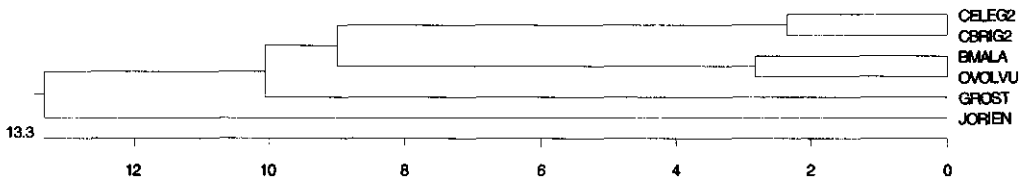


Fig. 1. (A) Comparison of the predicted amino acid sequence of *gpd* gene from *G. rostochiensis* (GROST, accession no. AF004522) with other predicted amino acid sequences from *C. elegans* (CELEG2, accession no. X15254), *Caenorhabditis briggsae* (CBRIG2, accession no. R04501), *Onchocerca volvulus* (OVOLVU, accession no. U96177), *Brugia malayi* (BMALA, accession no. U18137), *Jaculus orientalis* (JORIEN, accession no. X87226). (B) Phylogenetic tree of GAPDH amino acid sequences as caculated by Clustal-method with PAM250 residue weight table.

3. Results

3.1. Isolation of the cDNA of the *gpd* gene from the plant-parasitic nematode *G. rostochiensis*

N-terminal amino acid sequencing of a 39 kD protein from the potato cyst nematode *G. rostochiensis* resulted in a sequence of 46 residues:
KVGINGF(G)RIGRLA(L)RAAVEK(D)T(?)DV(V)AINDP(F)I(N)(L)DYMV(Y)MF
(residues or ? between brackets, not identifiable with high confidence). This sequence showed high degree of homology with GAPDH proteins from different species. On the basis of this sequence, a degenerate gene-specific forward primer gpd-f was designed. A second reverse primer gpd-r was designed from a conserved amino acid region at the C-terminus in *gpd-1* through *gpd-4* of *C. elegans*. Using gpd-f and -r, a 1046bp (AF004522) PCR fragment was amplified from a cDNA library. This fragment encodes a protein of 324 amino acids, which showed high homology with GAPDH proteins from different species (Fig. 1A).

At amino acid level, GAPDH from *G. rostochiensis* shares 80.6% similarity with predicted amino acid sequence of *gpd-2* from *C. elegans* (CELEG2, accession no. X15254), 80.2 % similarity with GAPDH from *Caenorhabditis briggsae* (CBRIG2, accession no. R04501), 79.6% with *O. volvulus* (OVOLVU, accession no. U96177), 79.3% with *B. malayi* (BMALA, accession no. U18137), 73.5% with *Jaculus orientalis* (JORIEN, accession no. X87226). The GAPDH from *J. orientalis*, jerboa, is the most closely related vertebrate GAPDH. In the highly conserved catalytic domain, a stretch of 99 aa was, except for 5 aa residues, identical to *C. elegans* GAPDHs (Fig. 1A. Amino acid residues 184-282). To visualize the degree of relatedness between different GAPDH amino acid sequences a phylogenetic tree was made using the Clustal-method with PAM250 residue weight table (Fig. 1B).

3.2. Cloning of the 5'- flanking region of *gpd*

Nested primers gpd-up1, -up2 were used in combination with either PCDNA-f or PCDNA-r to amplify fragments from a plasmid preparation of a *G. rostochiensis* genomic library. Both combinations resulted in specific PCR products of about 1 kb. The two clones were sequenced from both directions. In spite of being cloned in opposite orientation in the pZErO vector, these fragments appeared to be identical. The DNA sequences overlapped with the 91 bp 5' end of the cDNA sequences. A 745 bp 5'-flanking region of *gpd* was obtained (Fig.2) (this fragment was amplified also from genomic DNA, data not shown).

3.3. 5' rapid amplification of cDNA ends

5' RACE was used to determine whether *trans*-splicing of mRNA occurs in *G. rostochiensis*. This resulted in the identification of a 22-nucleotide sequence present at the 5' end of the *gpd* cDNA. Five clones were sequenced, and the 22 bp stretch was shown to be identical to the SL1 sequence of *C. elegans*. This sequence was not present in the 5'- flanking genomic region of the *gpd* gene (Fig.2). The splicing acceptor site "TTTTTAGA" is consistent to the consensus of *C. elegans* splicing acceptor site

"TTTTCAGA/G" [16], except a T residue in the place of a C residue before the most conserved "AG". No typical 5' splicing signal AGGTRAGTTT could be found in the 5' genomic sequence of *gpd*.

```
-745  TCTACGTACATACTATAATATTTTTTCTCTCTTCTCTTCGCTTTCCCATCTCATTTGTCATTGCCGGTTCGTGTCTCGCGAGTGAGCGA
-655  CGAAAAGGCCGAATCAATTGGTTGGGGAAGGGCAGGGAGGGGGTTAAATGTGCTGTTCTTTGGAGGGAAACCCACCAACCACCGCTCAT
-565  TCGTTCTCTGAGCTCAGTCCAACACATCGCCTCATCCCTTCCTTTTCATTTCAGTTAACCATTTCTGTTTACGCTCTGGTATATGGGGA
-475  ATTATTTTTGTTGACTTTATTTCAGTTCTCTTTTCTTACACTGAGCCAGCACATTCTTATTCAAAAACAGTTTGCAGTTTGTCCTCTTC
-385  TTTAACTTTCAGTAGTAGTGGACTCTAGCTCTGACAAATTAGAACTCCAATTCGGGGTCCCCAGAACCGAACCGAAAAAAGGCGGGTGC
-295  CCGACAATCAAAATCTTTCGAGGGGTTGGAACCGACACCGACAACCCTATTTAGAGAATTTGTTGAAGGATGCTGTTATAACTTATAAG
-205  TAGTGTGCCATGCCTCCTCCGGTTTTGCGAAAATGTTTCTTAAACACGCGTCTTCGTGTCAAACCCATGTAACGAATGTTATTGATTTG
-115  CTGTTCAGATTAGTTTTTGTTTATGCTCTTATGTGCATGCATTTTTAGATTATTGACTTTGCATATTTCCAACCAAGTTCATTTGACCA
                                  GGTTTAATTACCCAAGTTTGAG↑
 -25  TATTTTCCTGCTAAACATAGCAAAAATGGTGAAACCGAAGGTCGGCATTAATGGCTTTGGACGCATTGGGCGCTTGGCGTTGCGCGCTG
 +66  CGGTTGAGAAGGACACCGTTCAGGTGGTGGCGATCAATGACCCGTTCATCGAACTCGACTACATGGTATACATGTTCAACTACGACTCG
+156  ACCCATGGCCGCTTCAATGGCAAAATTTCGACAAGCGCCGGCAATTTGGTCGTTGAGAAAGAGGGGAAGGCCACGCACACCATCAAGGT
+246  GTTCAACCTCAAGGACCCGGCCGAGATCAAATGGGCTGAGGTGGGCGCGGAATATGTGATCGAGTCCACCGGGGTGTTCACTACCATTG
+336  AGAAGGCTTCGGCACACTTGAAGGGGGGCGCCAAGAAGGTGGTCATCTCTGCTCCGTCCGCTGATGCACCGATGTACGTGATGGGCGTC
+426  AACGAGGACAAATATGACCCGGCCAAGGACAACGTGATTAGCAACGCTTCGTGCACCACCAACTGCCTTGCGCCGCTGGCCAAAGTGAT
+516  CAACGACGAGTTTGGCATCATCGAAGGGTTGATGACCACTGTTCACGCAGTAACTGCCACTCAGAAAACGGTGGACGGACCGAATGGGA
+606  AGCAGTGGCGCGACGGGCGTGGCGCCGCGCAGAACATCATCCCTGCCAGCACCGGGGCAGCCAAAGCAGTGGGCAAGGTCATTCCTGAG
+696  CTGAACGGCAAATTAACCGGCATGGCATTCCGTGTGCCGACCCCGAACGTTTCCGTCGTCGACCTGACCGCTCGTTTGGAGAAGCCGGC
+786  CTCATTGGACGCCATCAAGGCGGCGGTGAAGAAGGCTGCCGAAGGGAATTTGAAGGGCATTTTGGGTTACACAGAGGACCAGGTGGTGT
+876  CCACGGACTTTCTTGGAGACAGTCGCTCGTCGATCTTCGACGCTGGGGCGTGCATCTCGTTGAACCCGCACTTTGTCAAGTTGGTCAGC
+966  TGGTACGACAATGAATTT
```

Fig. 2. The 5'-flanking and coding nucleotide sequence of *gpd* gene from *G. ros*. Nucleotides are numbered from the ATG start codon with A being position +1. ATG start codon, TATA promoter elements, CAAT promoter elements and *trans*-splicing acceptor site are boxed. CpG doublets surround the TATA and CAAT box are in bold. The *trans*-spliced leader (SL-1) sequence is shown in italics. Nested primers used both for 5'RACE and amplification of the 5'-flanking region are underlined. SL-1 is *trans*-spliced to mRNA at nucleotide A indicated by an upward arrow directly after the boxed *trans*-splicing acceptor site.

3.4. Potential cis-regulatory elements of the 5'- flanking region of *gpd*

130 bp upstream of the splicing acceptor site, two TATA boxes - RNA polymerase II subunit TBP (TATA box binding protein) binding sites - are present. 65 bp upstream of the TATA boxes, there were two typical CAAT boxes, which are a common motif in eukaryotic promoters. Another remarkable feature is the CpG islands, which contain concentrations of CpG doublets and often surround the promoters of constitutively expressed genes; eight such motives were found upstream of one of the two TATA boxes and six immediately downstream of it. It is noted that the ATG translation start site was preceded with four adenine residues (Fig.2), a phenomenon frequently observed in *C. elegans* genes [16].

3.5. Transformation of *C. elegans* with *gpd* promoter GFP fusion construct

Construct gpd95.75 was made by fusing the 769 bp *gpd* 5' upstream region in frame with *gfp*. This construct was injected into *C. elegans* adult hermaphrodite gonads together with *rol*-6 containing construct pRF4. Transgenic lines were established. About 15-40% of the progeny (F1) inherited the plasmid array. Four such lines were kept. About 40 transgenic animals were examined. Though the intensity of GFP expression varied between different individuals, all lines showed similar expression patterns and GFP was mainly localized in body wall muscle cells. In some transgenic nematodes, GFP expression was also detected in vulva muscle cells (Fig. 3.a, V.). GFP expression was observed in embryos and different larval stages. In embryos, the GFP expression was seen as many bright spots (Fig. 3.b). In the J1 larval stage just before hatching, GFP expression was also visible in body wall cells. Upon close inspection, clear green striated muscle fibers were observed in body wall muscles (Fig. 3.c, 3.d). The mosaic expression of GFP in transgenic *C. elegans* is not unusual. It could be due to the mosaic inheritance of the extrachromosomal array, or by other mechanisms preventing the introduced genes to be expressed, as evidenced by mosaic expression of integrated genes [17].

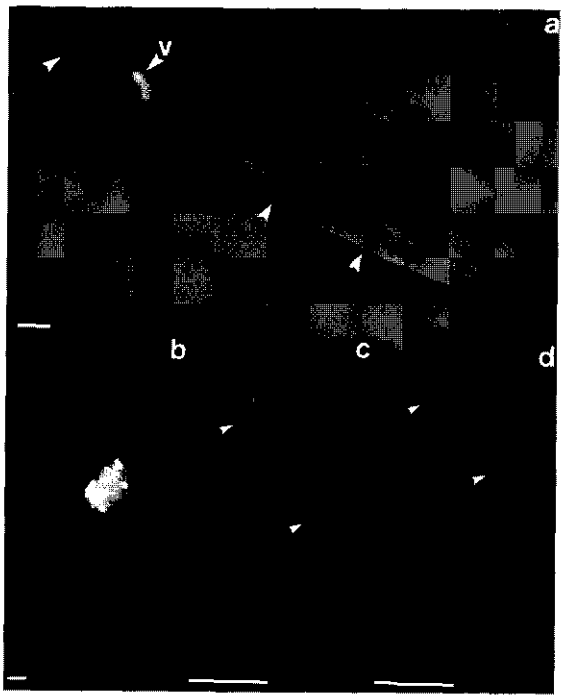


Fig. 3. GFP expression driven by the *gpd* promoter of *G. rostochiensis* in *C. elegans*. (3a) GFP expression in an adult hermaphrodite (arrows point to fluorescent fibers, V = vulva, bar = 80 µm (3b) *C. elegans* eggs expressing GFP, bar = 16 µm. (3c and 3d) GFP expression in muscle fibers in adult hermaphrodite at higher magnification. Bar = 80 µm.

4. Discussion:

Knocking out individual pathogenicity factors is the obvious way to establish the importance of these factors in the interaction plant parasitic nematode - plant. No protocol is available yet to transform plant-parasitic nematodes. By the identification of a homologous promoter that could drive gene expression in plant-parasitic nematodes, a first step is made towards transformation of these parasites. To that end , we cloned a

GAPDH-encoding gene and its 5'-flanking region from the potato cyst nematode *G. rostochiensis*, which is the first *gpd* gene isolated from plant parasitic nematode species. The deduced amino acid sequence of this *gpd* gene showed high homology with GAPDH proteins from other species. Using GFP as a reporter, it was shown that a 745 bp fragment of the *G. rostochiensis gpd* promoter is able to drive transcription in transgenic *C. elegans*. Both in larvae and adults of C. *elegans*, GFP accumulation was observed in body wall muscles mainly. The observation that a truncated promoter - even when it includes only several hundred base pairs - still results in tissue specific expression patterns in *C. elegans* is not unusual [18,19].

Previous immunological studies in preparasitic second-stage juveniles of *G. rostochiensis* revealed a similar expression pattern of *gpd* [De Boer, unpublished results]. *gpd-2* and *-3 from C. elegans* are preferentially expressed in body wall muscles during postembryonic larval development, too [20]. Unfortunately, it is hard to directly compare the promoter regions of *gpd-2* and *-3* from *C. elegans* on the one hand and *gpd* from *G. rostochiensis* on the other because *gpd-2* and *-3* are located downstream in a single operon proceeded by the *mai-1* gene [9]. It is no use comparing the transcriptional control of a polycistronic mRNA precursor with the control of a single gene. Nevertheless, our results indicate that certain regulatory elements in the *gpd* promoter region are recognized by *trans*-acting factors from *C. elegans*.

The mRNA coding for GAPDH in *G. rostochiensis* was shown to be preceded by a spliced leader. The first 22 nucleotides at the 5'-end of mature mRNA of GAPDH were identical to the *trans*-spliced leader SL1 of *C. elegans* [21]. *Trans*-splicing of mRNAs is a common phenomenon in *C. elegans*; it is estimated that at least 70% of the mRNAs have a SL1 sequence [22]. The free-living nematode *C. elegans* is only distantly related to obligatory plant-parasitic nematodes. Among the latter, spliced leaders were shown in a few occasions only. In *Meloidogyne incognita*, a putative esophageal gland protein gene was *trans*-spliced to a splicing leader sequence differing at one base pair with *C. elegans* SL1 [5]. In *G. rostochiensis*, a potato cyst nematode, a SL1 encoding gene was present in tandem repeating units together with the 5S ribosomal RNA gene [23]. In this paper we show that the presence of SL1 genes in *G. rostochiensis* resulted in mRNA containing a SL1 sequence at its 5'-end. Presumably *trans*-splicing to SL1 is a common phenomenon in *G. rostochiensis*, just as it is in *C. elegans*.

In *C. elegans*, SL1 tends to be spliced very close to the translation start site; in more than 95% percent of surveyed genes, splicing leaders are placed within 30bp of the translation start site [16]. In *B. malayi*, similar very short 5'-UTRs were also observed [24]. Thus SL1 appeared to be involved in translation initiation. However the distance from the splicing site to the translation start site of *gpd* mRNA is 66 bp, much longer than reported from the above two species. The significance of a long 5'-UTR to translation initiation is not clear yet.

Within the ancient phylum Nematoda, free-living nematodes such as *C. elegans* and plant-parasitic cyst nematodes diverged more than 100 million years ago. Nevertheless, we identified some highly conserved characteristics between these species: *G. rostochiensis gpd* is *trans*-spliced to a spliced leader identical to SL1 from *C. elegans*, the splicing acceptor site of *gpd* is virtually identical to *C. elegans*. GAPDHs from both species are about 80% identical at amino acid level. The non-conserved part of our results, the putative promoter region of *gpd*, is a good starting point for the

transformation of cyst nematodes.

References

[1] Smant G, Goverse A, Stokkermans JPWG, et al., Potato root diffusate-induced secretion of soluble, basic proteins originating from the subventral esophageal glands of potato cyst nematodes. Phytopathology 1997; 87: 839-845

[2] Smant G, Stokkermans JPWG, Yan Y, et al., Endogenous cellulase in animals: isolation of β-1,4-endoglucanase genes from two species of plant-parasitic cyst nematodes. Proc. Natl. Acad. Sci. USA 1998; 95: 4906-4911

[3] Giorgi C, Luca F, Vito M, et al., Modulation of expression at the level of splicing of cut-1 RNA in the infective second-stage juvenile of the plant parasitic nematode *Meloidogyne artiellia*. Mol. Gen. Gen. 1997;253:589-598.

[4] Novitski CE, Brown S, Chen R, Corner AS, Atkinson HJ, MePherson MJ. Major sperm protein genes from *Globodera rostochiensis*. J. Nematol. 1994;25:548-554.

[5] Ray C, Abbott AG, Hussey RS. Trans-splicing of a *Meiloidogyne incognita* mRNA encoding a putative esophageal gland protein. Mol. Biochem. Parasitol. 1994;68:93-101.

[6] Sirover MA. Role of the glycolytic protein, glyceraldehyde-3-phosphate dehydrogenase, in normal cell function and in cell pathology. J. Cell. Biol. 1997;66: 133-140.

[7] Yarbrough PO, Hayden MA, Dunn LA, Vermersch PS, Klass MR, Hecht RM. The glyceraldehyde-3-phosphate dehydrogenase gene family in the nematode, *Caenorhabditis elegans*: isolation and characterization of one of the genes. Biochim. Biophys. Acta. 1987; 908: 21-33.

[8] Tso JY, Sun XH, Kao TH, Reece KS, Wu R. Isolation and characterization of rat and human glyceraldehyde-3-phosphate dehydrogenase cDNAs: genomic complexity and molecular evolution of the gene. Nucleic Acids Res. 1985;13: 2485-2502.

[9] Spieth J, Brooke G, Kuerston S, Lea K, Blumenthal T. Operons in *C. elegans*: Polycistronic mRNA precursors are processed by *trans*-splicing of SL-2 to downstream coding regions. Cell 1993;73:521-532.

[10] Boer JM, Smant G, Goverse A, Davis EL, et al. Secretory granule proteins from the subventral oesophageal glands of the potato cyst-nematode identified by monoclonal antibodies to a protein fraction from second-stage juveniles. Mol. Plant Microbe. Interact. 1996; 9: 39-46.

[11] Brenner S. The genetics of the nematode *Caenorhabditis elegans*. Genetics 1974;77:71-94.

[12] Harlow E, Lane D. Antibodies - A laboratory manual. Cold Spring Habor Laboratory Press, New York. 1988; 726.

[13] Schagger H, Von-Jagow G. Sodium dodecylsulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100kD. Anal. Biochem. 1987; 166: 368-379.

[14] Curran J, Baillie DL, Webster JM. Use of genomic DNA restriction fragment length differences to identify nematode species. Parasitology 1985;90:137-144.

[15] Mello CC, Kramer JM, Stinchcomb D, Ambros V. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. The EMBO Journal 1991;10:3959-3970.

[16] Blumenthal T, Steward, K. RNA processing and gene structure. In: Riddle DL, Blumenthal T, Meyer BJ, Priess JR, editors. C. ELEGANS II. Cold Spring Harbor Laboratory Press, 1997; 117-146.

[17] Mello C, Fire A. DNA transformation. In: Shakes D, Epstein H, editors. Methods in Cell Biology*: C. elegans*. Academic Press, 1995; 478.

[18] Okkema PG, Harrison SW, Plunger V, Aryana A, Fire A. Sequence requirements for myosin gene expression and regulation in *Caenorhabditis elegans*. Genetics;1993: 385-404.

[19] Macmorris M, Broverman S, Greenspoon S, et al. Regulation of vitellogenin gene expression in transgenic *Caenorhabditis elegans*: short sequences required for activation of the vit-2 promoter. Mol. Cell. Biol. 1992;12: 1652-1662.

[20] Yarbrough PO, Hecht RM. Two isoenzymes of glyceraldehyde-3-phosphate dehydrogenase in *Caenorhabditis elegans*. J.Biol. Chem. 1984;259:14711-14720.

[21] Krause M, Hirsh D. A trans-spliced leader sequence on actin mRNA in *C. elegans*. Cell 1987;49: 753-761.

[22] Zorio DAR, Cheng NN, Blumenthal T, Spieth J. Operons as a common form of chromosomal organization in *C. elegans*. Nature 1994;372: 270-272.

[23] Stratford R, Shields R. A trans-spliced leader RNA sequence in plant parasitic nematodes. Mol. Biochem. Parasitol. 1994;67: 147-155.

[24] Blaxter ML. The filarial genome project. Parasitol. Today 1995;11: 811-812.

Chapter VII

Summary and concluding remarks

Qin, L., Smant,G., Bakker,J. and Helder,J
An adapted version of this chapter and chapter I is published in " **Biology of Plant-Microbe Interactions, Volume 3**"- 2001 IC-MPMI Congress Proceedings.
The identity and function of cyst nematode-secreted proteins in pathogenesis.

Summary and concluding remarks

Section 1.01 Summary

One of the fundamentals of this thesis is the development of a new strategy to identify pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*. This was needed because the strategy used so far - the antibody-based isolation and identification of putative pathogenicity-related proteins followed by the cloning of the corresponding cDNAs - appeared to be suitable for a limited number of pathogenicity factors only. As a starting point, three independent selection criteria were defined. Firstly, putative parthogenicity factors should be expressed solely or predominantly in the only infective stage of the potato cyst nematode, the pre-parasitic second stage juvenile. Secondly, pathogenicity-related genes were hypothesised to be specifically expressed in the esophageal glands of pre-parasitic second stage juveniles. Thirdly, pathogenicity factors will act outside the nematode and, as a consequence, the predicted proteins should be preceded by a signal peptide for secretion. It should be stressed that these criteria are not necessarily watertight; there could be pathogenicity factors that do not meet all criteria. However, this strategy allowed us to efficiently select putative pathogenicity factors among thousands of expressed genes. Moreover, application of these criteria resulted in the isolation of an unprecedented number of putative pathogenicity factors from this obligatory plant parasite.

cDNA-AFLP technology was used to compare mRNA pools from five highly synchronous different developmental stages. Visualisation of the expression patterns of thousands of genes allowed us to select for genes predominantly expressed in infective juveniles (the first selection criterion). The spatial expression patterns of the selected genes were then checked by *in situ* hybridization. Finally, genes that are both up-regulated in the infective J2 stage and specifically expressed in the dorsal esophageal gland of the infective J2 were checked for the presence of a signal peptide for secretion. This strategy resulted in the identification of the first three dorsal gland specific nematode secretory genes. Their presence in the dorsal glands of infective juveniles suggests that these proteins are probably involved in feeding cell initiation, and not in migration in the plant root or in protection against plant defence responses (Chapter 2).

cDNA-AFLP allows for the comparison of corresponding transcript-derived fragments (TDFs), relatively short cDNA fragments, between different mRNA pools. In order to link expression data directly to cDNA sequence data, and *vice versa*, a bi-directional link between TDF identifiers and cDNA sequence data (including ESTs) would be very useful. For this purpose, we have developed a computer program named GenEST that predicts the sizes of virtual transcript-derived fragments (TDFs) of *in silico* digested cDNA sequences retrieved from databases. The vast majority of the resulting virtual TDFs could be traced back among the thousands of TDFs displayed on cDNA-AFLP gels. As a consequence, cDNA sequence databases can be screened very efficiently to identify genes with relevant expression profiles. On the other hand, using the restriction enzyme recognition sites, the primer extensions and the estimated TDF size as identifiers, the DNA sequence(s) corresponding to a TDF with an interesting expression pattern can be readily identified from sequence databases. Using this powerful bioinformatics tool, various novel pathogenicity

factors were identified. We believe that this tool would be very useful for functional genomics studies in other systems as well (Chapter 3).

In Chapter 4, we have identified from the potato cyst nematode the first expansin gene outside land plants. The C-terminal half of the protein encoded by this nematode gene shares significant homology with plant β expansins, whereas the N-terminal half appears to be a family II cellulose-binding domain, which is specific for bacterial cellulases. Expansin activity was found in the infective second stage juveniles. Moreover, strongly enhanced expansin activity was detected in transgenic tobacco plant expressing this nematode expansin gene. The nematode expansin gene was expressed only in the subventral secretory gland and specifically up regulated in the infective second stage juveniles. A signal peptide for secretion is predicted for this nematode protein. This nematode expansin presumably facilitates the intracellular migration through plant roots by loosening plant cell walls.

Using the strategy as described in Chapter 2 in combination with the computer program developed in Chapter 3 resulted in the identification of RanBPM (Ran-Binding Protein in Microtubule organization)-like gene family with - at least - 8 members. These infective-stage specific genes are predominantly expressed in the dorsal esophageal gland of pre-parasitic second stage juveniles. The proteins encoded by these genes are probably secreted by the nematode into plant cells and might be involved in the syncytium induction by changing the dynamic instability of microtubules (Chapter 5).

To assess the relative importance of putative pathogenicity factors from the potato cyst nematode, a transformation protocol is urgently needed. Towards this end, we have cloned the promoter region of the glyceraldehyde-3-phosphate-dehydrogenase gene (GAPDH). We have shown that this promoter is also functional in *Caenorhabditis elegans*. This promoter region of GAPDH could be a valuable tool to drive gene expression in transgenic *G. rostochiensis* and other related plant-parasitic nematode species (Chapter 6).

Section 1.02 Assessing the importance of pathogenicity factors

It has become clear from this research and other studies that, like many fellow parasitic microbes, the plant parasitic nematodes produce a battery of pathogenicity factors to infect their host plants. More and more putative pathogenicity factors have been or will be cloned from various plant parasitic nematode species. The way in which these factors were identified implies that we mostly do not know their exact function nor their importance in pathogenicity. There is an urgent need for reliable methods allowing us to evaluate the importance of candidate parasitism genes. Here, I would like to discuss several possible approaches.

(a) Nematode transformation

A transformation protocol of the potato cyst nematode would be very useful either to knockout one putative pathogenicity factor or test whether one protein may function as an avirulence factor. For several free-living nematode species, such as *C. elegans*, genetic transformation achieved by microinjection of DNA into the developing gonad of the nematode is a routine technique. However, for plant parasitic nematodes there are several major obstacles:

- they are obligatory plant parasites, which means it is not possible to raise them in large numbers on a defined medium in petri-dishes;

- they have a long life cycle, followed by a diapause of several months. So a cyst nematode life cycle takes months in stead of a couple of days as for *C. elegans*;
- there is no selective markers available for easy identification of the transformants;
- the female body is surrounded by a thick cuticle and no internal structures can be seen.

Despite these difficulties, many labs have attempted to transform these nematodes. In Chapter 6 of this thesis, the promoter of GAPDH gene was fused to a reporter gene green fluorescence protein (GFP) and this promoter was shown to be functional in *C. elegans*. Recently, this construct was delivered into the potato cyst nematode by biolistic bombardment and GFP expression was detected in treated preparasitic J2 (Jone Jones, personal communication). However, the transformation efficiency was rather low and further optimisation is needed.

(b) Double-stranded RNA

Introduction of double-stranded RNA (dsRNA) has been shown to specifically disrupt the activity of genes containing homologous sequences in *C. elegans* [1]. dsRNA can be delivered in several ways to achieve this gene silencing effect: (a) it can be micro-injected into the body cavity of the nematode; (b) bacteria transformed with a dsRNA-encoding construct can be fed to the nematode; (c) the nematode can be soaked in solutions containing dsRNA. Because of the difficulties inherent in micro-injecting plant parasitic nematodes, the first method is apparently not suitable. The second option is unpractical because obligatory plant parasitic nematodes do not feed on bacteria. To overcome this difficulty, host plants can be transformed with constructs encoding dsRNA. When nematodes are feeding on the transformed plants, they will automatically take up dsRNA. It was shown that the 21-23 nucleotide fragments of dsRNA are guiding targeted mRNA degradation [2;3]. Such small dsRNA molecules should easily pass the narrow stylet opening of the nematode. It is also worth testing whether soaking the plant parasitic nematodes in dsRNA solution will also have the same effect as for *C. elegans*.

(c) Inhibitory antibody

An ingenious approach is to use the so called 'plantibody' – to express *in planta* single chain antibody (scAb) that has high binding affinity to the candidate pathogenicity factors in order to inactivate these nematode proteins [4]. In this way, the functionality of putative pathogenicity factors could be assessed. It is probably feasible and likely to be beneficial to express both scAb and dsRNA at the same time and in the same nematode feeding sites in order to achieve the maximum silencing effect.

(d) Transient transformation I: PVX

As an alternative to knock out genes, putative pathogenicity factors can be over-expressed in plants either transiently or stably. The phenotypes resulted from over-expression of the nematode proteins could give an indication of its role in parasitism.

Virus expression vectors, such as the potato virus X (PVX) vector, can be used to transiently transform plants [5]. Several candidate parasitism genes identified from the potato cyst nematode either by cDNA-AFLP or EST technique (see Chapter 2, 3, 4, 5) have been cloned into the PVX vector and transformed into *Agrobacterium tumefaciens*. *Agrobacterium* colonies containing the recombinant constructs were used to infect several plant species, including *Nicotiana benthaminana* and tomato. In

some cases, the infected plants displayed unusual phenotypes (Ling Qin & Aska Goverse, unpublished). However, these observations were complicated by the fact that the expression of non-pathogenicity-related genes in some cases also inflicted similar symptoms on plants. It might be that under viral infection, plants may respond abnormally to the high level expression of certain foreign proteins. Therefore, it is hard to draw firm conclusions from such experiments.

(e) Transient transformation II: protoplasts

The difficulties related to virus-mediated transient gene expression can be avoided by the transformation of protoplasts with plasmids encoding putative pathogenicity factors. The resulting phenotype can then be interpreted without the complicating factor of a viral infection. If a certain intracellular process is suspected to be altered by the nematode gene products, this would be a powerful approach. However, the physiology of protoplasts is rather different as compared to that of intact cells within the plant. The applicability of this approach may be rather limited.

(f) Stable plant transformation

Although stable plant transformation is more time consuming than transient transformation, it has several advantages. First, it is easier to interpret the results because no viral infection is involved. Second, the effect of the ectopic expression of the nematode genes can be observed under natural conditions at various tissues and during the developmental process. A special example of this strategy is described in Chapter 4. To prove that the plant expansin homologue from the nematode has expansin activity, tobacco plant was transformed with the nematode expansin gene. Constitutive expression of *Gr-ExpB1* in tobacco resulted in a significant increase of expansin activity. In this way, we were able to confirm the nature of this putative pathogenicity factor.

Section 1.03 The impact of genomics and functional genomics on plant nematology research

Development of new and powerful technologies has always been the driving force for the progresses in biological research. Plant nematology research is no exception. Equipped with microscopes, plant nematologists were able to observe the plant nematode interaction at the cellular level. With the aid of electron microscopes, many details of the intimate plant-nematode interaction have been revealed at the subcellular level. In recent years, various molecular biology techniques have enabled plant nematologists to study the molecular basis of this intriguing interaction. The trend in the past is going deeper and deeper from the whole organism down to single molecules. Nowadays, using genomics and functional genomics tools, one can not only study single molecules, but also can look wider at the same time – it is possible now to study a large number, if not the whole set, of genes simultaneously.

The availability of complete sequence information of all coding regions in a genome is crucial if we wish to obtain a fundamental, mechanistic insight in any particular organism. This has already been achieved for a number of organisms including the free-living nematode *C. elegans* [6-8], which serves as a model system for developmental biologists. The extreme high cost (estimated 0.2 US$ per sequenced and annotated base, for a nematode genome of around 100 Mb, this means roughly 200 million dollars!) at this time prohibits the sequencing of a pathogen like

the potato cyst nematode or any other plant parasitic nematode for that matter. However, if the cost of sequencing decrease as drastically as it does today, maybe in ten years time, this will be feasible.

For the time being, the best way to discover a large number of genes with a modest investment is the expressed sequence tag (EST) approach [9]. Because many genes are developmentally regulated, it is better to construct cDNA libraries used for generating ESTs from several distinct life stages of the nematode, especially from the parasitic stages. A good quality gland-specific library would lead to the sequencing of a higher percentage of parasitism-related genes. Recently, a big EST project to sequence from several root knot nematode species (*Meloidogyne* spp.) and the soybean cyst nematode *Heterodera glycines* has started (http://www.nematode.net/). Around 15,000 ESTs will be generated from various stages from each of these nematodes and the EST data will be publicly available. This will be a very useful resource for further studies.

The limitation of EST approach is that it will not reveal the potential functions of the genes if they do no share homologies with functionally annotated genes in the database. It is therefore necessary to link ESTs with functional genomics techniques, which can reveal the expression patterns of a large number of genes. In Chapter 3, we have developed software GenEST to link ESTs with cDNA-AFLP expression profiles. Using GenEST, ESTs, which are predicted by the program to have interesting expression patterns, can be quickly identified. *Vice versa*, using the identifiers of a band displayed on a cDNA-AFLP gel, the DNA sequence(s) corresponding to this band can be obtained. As an alternative, the expression pattern of ESTs can also be analyzed using the ever-popular cDNA microarray technique.

Equally if not even more informative is the proteomics approach, *i.e.*, the profiles of a large number of proteins. Due to post-transcriptional regulation, mRNA stability and some other processes, the protein content is not necessarily the same as that of mRNA. Above all, it is the proteins that carry out various physiology functions. Therefore, the direct analysis of proteins bears more biological relevance. In the past, careful 2D-protein gels are used to analyze the changes of hundreds of polypeptide during various developmental stages and protein spots showing interesting pattern changes were found [10]. Due to technical difficulties at that time, it was not possible to determine the identity of these proteins. Today, the 2D-gels are more sensitive and have higher resolutions. The amino acid sequences from minute amounts of proteins can be readily determined by mass-spectrometry technology. Proteomics of plant nematology will certainly become an interesting and fruitful research area in the coming years.

Genomics and functional genomics are changing the way of conducting research in plant nematology. Choosing the right and developing new bioinformatics tools are becoming more and more important to organize the enormous amount of sequence data and expression profile data. Appropriate software is also helpful to discover the genes that are important for the pathogenicity of the nematode, as shown in Chapter 3 of this thesis.

References

1. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC: **Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans**. *Nature* 1998, **391:**806-811.
2. Zamore PD, Tuschl T, Sharp PA, Bartel DP: **RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals**. *Cell* 2000, **101:**25-33.
3. Hamilton AJ, Baulcombe DC: **A species of small antisense RNA in posttranscriptional gene silencing in plants**. *Science* 1999, **286:**950-952.
4. Schots A, De Boer JM, Schouten A, Roosien J, Zilverentant JF, Pomp H, Bouwman L, Overmars H, Gommers FJ, Visser B, Stiekema WJ, Bakker J: **'Plantibodies': a flexible approach to design resistance against pathogens**. *Neth.J.Pl.Path.* 1992, **Supplement 2:**183-191.
5. Hammond-Kosack KE, Staskawicz BJ, Jones JDG, Baulcombe DC: **Functional expression of a fungal avirulence gene from a modified potato virus X genome**. *Molecular Plant-Microbe Interactions* 1995, **8**: 181-185.
6. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG: **Life with 6000 genes**. *Science* 1996, **274:**546,563-546,567.
7. The C.elegans Sequencing Consortium: **Genome sequence of the nematode C. elegans: a platform for investigating biology.** *Science* 1998, **282:**2012-2018.
8. Kaul S, Koo HL, Jenkins J, Rizzo M, Rooney T, Tallon LJ, et al.: **Analysis of the genome sequence of the flowering plant Arabidopsis thaliana**. *Nature* 2000, **408:** 796-815.
9. Adams MD, Soares MB, Kerlavage AR, Fields C, Venter JC: **Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library**. *Nat.Genet.* 1993, **4:**373-380.
10. De Boer JM, Overmars HA, Bakker J, Gommers FJ: **Analysis of two-dimensional protein patterns from developmental stages of the potato cyst-nematode, *Globodera rostochiensis***. *Parasitology* 1992, **105:**461-474.

LIST OF PUBLICATIONS

1. **Qin,L.**, Prins,P., Jones,J.T., Popeijus,H., Smant,G., Bakker,J. and Helder,J. 2001. GenEST, a powerful bi-directional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP. Nucleic Acids Research 29:1616-1622

2. **Qin,L.**, Overmars,H., Helder,J., Popeijus,H., Rouppe Van Der Voort,J.N.A.M., 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.

3. **Qin,L.**, Smant,G., Stokkermans,J., Bakker,J., Schots,A., and Helder,J. 1998. Cloning of a trans-spliced glyceraldehyde-3-phosphate-dehydrogenase gene from the potato cyst nematode Globodera rostochiensis and expression of its putative promoter region in Caenorhabditis elegans. Molecular and Biochemical Parasitology 96:59-67.

4. **Qin,L.**, Smant,G., Stokkermans,J.P.W.G., Bakker,J., Schots,A., Helder,J. 1998. Characterization and functional analysis of a putative promoter region of a trans-spliced glyceraldehyde-3-phosphate-dehydrogenase gene from the potato cyst nematode *Globodera rostochiensis*. Nematologica 44 -p.564.

5. **Qin,L.**, Kudla,U., Goverse,A., Popeijus,H., Roze,E., Overmars,H., Smant,G., Schots,A., Bakker,J. and Helder,J. Identification of an expansin-like protein from the potato cyst nematode. Submitted.

6. Yan,Y.T., 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.

7. Meijer,A.H., de Kam,R.J., **Qin,L.** and Hoge,J.H.C. 1998. Osjab1: (Accession No. AF072849), A rice homolog of the human JAB1transcriptional coactivator (PGR98-149), Plant Physiol. 117: 1526

8. Meijer,A.H., Scarpella,E., Dijk,V.E., **Qin,L.**, Taal,A.C., Rueb,S., Harrington,S.E., McCouch,S.R., Schilperoort,R.A., Hoge,J.C. 1997. Transcriptional repression by Oshox1, a novel homeodomain leucine zipper protein from rice. Plant Journal 11:263-276

9. Zhang,K.J, Lin,T.J, **Qin,L**, Liu,G.T. 1995. Effect of schisanhenol on function and surface shape of rat neutrophils. The Proceedings of Chinese Pharmacology. 16(3):234-8.

10. Zhang,T.F, **Qin,L.** 1993. Progress in the study of protein kinase C. The Proceedings of the Physiological Sciences (Chinese) 24(2):123-6.

11. **Qin, L.**, Smant,G., Bakker,J. and Helder,J. 2001. The identity and function of cyst nematode-secreted proteins in pathogenesis. "Biology of Plant-Microbe Interactions, Volume 3" - 2001 IC-MPMI Congress Proceedings. In press.

12. **Qin,L.**, Prins,P., Jones,J.T., Popeijus,H., Smant,G., Bakker,J. and Helder,J. 2001. GenEST, a powerful bi-directional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP. Plant & Animal Genome IX Genomics – The international conference on the status of Plant & Animal Genome Research, San Diego, CA. –p.62

13. **Qin,L.**, Prins,P., Jones,J.T., Popeijus,H., Smant,G., Bakker,J. and Helder,J. 2001. GenEST, a powerful bi-directional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP. Agricultural Microbes Genome II, San Diego, CA.. –p.17

14. **Qin,L.**, Overmars,H., Smant,G., Rouppe Van Der Voort,J.N.A.M., van Koert,P., Schots,A., Bakker,J. and Helder,J. 2000. An efficient cDNA-AFLP-based strategy for the identification of putative pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*. 6th International Congress of Plant Molecular Biology. Quebec Canada. S22-80

15. **Qin,L.**, Overmars,H., Smant,G., Rouppe Van Der Voort,J.N.A.M., van Koert,P., Schots,A., Bakker,J. and Helder,J. 2000. An efficient cDNA-AFLP-based strategy for the identification of putative pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*. Proceedings of the 39th Annual Meetings of Society of Nematology. Quebec Canada. -p.60

16. **Qin,L**, Overmars,H., Smant,G., Bakker,J., Schots,A.., Helder,J. 1999. Identification of pathogenicity factors by large scale monitoring of gene expression in potato cyst nematodes using cDNA AFLP. Book of abstracts: 9th international congress Molecular Plant-Microbe Interactions, Amsterdam -p.81

17. **Qin,L**, Bakker,J., Schots,A.., Helder,J. 1999. Genome-wide gene expression monitoring using Anchored RNA fingerprinting: discover pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*. Book of abstracts: Cambridge Healthtech institute's Genome Series: Molecular biology's Role in Enhancing Agricultural Productivity. Crop Production and Crop Protection, Amsterdam. -p.800 - 801

18. **Qin,L.**, Smant,G., Stokkermans,J.P.W.G., Bakker,J., Schots,A., Helder,J. 1998. Characterization and functional analysis of a putative promoter region of a trans-spliced glyceraldehyde-3-phosphate-dehydrogenase gene from the potato cyst nematode *Globodera rostochiensis*. Proceedings of the 24th International Nematology Symposium. –p.95

ACKNOWLEDGEMENTS

The keyword of science nowadays is cooperation and team effort is often necessary to crack the problems we are facing. Everybody must be ready to help and to be helped, to inspire and to be inspired. During the four years of my thesis research, I am very fortunate to have been helped and inspired by many brilliant people.

Firstly, I appreciate very much the general guidance, support and trust from Jaap Bakker. His creative way of thinking, enthusiasm, encouragement and flexibility are very stimulating for my research. The green light from him and Hans Helder to my proposal to change the direction of my research at the beginning of my second year is crucial in the success of this thesis.

I have benefited a lot from the inspiring supervision of Hans Helder from the very beginning. His clear (helder) mind, preciseness and writing skills are behind all the papers I have written and also this thesis. I am fortunate to have the freedom he has given to me in conducting my research and enjoyed his easy manner and willingness to help. I can still clearly remember how he has helped me to go through the formidable red tape in order to start my AIO work (also many thanks to Fred Gommers). During our many travelling together, his master of foreign languages has certainly made a lot of things a breeze for me.

Without the first interview with Arjen Schots, I would not be able to write my thesis here now. His insightful suggestions and interests in nematode 'spits' have also contributed to my work. I am also very grateful that I am allowed to work in the lab of LMA.

The two intelligent, friendly and enthusiastic AIOs Geert Smant and Aska Goverse, whom I met on the first day of my visit to this lab, have been very helpful all through the years of my stay here. It is my pleasure to work with them on the same projects and to supervise several students together. I value very much of their sharing with me the passion for nematode 'spits' and their own experiences of the life as an AIO (so I would be prepared for the famous 'AIO dip'). I also appreciate the help I have received from Geert in writing the MPMI and NAR papers. I am looking forward to more fruitful cooperation with them in the future.

I am grateful to Hein Overmars for the beautiful *in situ* experiments he has performed for me (apparently the editors of MPMI shared the same aesthetic taste with us). It is always a lot of fun during our gezellige mensa dinners together.

Pjotr Prins is the one to whom I have written and from whom I have received the most e-mails in my entire surfing career. Only very recently have I realized that this is perhaps the most efficient way to communicate with genuine computer experts.

I would also like to thank my multi-cultural students: Erwin Roze, Anneta Karczmarek, Ehab Mohammed Abdella and Wouter Groeninck for their contributions to various

projects. The nice results from Ula Kudla, the first Ph.D. whom I have the honor to supervise, is crucial for chapter V.

The contribution from Jeroen Rouppe van der Voort, Paul van Koert and Peter van Zandvoort are critical for the success of cDNA-AFLP. Also my thanks go to Rikus Pomp for many practical things in the lab of LMA; to Jack Stokkermans for his help at the start of my research; to Hans Keller, Jan Roosien, Herman Popeijus and Makedonka Dautova for many interesting discussions.

I am grateful to Lisette Groeneveld for her efficient administrative support and for her speaking Dutch to me relentlessly all through the years. Annette Agelino has consolidated both and is the secret behind my fluency in speaking 'de kat krabt de krullen van de trap'.

Thanks to all in the laboratory at Nematology and LMA, which by now have grown too big for me to name them one by one. I have enjoyed very much the friendly, open and helpful atmosphere in the lab and I believe that it will remain like this in the future.

I have learned very much through cooperation with brilliant researchers outside the lab of nematology. My thanks to Christian Bachem (Plant Breeding, WU), Jeroen Nieuweland (Nijmegen University), Frank Takken (Phytopathology, WU), Jan Carette (Molecular Biology, WU), Stephen Sturzenbaum and Andrew Gane (Cardiff University).

I consider myself very fortunate that I could meet regularly with excellent scientists from the European and American nematology community to share ideas, enthusiasm and conduct stimulating discussions (often at beautiful places as well). My thanks to Lieve Gheysen, Tom Tydgat (Gent University), Jone Jones (SCRI) and the members of the 'legendary' spit group: Jan de Boer, Thomas Baum (Iowa State University), Dick Hussey (University of Georgia) and Rick Davis (North Carolina State University).

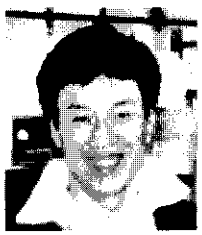
I am grateful to LEB fund and STW fund for their generous support that allowed me to attend many highly informative conferences.

While everybody gets older every year including myself (unfortunately), somehow my corridor people (Rijnsteeg 15A, Ya! Ya! Ya!) manage to get younger all the time. Thanks for their interests in my work.

I wish to thank my Chinese friends in Wageningen and else where and the Tulip Magazine editors for their kind interests.

Above all, I thank my dear wife, my parents, my sister and all other family members for their love, support and confidence in me. My last thank goes to my little boy, for his patience and understanding that allows me to stay selfishly for hours up in the attic and for his never-ending interest in my computer, which I am forever grateful, did not lead to the total loss of this thesis.

CURRICULUM VITAE



Ling Qin was born on October 18th 1967 in Beijing, China. He studied biochemistry and molecular biology at the Beijing University, where he obtained his Bachelor of Science degree as Biochemistry major in 1991. From 1991 to 1993, he worked as a research scientist at Institute of Materia Medica, China Academy of Medical Science on drug development research. From 1993 to 1995, he worked on various projects at China United Biotechnology Corporation.

He came to the Netherlands in 1995 and studied plant molecular biology at Leiden University, where he obtained his Master of Science degree in Biology with distinction (*cum laude*) in February 1997. In the same month, he was appointed as Assistant in Opleiding (as PhD researcher is called in the Netherlands) in the Laboratory of Nematology, Graduate School of Experimental Plant Science (EPS) of Wageningen University. The thesis is the result of four years of research carried out in this laboratory, which was financed by the European Community grants BIO4-CT96-0318 (ARENA-Basis and Development of Molecular Approaches to Nematode Resistance) and QLK5-1999-01501 (NONEMA).

Since April 2001, the author works as a Post-doc researcher in the same laboratory.

Ling Qin 9-4-98
Nema, LUW

Proposal and Research Plan

Best Opportunity for Identifying Secretions from Plant Parasitic Nematode

The hunt for nematode secretions is on in many different labs all over the world. The Mab approach started more than 10 years ago, has resulted in cellulase in our lab. But this method has serious limitations to be a productive way in getting the secretions and is also very time-consuming and labour intensive. Realising this, people now are trying alternatives at the DNA level like random EST done in Scotland, yeast signal peptide selection system in Hussey's lab. There is not enough data yet to evaluate these options, although it is already clear that certain disadvantages do exist. Here I propose, what in my eyes, a method that offers the best opportunity at this stage.

cDNA AFLP was developed recently in Plant breeding department in LUW (Plant J, 1996, 9(5), 745-753; the whole procedure is available at http://www.spg.wau.nl/pv/research.htm), which has certain advantages over the differential display method. And most importantly, Peter and Paul are very experienced in AFLP technique.

I intend to use this method to look for differentially expressed genes of cDNA populations between water- hatched (or water soaked cysts) and PRD- hatched J2 of G. ros. mRNA can readily be extracted from large quantities of J2 or cysts using Trizol and converted to cDNA primed with oligo-dT primer. RT-PCR could be done to check if expression of cellulases differs. If it does, this will give you some confidence that other pathogenecity-related factors might also be expressed at different levels, or more ideally-positive against negative.

Next, as Herman is trying or will try to make cDNAs from scaled-down material. If it works fine, we can also perform cDNA AFLP on two populations of anterior and posterior nematode cDNAs, or even on gland specific cDNAs.

If interesting differentially expressed bands can be seen, we can get the sequence information by directly sequencing PCR amplified DNA from gel. Then homology search in the data base may provide us with some useful information concerning the identity of this differentially expressed gene. Full length clone can be picked up from cDNA library using PCR. In situ experiments could generate more evidence.

* On 9-4-2001, exactly the same day three years later after this proposal , the author has finished writing this thesis and started a Post-doc research job. This may well be just a coincidence or maybe not.