Characterization of genes coding for small hypervariable peptides in *Globodera rostochiensis*

Nikkie E.M. van Bers

Promotor:	Prof. dr. ir. J. Bakker Hoogleraar in de Nematologie In het bijzonder de fysiologie en de moleculaire ecologie van nematoden, Wageningen Universiteit
Copromotoren:	Dr. ir. G. Smant Universitair Docent Laboratorium voor Nematologie, Wageningen Universiteit
	Dr. Ir. A. Goverse Universitair Docent Laboratorium voor Nematologie, Wageningen Universiteit
Promotiecommissie:	Prof. dr. W.J. Stiekema Wageningen Universiteit
	Dr. H. Franssen Wageningen Universiteit
	Dr. J. T.Jones Scottish Crop Research Institute, Dundee, Schotland
	Dr. ir. B.P.H.J. Thomma Wageningen Universiteit

Dit onderzoek is uitgevoerd binnen de onderzoeksschool “Experimental Plant Sciences”

Characterization of genes coding for small hypervariable peptides in *Globodera rostochiensis*

Nikkie E.M. van Bers

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. Dr. M.J. Kropff,
in het openbaar te verdedigen
op dinsdag 17 juni 2008
des morgens te 11.00 uur in de Aula.

Van Bers, N.E.M. (2008)

Characterization of genes coding for small hypervariable peptides in *Globodera rostochiensis*

PhD thesis Wageningen University, Wageningen, The Netherlands

With summaries in English and Dutch

ISBN: 978-90-8504-957-9

Table of contents

Chapter 1

General introduction

Nematodes

The term nematode literally means “thread-shaped”, and refers to the members of the phylum Nematoda or unsegmented roundworms. This terminology may seem contradictorily, but the latter is related to the cross-section of the nematode rather than to its gross appearance (Gibbons, 2002). Nematodes are diverse organisms, that are found in almost all habitats, adopt life styles ranging from free-living to both animal and plant parasitic, and range in length from microscopic to several metres (reviewed in (Gibbons, 2002), (Blaxter, 2003).

Plant parasitic nematodes

Plant parasitic nematodes are responsible for an estimated annual yield loss of billions of US dollars in agriculture worldwide (Sasser & Freckman, 1987). Among them, the root knot and cyst nematodes are considered economically most devastating. Root knot nematodes have a wide host range, while the host range of individual cyst nematode species is restricted to only one or a few related families of plant species (Evans & Stone, 1977). Cyst nematodes comprise at least 17 genera, of which the *Heterodera* and *Globodera* have been studied in most detail.

Life cycle

The life cycle of the cyst nematodes consists of four juvenile stages followed by the adult stage (Lee, 2002)Williamson, 1996 #26; Turner, 1998 #546; Zuncke, 1998 #542}. The first stage juvenile nematode is contained inside the egg shell (Figure 1a), and this is where its first moult, into a second stage juvenile or J2, takes place. Hatching is mostly stimulated by root exudates of host plants, which may also help the infective juvenile in localizing the host plant (Sharma & Sharma, 1998).

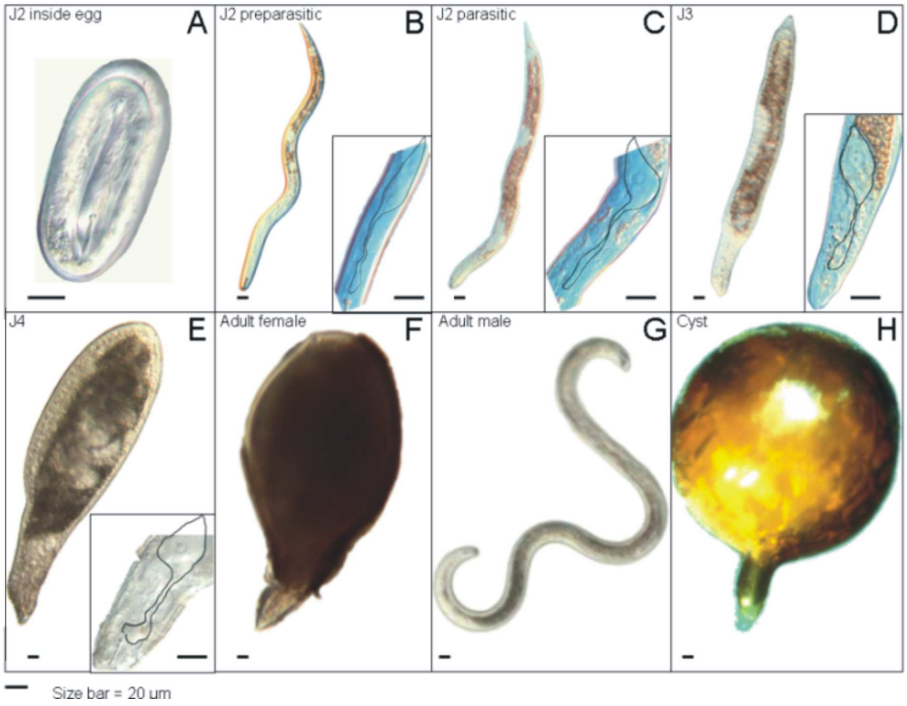


Figure 1: Life cycle of the potato cyst nematode. Microscopic images of: juvenile inside the egg (A), preparasitic juvenile (B) with an enlargement of the dorsal eosophageal gland (inset, dorsal gland encircled), J2 parasitic (C), J3 (D), J4 (E), adult female and male (F and G) and cyst (H).

Upon hatching the pre-parasitic J2 (Figure 1b) migrates towards and subsequently invades the plant root in the cell elongation zone (von Mende *et al.*, 1998). Cyst nematodes migrate intracellularly through root tissue by the enzymatic weakening of the cell wall structure and perforation of the cell walls with their oral stylet. Secretions produced in the two subventral esophageal glands and one dorsal esophageal gland that are released through the oral stylet assist in this enzymatic weakening of the cell wall. The plant cell wall is a complex composite of polysaccharides and proteins. Nematode migration is believed to be facilitated by the secretion of a cocktail of enzymes for modification of all major components of the plant cell wall (e.g. pectate lyases, expansins and endo-ß-1,4-glucanases)(reviewed in (Gheysen & Jones, 2006).

Soon after feeding cell induction, the nematode loses its ability to move and becomes sedentary, while feeding from the plant. During this feeding process the nematode undergoes three additional moults, into J3, J4 and finally into the adult stage (Figure 1d-1g, respectively). Adult males regain their mobility and leave the root to mate with the female that remains sedentary. Females retain the fertilized eggs inside their body wall, which eventually hardens to form a protective cyst (Figure 1h). Eggs inside the cyst can stay dormant for up to several years.

Plant-nematode interaction

Upon reaching the vascular cylinder the parasitic J2 (Figure 1c) transforms a competent cell, usually a procambium cell, into a feeding cell, called a syncytium. The feeding site induced by cyst nematodes, is the result of a progressive protoplast fusion of up to 200 cells as a consequence of local cell wall degradation. Although a role for enzymes of nematode origin in this cannot be excluded, it seems more likely that cell wall modifying proteins of plant origin are responsible for the cell wall degradation observed in syncytium formation (Goellner *et al.*, 2001, {Wieczorek, 2006 #804)A syncytium is a metabolically highly active cell complex, with a dense cytoplasm containing numerous organelles and small vacuoles. Its multiple nuclei are enlarged and amoeboid (Golinowski *et al.*, 1997, Golinowski *et al.*, 1996), resulting from endoreduplication, in which DNA synthesis occurs in the absence of nuclear or cellular division (Gheysen *et al.*, 1997).

Cell cycle

Cell cycle reactivation in host cells has been shown to be essential for feeding site formation (de Almeida Engler *et al.*, 1999). Several cell cycle related genes like *cdc2a*, *cycB1;1* and *cycA2;1* are induced in the early stages of syncytium formation (Goverse *et al.*, 2000a, Niebel *et al.*, 1996). The cell cycle in developing syncytia progresses at least until the end of the G2-phase, in which the cell organelles are duplicated

(Golinowski et al., 1997). Oryzalin treatment inhibits microtubule polymerization, thereby resulting in blockage of the cell cycle in late G2 phase, which prevents mitosis to occur. Smaller syncytia are formed in plants grown on medium containing oryzalin, indicating that mitosis is important in the process of syncytium formation. However, it is still controversial whether mitosis takes place *inside* the syncytial cell, or only in the peripheral cells that are being prepared for incorporation into the feeding cell complex (de Almeida Engler et al., 1999, Gheysen et al., 1997).

Plant hormones

The plant hormones auxin and cytokinin are considered to be key factors in controlling cell cycle progression in plants, i.e. by regulating the expression of cell-cycle related genes (Goverse et al., 2000a, Vanneste *et al.*, 2005). Exogenous application of the plant hormone auxin can mimic many of the cellular changes that occur in nematode-induced feeding sites, e.g. cell wall breakdown, endoreduplication, cell enlargement and lateral root initiation (reviewed in (Goverse *et al.*, 2000b). Auxin has been shown to be essential for feeding site formation in several studies. First, the auxin-insensitive tomato mutant *diageotropica* is *de facto* resistant to the potato cyst nematode, and both the auxin responsive promoter trap line 5-1E1 and the auxin responsive element *Dr5* are activated inside and around young syncytia induced by the beet cyst nematode *Heterodera schachtii* on *Arabidopsis thaliana* roots (Goverse et al., 2000b); (Karczmarek *et al.*, 2004). Furthermore, disturbed local auxin gradients result in abnormal feeding sites in which expansion seems to be inhibited, as was shown for the tomato-*Globodera rostochiensis* interaction by application of the auxin transport inhibitor NPA and by use of polar auxin transport mutants of *Arabidopsis* in its interaction with *H. schachtii* (Goverse et al., 2000b). Moreover, the *WRKY23* transcription factor, which is involved in the regulation of auxin flux, is highly expressed in early syncytia induced by the beet cyst nematode (Barthels *et al.*, 1997).

Cyst nematodes flourish on the *Arabidopsis* overproducing mutants *eto1—3* (Goverse

et al., 2000a). This indicates that ethylene promotes feeding cell formation, which may be mediated by facilitating cell wall degradation or by regulating auxin transport and accumulation (Goverse et al., 2000a).

Nematode effectors

Nematode effectors are the products of genes coding for secreted proteins that are involved in parasitism of the nematode. The amphids and the large subventral and dorsal esophageal glands of the cyst nematode are its main secretory organs. The esophageal glands are packed with secretory granules containing putative effector molecules. The oral stylet delivers the nematode effectors into the host cell apoplast and cytoplasm. Some of these effectors may be recognized by the plant immune system resulting in a potent defense response, while others may mask recognition of the nematode by the plant defense system or actively suppress plant defense (see below). Furthermore, (a subset of) the effectors will promote virulence by facilitating the ingestion of plant material or by the reprogramming of host cells into feeding cells (compatible interaction) (Jones & Robertson, 1997, Hussey, 1989, Wyss, 1992).

Nematode effectors involved in feeding cell formation

At the onset of feeding, the number of secretory granules in the subventral esophageal gland decreases, indicating that this gland mainly produces factors facilitating hatching and migration (see section **migration**). Upon feeding site formation, the number of granules in the dorsal esophageal gland increases (compare figure 1b to 1c-e), indicating the importance of this gland in the actual formation of the feeding site (Hussey & Grundler, 1998).

Bioactive nematode peptides

Goverse *et al* showed in 1999 that the fraction of small proteins (<3kDa, which is about 30 amino acids) present in the secretions of *Globodera rostochiensis* co-stimulate cell

proliferation of tobacco protoplasts. Co-stimulation is a term used for growth factors which solely do not induce cell proliferation, but enhance the proliferative response of cells upon mitogenic stimulation.

Different conventions exist on the definition of a peptide (Wikipedia online dictionary). One convention is that peptide chains that are short enough to make synthetically are called peptides rather than proteins. However, with the improvement of techniques for the generation of synthetic peptides, this definition seems to be outdated. A somewhat less arbitrary definition is that a peptide is an amino-acid molecule without secondary structure, while another convention defines peptides as proteins consisting of <50-100 amino-acids. Here, we will use the term peptide when referring to small proteins (<150 amino-acids) without any predicted secondary structure, or for small proteins showing sequence similarity to a peptide.

The *G. rostochiensis* peptides required the presence of auxin and cytokinin, but did not significantly increase the sensitivity of the protoplasts to the plant hormones at the concentrations used.

Several genes encoding small proteins have been shown to be expressed in the dorsal esophageal glands of cyst nematodes. SYV-46 is a 139 amino acid polypeptide, expressed in parasitic stages of *Heterodera glycines*. The polypeptide is unique to *H. glycines*, however it does share a C-terminal motif with the CLAVATA3/ESR-related (CLE) protein family in *Arabidopsis* (Olsen & Skriver, 2003, Wang *et al.*, 2001). CLV3, a founding member of the CLE protein family, functions as a negative regulator of *WUSCHEL* (*WUS*). *WUS* is a transcription factor, restricting the size of the stem cell population in the shoot and floral meristems (Cock & McCormick, 2001, Schoof *et al.*, 2000). Expression of *Hg-SYV46* rescued the mutant phenotype of the clv3-1 mutant, suggesting a similar functionality for both peptides (Wang *et al.*, 2005).

The ubiquitin extension protein is a peptide of about 122 amino acids, consisting of a N-terminal signal peptide for secretion, a central highly conserved mono-ubiquitin domain, and a variable positively charged C-terminal domain (Tytgat *et al.*, 2004).

Fusion of the ubiquitin extension protein of *Heterodera schachtii*, Hs-ubi1, with GFP shows that the C-terminal domain is targeted to the nucleolus. The nucleolus is a subcompartment of the nucleus and its main function is in ribosome biogenesis, but it is also involved in maturation of some tRNAs and plays a role in cell cycle regulation and small RNA formation (Raŝka *et al.*, 2006).

Plant defense

Plants employ an elaborate array of defensive systems against pathogens of which the first one to be encountered by pathogens is the physical barrier provided by the waxy cuticle and the cell wall. Additional to this first line of defense, plants have an immune system that shares similarities with the animal innate immune system (Ausubel, 2005). In 2006, Jones and Dangl presented a *zigzag* model illustrating that the output of the plant immune system depends on the interplay between plant defense and pathogen effector molecules (figure 2). Plant pattern recognition receptors (PRRs) perceive so-called conserved pathogen-associated molecular patterns (PAMPs), like flagellin or chitin, thereby inducing PAMP-triggered immunity (PTI)(phase 1, figure 2)(Jones & Dangl, 2006). PTI is also known as basal defense and occurs early in the plant-pathogen interaction (<10 minutes after contact). Basal defense is featured by a.o. callose deposition, production or accumulation of ethylene and reactive oxygen and nitrogen species and change in gene expression (Ingle *et al.*, 2006). Reactive oxygen and nitrogen species may be directly damaging to the pathogen, but they are also important molecules in plant defence signalling pathways.

A more specific form of defense is the R-gene mediated effector triggered immunity (ETI), which occurs later in the host-pathogen interaction (2-3 hours), upon delivery of pathogen effectors into the host cytoplasm or apoplast (phase 3, figure 2). Effectors that elicit ETI are known as avirulence effectors. ETI shares some similar features with PTI, and may even share some molecular meachanisms (Abramovitch *et al.*, 2006). However, the most characteristic feature of ETI, which is generally not observed in PTI,

is the localized programmed cell death known as a hypersensitive response (HR).
Recognition of pathogen effectors by R-gene products can be described by two models, the gene-for-gene model and the guard-model. In a gene-for-gene interaction, a resistance gene is only effective if a specific effector is produced by the pathogen. In the guard model, presence of the pathogen is sensed by R-gene products through a modified state of other (guarded) host proteins. These host proteins may be the virulence target of the effector protein in compatible interactions.

Defense in the plant-cyst nematode interaction

The interaction between non-host *Arabidopsis thaliana* and the soybean cyst nematode *Heterodera glycines* is characterized by a prolonged invasion and migration period, and browning of the tissue at the infection site. The juveniles that manage to reach the vascular cilinder are confronted with a strong HR, which leaves them almost no possibility of selecting a suitable initial feeding cell. Additionally, callose or callose-like material is deposited in the vascular cilinder and around the nematodes' head, thereby trapping nematodes that are completely inside the root (Grundler *et al.*, 1997). Furthermore, hydrogen peroxide (H_2O_2) is produced by damaged cells (wound response) but also by cells that are neither in contact with the nematode nor with the syncytium (Waetzig *et al.*, 1999).
Host resistance against cyst nematodes is often characterized by local cell death in cells at the periphery of the initial feeding cell, thereby inhibiting the expansion of the syncytium. The timing of this response suggests that the avirulent nematode effectors are recognized early in feeding site formation (Cabrera Poch *et al.*, 2006).
The few nematodes that do develop on plants showing a defense response are mostly male, which is a sign of poor nutrition for the nematode (Williamson & Hussey, 1996).
Up to now, 19 R-genes have been mapped, conferring resistance against potato cyst nematodes (Caromel *et al.*, 2003, Bakker *et al.*, 2003, Rouppe Van Der Voort *et al.*, 1998, Rouppe Van Der Voort *et al.*, 1997, Kreike *et al.*, 1996, Kreike *et al.*, 1994 and

Bakker, 2002). Interestingly, those genes belong to two different classes of NB-LRR-genes, either the class of the LZ-NB-LRR proteins, containing a leucine zipper (LZ) domain, a nucleotide binding domain (NB) and a leucine rich repeat region (LRR) or to the class of Tir-NB-LRR proteins containing a TIR (Toll-interleukin 1 receptor) domain (Williamson & Kumar, 2006, Ausubel, 2005).

The H1 resistance gene against *Globodera rostochiensis* and its as yet unidentified elicitor comply with the gene-for-gene model, which explains the recognition specificity of major R genes. Recently, the Rbp-1 protein from *Globodera pallida* was shown to induce a HR on *Nicotiana benthamiana* leaves expressing Gpa2, a resistance gene conferring resistance to *G. pallida* (Moffet, 2007;Sacco *et al.*, 2007). Rbp-1 shows strong homology to the SPRY domain of the Ran-binding protein in the microtubule-organizing center (RanBPM), which is potentially involved in protein-protein interactions (Rehman, 2008). RanGTPase Activating Protein (RanGAP) interacts with the CC domain of Gpa2 (Sacco *et al.*, 2007). This interaction may lead to changes in RanGAP related targets, ultimately inducing a HR response. RanGAP and RanGTPase are involved in the hydrolysis of RanGTP in RanGDP, which at least in some cases is important for the release of Ran from its receptor (Steggerda & Paschal, 2002). RanGTP is involved in nuclear trafficking and cell plate and mitotic spindle formation. A large family of genes with sequence similarity to *Rbp*-1, the *SPRYSEC* family, has been shown to be specifically expressed in the dorsal esophageal gland of *G. rostochiensis* (Qin, 2001, Qin *et al.*, 2000)(Rehman, PhD thesis).

Avoidance or suppression of host defense

Microbes succesfull in pathogenicity deploy a wide array of effectors to succesfully interfere with the plant immune system, which is called Effector Triggered Susceptibility (ETS)(phase 2 and 4, figure 2)(reviewed in (Abramovitch & Martin, 2004) and (Jones & Dangl, 2006).

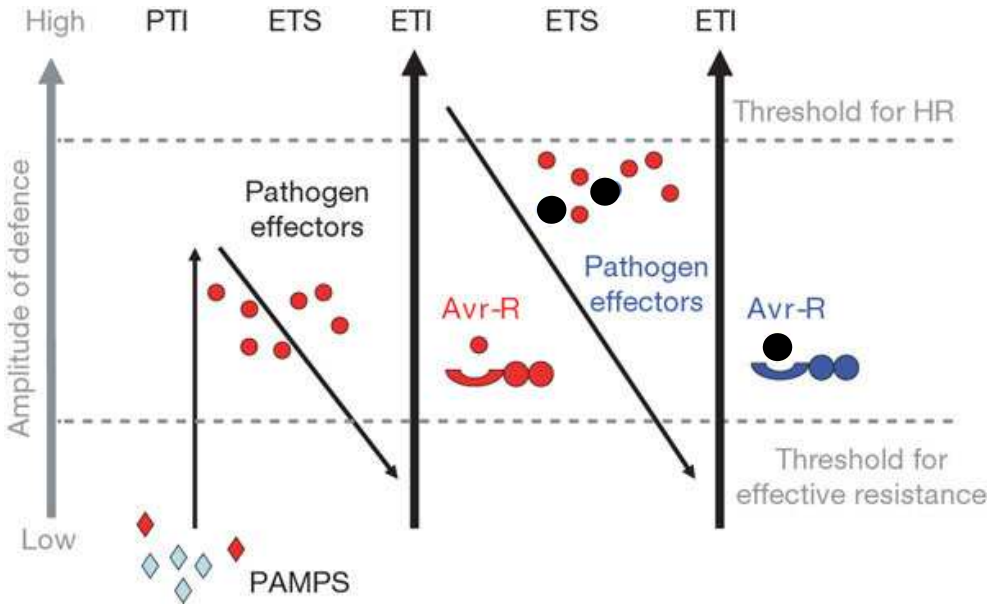


Fig 2: A zigzag model illustrates the quantitative output of the plant immune system. In phase 1, the plant detects pathogen associated molecular patterns (PAMPs, diamonds) to trigger pathogen-triggered immunity (PTI). In phase 2, succesful pathogens deliver effectors (circles that interfere with PTI, or otherwise enable pathogen nutrition and dispersal, resulting in effector-triggered susceptibility (ETS). In phase 3, one effector is recognized by an NBS-LRR protein, activating effector-triggered immunity (ETI), which is an amplified version of PTI. In phase 4, pathogen isolates are selected that have lost the recognized effector, and perhaps gained new effectors (black circle) that can help pathogens to suppress ETI. Selection favours new NB-LRR alleles that can recognize one of the newly acquired effectors, resulting again in ETI (phase 3)(Figure is adapted from Jones &Dangl, 2006).

As sedentary nematodes reside for an extended period of time in extremely close contact with their host, it is expected that the nematode secretes effectors involved in avoidance or suppression of host defense (Williamson & Kumar, 2006). This secretion can be either at its own surface or into the host cell cytoplasm or apoplast. The nematode cuticle surface is overlaid with a coat largely made up of proteins. This surface coat may mimic host tissues in an attempt to avoid eliciting a defense response (Jones & Robertson, 1997), which may ultimately result in ETS. The

composition of the nematode surface coat seems to change during nematode development, and the surface coat of *G. rostochiensis* is shown to contain several antioxidant enzymes, like thioredoxin peroxidase, peroxiredoxin and superoxide dismutase (Speigel, 1995; (Jones *et al.*, 2004, Robertson *et al.*, 1999, Robertson *et al.*, 2000). Furthermore, a glutathione peroxidase is expressed in the hypodermis of *G. rostochiensis* juveniles, from where it is probably secreted onto the nematode surface coat (Jones et al., 2004).

Aditionally, cyst nematodes may suppress plant defense by secreting effectors into the host cell apoplast or cytoplasm. Chorismate mutase (CM) is a secreted nematode effector produced in the esophageal gland of both root-knot and cyst nematodes (Lambert *et al.*, 1999). CM may function in manipulating the plant's shikimate pathway, which is, besides controlling cell growth and development, involved in defense (Bekal *et al.*, 2003). CM may suppress plant defense by breaking down chorismate-derived compounds, like salicylic acid or phenolic phytoalexins.

A class of genes found in several cyst and root knot nematode species, shows significant homology to allergen antigen 5 found in hymenopteran insect venom (Gao *et al.*, 2001), and the gene products are therefore called venom allergen like proteins (vap). The exact function of *vap* in parasitism is unknown, but it shows sequence similarity to both pathogenesis related protein 1 (PR1) from plants and Ion-1 in *C. elegans* (Morita *et al.*, 2002). PR1 is one of the key components in the salicylic acid mediated defense signalling in plants.

Sequence diversification

Genes encoding pathogen effectors as well as host defense genes are subject to strong evolutionary pressures. Pathogens overcome recognition by host defense genes by losing the cognate avirulent effector. This may be achieved by rapid sequence variation; however, the extent of this variation is constraint by its impact on effector functionality. Host defense genes will have to keep up with the changing effector

molecules. Evidence for the rapid sequence divergence of both effector and defense molecules when involved in direct interaction (Jones & Dangl, 2006, Dodds *et al.*, 2006)(see also chapter 6 of this thesis). Accelerated evolution in effector genes is found in spider and scorpion venom peptides. Similarly, avr genes of fungi and oomycetes show a high sequence divergence (Escoubas, 2006, Zhijian *et al.*, 2006, Dodds et al., 2006, Staats *et al.*, 2007).
Antimicrobial peptides and some R genes, like the L genes from flax rust, are defense molecules undergoing accelerated evolution to cope with the ever changing effector molecules of pathogens (Nicolas *et al.*, 2003, Dodds et al., 2006).

SL-transsplicing

Numerous nematode genes undergo two different forms of splicing. In addition to *cis*-splicing, which involves in the removal of introns from pre-mRNA, nematode genes undergo *trans*-splicing. In *trans*-splicing, a 22 or 23 nucleotide leader sequence, known as spliced leader or SL, is spliced onto the 5'-end of the pre-mRNA of a wide range of nematode transcripts. In *Caenorhabditis elegans*, trans-splicing is involved in translation initiation, and may contribute to the stability of the mRNA. Furthermore, *trans*-splicing is essential for resolving the pre-mRNA of genes organized in operons (Blumenthal, 1998). SL sequences are also found on the 5'-end of transcripts of many plant parasitic nematodes (Stratford & Shields, 1994, Ray *et al.*, 1994, Mitreva *et al.*, 2004).

Outline of this thesis

In the past decade, a wide variety of approaches have been applied successfully to identify putative effectors in plant parasitic nematodes. This resulted in a large amount of gene sequences, some concerning effectors with a known function, but the majority of the sequences identified has no significant sequence similarity to proteins functionally annotated in the database. Unraveling the functional role of those pioneers is the challenge that started only recently.

In chapter 2 we study some of these pioneers, which are encoded by the *SECPEP1* gene family. Sequence analysis and the expression profile of this gene family resulted in a hypothetical function for the gene products in feeding site formation. This function is further studied for *SECPEP3* by ectopic *in planta* (over)expression in *Arabidopsis* and potato, described in chapter 4.

In chapter 3 we describe the discovery and functional analysis of *NEMPEP*. *NEMPEP* encodes for an unrelated peptide gene, which is believed to be involved in the parasitic interaction with the plant.

Nearly all known members of the *SECPEP* gene family are subject to trans-splicing. In chapter 5 we analyse the diversity of splice leader sequences trans-spliced onto the 5'-end of *SECPEP1*.

References

Abramovitch, R. B., Anderson, J. C. and Martin, G. B. (2006) Bacterial elicitation and evasion of plant innate immunity. *Nat Rev Mol Cell Biol,* **7,** 601-611.

Abramovitch, R. B. and Martin, G. B. (2004) Strategies used by bacterial pathogens to suppress plant defenses. *Current Opinion in Plant Biology,* **7,** 356-364.

Ausubel, F. M. (2005) Are innate immune signaling pathways in plants and animals conserved? *Nature Immunology,* **6,** 973-979.

Bakker, E., Butterbach, P., Rouppe Van Der Voort, J., Van Der Vossen, E., Van Vliet, J., Bakker, J.*, et al.* (2003) Genetic and physical mapping of homologues of the virus resistance gene Rx1 and the cyst nematode resistance gene Gpa2 in potato. *Theoretical and Applied Genetics,* **106,** 1524-1531.

Bakker, J. (2002) Durability of resistance against potato cyst nematodes. *Euphytica,* **124,** 157-162.

Barthels, N., Van Der Lee, F. M., Klap, J., Goddijn, O. J. M., Karimi, M., Puzio, P.*, et al.* (1997) Regulatory sequences of Arabidopsis drive reporter gene expression in nematode feeding structures. *Plant Cell,* **9,** 2119-2134.

Bekal, S., Niblack, T. L. and Lambert, K. N. (2003) A chorismate mutase from the soybean cyst nematode Heterodera glycines shows polymorphisms that correlate with virulence. *Mol Plant Microbe Interact,* **16,** 439-446.

Blaxter, M. L. (2003) Nematoda: Genes, Genomes and the Evolution of Parasitism. *Advances in Parasitology,* **54,** 101-195.

Blumenthal, T. (1998) Gene clusters and polycistronic transcription in eukaryotes. *Bioessays,* **20,** 480-487.

Cabrera Poch, H. L., Manzanilla Lo?pez, R. H. and Kanyuka, K. (2006) Functionality of resistance gene Hero, which controls plant root-infecting potato cyst nematodes, in leaves of tomato. *Plant, Cell and Environment,* **29,** 1372-1378.

Caromel, B., Mugnie?ry, D., Lefebvre, V., Andrzejewski, S., Ellisse?che, D., Kerlan, M. C.*, et al.* (2003) Mapping QTLs for resistance against Globodera pallida (Stone) Pa2/3 in a diploid potato progeny originating from Solanum spegazzinii. *Theoretical and Applied Genetics,* **106,** 1517-1523.

Cock, J. M. and McCormick, S. (2001) A large family of genes that share homology with CLAVATA3. *Plant Physiol,* **126,** 939-942.

de Almeida Engler, J., De Vleesschauwer, V., Burssens, S., Celenza, J. L., Jr., Inze, D., Van Montagu, M.*, et al.* (1999) Molecular markers and cell cycle inhibitors show the importance of cell cycle progression in nematode-induced galls and syncytia. *Plant Cell,* **11,** 793-808.

Dodds, P. N., Lawrence, G. J., Catanzariti, A. M., Teh, T., Wang, C. I. A., Ayliffe, M. A.*, et al.* (2006) Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences of the United States of America,* **103,** 8888-8893.

Escoubas, P. (2006) Molecular diversification in spider venoms: A web of combinatorial peptide libraries. *Mol Divers,* **10,** 545-554.

Evans, K. and Stone, A. R. (1977) A Review of the Distribution and Biology of the Potato

Cyst-Nematodes *Globodera rostochiensis* and *G. pallida*. *PANS,* **23,** 178-189.

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

Gheysen, G., Engler, J. D. A. and Montagu, M. V. (1997) Cell cycle regulation in nematode feeding sites. In: *Cellular and molecular aspects of plant-nematode interactions.* (Fenoll, C., Grundler, F. M. W. and Ohl, S. A., ed.^eds.). Kluwer Academic Publishers, pp. 120-132.

Gheysen, G. and Jones, J. T. (2006) Molecular Aspects of Plant-Nematode Interactions. In: *Plant Nematology.* (Perry, R. N. and Moens, M., ed.^eds.). Wallingford: CABI, pp. 234-254.

Gibbons, L. M. (2002) General Organisation. In: *The Biology of Nematodes.* (Lee, D. L., ed.^eds.). London: Taylor and Francis, pp. 31-59.

Goellner, M., Wang, X. and Davis, E. L. (2001) Endo- beta -1,4-glucanase expression in compatible plant-nematode interactions. *Plant Cell. 2001;,* **13,** 2241-2255.

Golinowski, W., Grundler, F. M. W. and Sobczak, M. (1996) Changes in the structure of Arabidopsis thaliana during female development of the plant-parasitic nematode Heterodera schachtii. *Protoplasma,* **194,** 103-116.

Golinowski, W., Sobczak, M., Kurek, W. and Grymaszewska, G. (1997) The structure of syncytia. In: *Cellular and molecular aspects of plant-nematode interactions.* (Fenoll, C., Grundler, F. M. W. and Ohl, S. A., ed.^eds.). Kluwer Academic Publishers, pp. 80-97.

Goverse, A., de Engler, J. A., Verhees, J., van der Krol, S., Helder, J. H. and Gheysen, G. (2000a) Cell cycle activation by plant parasitic nematodes. *Plant Mol Biol,* **43,** 747-761.

Goverse, A., Overmars, H., Engelbertink, J., Schots, A., Bakker, J. and Helder, J. (2000b) Both induction and morphogenesis of cyst nematode feeding cells are mediated by auxin. *Mol Plant Microbe Interact,* **13,** 1121-1129.

Grundler, F. M. W., Sobczak, M. and Lange, S. (1997) Defence responses of Arabidopsis thaliana during invasion and feeding site induction by the plant-parasitic nematode Heterodera glycines. *Physiological and Molecular Plant Pathology,* **50,** 419-429.

Hussey, R. S. (1989) *Disease-inducing secretions of plant-parasitic nematodes*.

Hussey, R. S. and Grundler, F. M. W. (1998) Nematode parasitism of plants. In: *The Physiology and Biochemistry of Free-living and Plant-parasitic Nematodes.* (Perry, R. N. and Wright, D. J., ed.^eds.). CAB International, pp. 213-243.

Ingle, R. A., Carstens, M. and Denby, K. J. (2006) PAMP recognition and the plant-pathogen arms race. *BioEssays,* **28,** 880-889.

Jones, J. D. G. and Dangl, J. L. (2006) The plant immune system. *Nature,* **444,** 323-329.

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

Jones, J. T. and Robertson, W. M. (1997) Nematode secretions. In: *Cellular and molecular aspects of plant-nematode interactions.* (Fenoll, C., Grundler, F. M. W. and Ohl, S. A., ed.^eds.). Kluwer Academic Publishers, pp. 98-106.

Karczmarek, A., Overmars, H., Helder, J. and Goverse, A. (2004) Feeding cell development by cyst and root-knot nematodes involves a similar early, local and transient activation of a specific auxin-inducible promoter element. *Molecular Plant Pathology,* **5,** 343-346.

Kreike, C. M., de Koning, J. R. A., Vinke, J. H., van Ooijen, J. W. and Stiekema, W. J. (1994) Quantitatively-inherited resistance to Globodera pallida is dominated by one major locus in Solanum spegazzinii. *Theoretical and Applied Genetics,* **88,** 764-769.

Kreike, C. M., Kok-Westeneng, A. A., Vinke, J. H. and Stiekema, W. J. (1996) Mapping of QTLs involved in nematode resistance, tuber yield and root development in Solanum sp. *Theoretical and Applied Genetics,* **92,** 463-470.

Lambert, K. N., Allen, K. D. and Sussex, I. M. (1999) Cloning and characterization of an esophageal-gland-specific chorismate mutase from the phytoparasitic nematode Meloidogyne javanica. *Mol Plant Microbe Interact,* **12,** 328-336.

Lee, D. L. (2002) Life cycles. In: *The Biology of Nematodes.* (Lee, D. L., ed.^eds.). London: Taylor and Francis, pp. 61-72.

Mitreva, M., Elling, A. A., Dante, M., Kloek, A. P., Kalyanaraman, A., Aluru, S.*, et al.* (2004) A survey of SL1-spliced transcripts from the root-lesion nematode Pratylenchus penetrans. *Mol Genet Genomics,* **272,** 138-148.

Moffet, P. (2007) Recognition and activation by the NB-LRR proteins Rx and Gpa2; involvment of a Ran GTPase-activating protein. XIII International Congress on Molecular Plant-Microbe Interacions, July 21-27, Sorrento, Italy

Morita, K., Flemming, A. J., Sugihara, Y., Mochii, M., Suzuki, Y., Yoshida, S.*, et al.* (2002) A Caenorhabditis elegans TGF-beta, DBL-1, controls the expression of LON-1, a PR-related protein, that regulates polyploidization and body length. *Embo J,* **21,** 1063-1073.

Nicolas, P., Vanhoye, D. and Amiche, M. (2003) Molecular strategies in biological evolution of antimicrobial peptides. *Peptides,* **24,** 1669-1680.

Niebel, A., de Almeida Engler, J., Hemerly, A., Ferreira, P., Inze, D., Van Montagu, M.*, et al.* (1996) Induction of cdc2a and cyc1At expression in Arabidopsis thaliana during early phases of nematode-induced feeding cell formation. *Plant J,* **10,** 1037-1043.

Olsen, A. N. and Skriver, K. (2003) Ligand mimicry? Plant-parasitic nematode polypeptide with similarity to CLAVATA3. *Trends in Plant Science,* **8,** 55-57.

Qin, L. (2001) Molecular genetic analysis of the pathogenicity of the potato cyst nematode *Globodera rostochiensis* In: *Laboratory for Nematology.* (ed.^eds.). Wageningen: Wageningen University.

Qin, L., Overmars, H., Helder, J., Popeijus, H., van der Voort, J. R., Groenink, W.*, et al.*

(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.

Raŝka, I., Shaw, P. J. and Cmarko, D. (2006) New Insights into Nucleolar Architecture and Activity. In: *International Review of Cytology.* (ed.^eds.). pp. 177-235.

Ray, C., Abbott, A. G. and Hussey, R. S. (1994) Trans-splicing of a Meloidogyne incognita mRNA encoding a putative esophageal gland protein. *Molecular and Biochemical Parasitology,* **68,** 93-101.

Rehman, S., Qin, L., Overmars, H., Petrescu, A. J., Thomma, B. P. H. J., Helder, J.*, et al.* (2008) Overexpression of a nematode SPRYSEC-gene in potato promotes the virulence of a nematode, a fungus and a virus. *in PhD thesis, Wageningen University, the Netherlands***,** 74-107.

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

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

Rouppe Van Der Voort, J., Lindeman, W., Folkertsma, R., Hutten, R., Overmars, H., Van Der Vossen, E.*, et al.* (1998) A QTL for broad-spectrum resistance to cyst nematode species (Globodera spp.) maps to a resistance to gene cluster in potato. *Theoretical and Applied Genetics,* **96,** 654-661.

Rouppe Van Der Voort, J., Wolters, P., Folkertsma, R., Hutten, R., Van Zandvoort, P., Vinke, H.*, et al.* (1997) Mapping of the cyst nematode resistance locus Gpa2 in potato using a strategy based on comigrating AFLP markers. *Theoretical and Applied Genetics,* **95,** 874-880.

Sacco, M. A., Mansoor, S. and Moffett, P. (2007) A RanGAP protein physically interacts

with the NB-LRR protein Rx, and is required for Rx-mediated viral resistance. *Plant Journal,* **52,** 82-93.

Sasser, J. N. and Freckman, D. W. (1987) A world perspective on nematology: The role of the society. In: *Vistas on nematology.* (Veech, J. A. and Dickson, D. W., ed.^eds.). Hyattsville, Maryland: Society of Nematologists, pp. p. 7-14.

Schoof, H., Lenhard, M., Haecker, A., Mayer, K. F. X., Ju?rgens, G. and Laux, T. (2000) The stem cell population of Arabidopsis shoot meristems is maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell,* **100,** 635-644.

Sharma, S. and Sharma, R. (1998) Hatch and emergence. In: *The cyst nematodes.* (Sharma, S., ed.^eds.). Dordrecht: Kluwer Academic Publishers.

Staats, M., van Baarlen, P., Schouten, A., van Kan, J. A. L. and Bakker, F. T. (2007) Positive selection in phytotoxic protein-encoding genes of Botrytis species. *Fungal Genetics and Biology,* **44,** 52-63.

Steggerda, S. M. and Paschal, B. M. (2002) Regulation of nuclear import and export by the GTPase ran. *International Review of Cytology,* **217,** 41-91.

Stratford, R. and Shields, R. (1994) A trans-spliced leader RNA sequence in plant parasitic nematodes. *Mol Biochem Parasitol,* **67,** 147-155.

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

Vanneste, S., Maes, L., De Smet, I., Himanen, K., Naudts, M., Inze, D.*, et al.* (2005) Auxin regulation of cell cycle and its role during lateral root initiation. *Physiologia Plantarum,* **123,** 139-146.

von Mende, N., Gravato-Nobre, M. J. and Perry, R. N. (1998) Host finding, invasion and feeding. In: *The cyst nematodes.* (Sharma, S., ed.^eds.). Dordrecht: Kluwer Academic Publishers.

Waetzig, G. H., Sobczak, M. and Grundler, F. M. W. (1999) Localization of hydrogen peroxide during the defence response of Arabidopsis thaliana against the plant-parasitic nematode Heterodera glycines. *Nematology,* **1,** 681-686.

Wang, X., Allen, R., Ding, X., Goellner, M., Maier, T., de Boer, J. M.*, et al.* (2001) Signal peptide-selection of cDNA cloned directly from the esophageal gland cells of the soybean cyst nematode Heterodera glycines. *Mol Plant Microbe Interact,* **14,** 536-544.

Wang, X., Mitchum, M. G., Gao, B., Li, C., Diab, H., Baum, T. J.*, et al.* (2005) A parasitism gene from a plant-parasitic nematode with function similar to CLAVATA3/ESR (CLE) of Arabidopsis thaliana. *Molecular Plant Pathology,* **6,** 187-191.

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

Williamson, V. M. and Kumar, A. (2006) Nematode resistance in plants: the battle underground. *Trends in Genetics,* **22,** 396-403.

Wyss, U. (1992) Observations on the feeding behaviour of *Heterodera schachtii* throughout development, including events during moulting. *Fundam. appl. Nematol.,* **15,** 75-89.

Zhijian, C., Feng, L., Yingliang, W., Xin, M. and Wenxin, L. (2006) Genetic mechanisms of scorpion venom peptide diversification. *Toxicon,* **47,** 348-355.

Chapter 2

SECPEP: A novel gene family coding for hypervariable effector peptides from the potato cyst nematode *Globodera rostochiensis*

Nikkie van Bers, Hein Overmars, Ling Qin, Aska Goverse, Jaap Bakker and Geert Smant

to be submitted

Abstract

Secretory cyst nematode effectors are essential for succesfull feeding cell formation in plants. *SECPEP*1 is a novel effector gene, specifically expressed in the dorsal esophageal gland of infective juveniles from the potato cyst nematode *Globodera rostochiensis.* Here we show that *SECPEP*1 is a member of a pioneer gene family that consists of at least eight other members. The *SECPEP* genes code for positively charged secreted peptides, ranging in size from 3 to 12 kDa. The mature peptide sequence of the SECPEPs appears to be hypervariable, while, remarkably, the 3'UTR, the introns, and the region coding for the signal peptide for secretion are conserved. We show that some SECPEP sites are under diversifying selection and that the SECPEP peptides share their positive charge, C-X-G γ-core motif and hypervariability with several classes of effector peptides involved in host defense of animals and plants.

Introduction

Potato cyst nematodes (*Globodera rostochiensis* and *G. pallida*) are obligatory parasites of a small range of Solanaceous plants, including economically relevant crops such as potato, tomato and eggplant. The infective juvenile invades roots of a plant in the cell elongation zone and transforms a competent host cell, usually a procambium cell, into a feeding structure called syncytium (von Mende *et al.*, 1998). While feeding from its syncytium the nematode develops into the adult stage, mates, and reproduces. Females retain the fertilized eggs inside their body wall, which after death hardens to form a protective cyst.

The prominent adaptations to plant-parasitism in nematodes are the oral stylet and the enlarged esophageal glands. The lobe and extension of the esophageal glands are packed with secretory granules containing effector molecules. The oral stylet delivers the nematode effectors into the host cell apoplast and cytoplasm. Some of these effectors may be recognized by the plant immune system resulting in a potent defense response, while others may mask recognition of the nematode by the plant or actively suppress plant defenses. Furthermore, (a subset of) the effectors will promote virulence by facilitating the ingestion of plant material or by the reprogramming of host cells into feeding cells (Jones & Robertson, 1997, Hussey, 1989, Wyss, 1992).

The feeding site of cyst nematodes is the result of a progressive protoplast fusion of up to 200 host cells as a consequence of local cell wall degradation. A syncytium is a metabolically highly active cell complex, with a dense cytoplasm containing numerous organelles and small vacuoles. Goverse *et al* (1999) showed that a fraction of small proteins (<3kDa) present in the stylet secretions of *Globodera rostochiensis* co-stimulates cell proliferation of tobacco protoplasts. Co-stimulation is a term used for growth factors which solely do not induce cell proliferation, but enhance the proliferative response of cells upon mitogenic stimulation. The nematode protein fraction required the presence of the plant hormones auxin and cytokinin for their stimulatory effect. Interestingly, both auxin and cell cycle reactivation in host cells have

been shown to be essential for feeding site formation (de Almeida Engler *et al.*, 1999; Goverse *et al.*, 2000; Karczmarek *et al.*, 2004), and several cell cycle related genes like *cdc2a*, *cycB1;1* and *cycA2;1* are induced in the early stages of syncytium formation (Goverse *et al.*, 2000, Niebel *et al.*, 1996).

Recently, Ithal *et al* (2007) reported on 73 cyst nematode gland-expressed genes, which are differentially expressed during the nematodes' life-cycle (Ithal *et al.*, 2007). The functional role in parasitism of the majority of these genes remains to be elucidated. We have generated about 17,000 developmental expression profiles of *Globodera rostochiensis* genes in order to identify genes putatively involved in parasitism (Qin *et al.*, 2001), Qin, unpublished results). In this analysis, gene expression in five successive developmental stages was monitored by use of cDNA-AFLP. The transcript derived fragment (TDF) with code A4 was one of the approximately 216 TDFs upregulated at the onset of parasitism, which may point to a role as an effector molecule in parasitism. A4 corresponds to *SECPEP1*, which is specifically expressed in the dorsal esophageal gland of preparasitic juveniles of *G. rostochiensis*. *SECPEP1* codes for a small, secreted peptide of about 5 kDa. We analyzed *SECPEP1* further and found that it is a member of gene family of at least nine genes of which the mature peptide sequence appears to be hypervariable. The 3' UTR, the introns and the region coding for the signal peptide for secretion are conserved. We propose two models explaining the effector function of the *SECPEP*s in parasitism, based on the structural characteristics of the peptides, and the developmental expression profile of the gene family.

Material and methods

Nematode collection, RNA isolation and cloning of full length cDNA

Preparasitic juveniles of *G. rostochiensis* pathotype Ro1-Mierenbos were collected as described previously (De Boer *et al.*, 1992). Parasitic second stage juvenile (J2) nematodes, third stage juveniles (J3) and fourth stage juveniles (J4) were collected respectively 13/14, 19 and 23 days post-inoculation (dpi). Adult female and adult male nematodes were collected at respectively 34 and 27 days post-inoculation (dpi). Extraction of the nematodes from plant material was performed as described previously for adult males and females by (De Boer et al., 1992). Total RNA was isolated using TRIzol reagent (Invitrogen, Breda, the Netherlands). Five µg of total RNA was the starting material for full-length, RNA ligase-mediated rapid amplification of 5' and 3' cDNA ends (RLM-RACE) by use of the GeneRacer Kit (Invitrogen), which was performed according to the manufacturers' protocol. Super-script III reverse transcriptase (Invitrogen) was used for cDNA synthesis, according to the suppliers' protocol. Full-length cDNA was amplified by using gene specific primers in combination with adaptor specific primers, and subsequently cloned into the pCR4-TOPO vector (Invitrogen). Transformed *E. coli* colonies were checked for presence of the expected insert by PCR with the same primers as used for the amplification. Plasmids were purified using the Wizard *Plus* Miniprep DNA Purification System (Promega Corporation, Madison, WI, USA) and sequenced at the Sequence Facility Wageningen (Wageningen, the Netherlands) or at BaseClear (Leiden, the Netherlands).

Amplification of genomic DNA

Genomic DNA was extracted from preparasitic J2 as described by (Curran *et al.*, 1985). Genomic DNA fragments corresponding to the *SECPEP* genes were amplified with a gene specific forward primer annealing on the mature peptide region and a reverse primer annealing in the 3'UTR region. The *SECPEP1* 5' flanking region was amplified with the gene specific A4 5'RV primer (CAAAGCCACGCCCAAAATGAGAAG) in

combination with the pcDNA-FW primer (GGTGACACTATAGAATACTCAAGCTATGCA) from a plasmid preparation of a *G. rostochiensis* genomic library (Qin *et al.*, 1998). On the product of this PCR, a second PCR with the same primers was performed, which resulted in the amplification of 800 nucleotides of the *SECPEC1* 5'flanking region. The *SECPEP9* 5' flanking region was amplified from the same genomic library, in a PCR with the nested primers 3'UTR (271-293)RV (CAAAATGCCTTCAGCAAAATGAC) and 1h-(71-91)RV (GGCTGACTTTTAGTGGAAAGC) in combination with the pcDNA-FW primer (GGTGACACTATAGAATACTCAAGCTATGCA). For amplication of the start codon and SP region of *SECPEP2—7*, a gene specific reverse primer was used in combination with a forward primer designed on a stretch of nucleotides conserved in the promoter regions of *SECPEP1* and *SECPEP9* (1/1h(707-730)FW: 5' TTAAGGGCTGAAATTGGCAAATAT 3'). Amplification products were cloned into the pCR4-TOPO vector (Invitrogen).

Semi-quantitive RT-PCR analysis

Transcript expression of the *SECPEP* genes was analysed in different parasitic stages by reverse transcription (RT)-PCR. Total RNA was extracted using TRIzol Reagent (Invitrogen). For cDNA synthesis, total RNA was treated with Turbo DNA-free (Ambion) to degrade contaminating genomic DNA. Superscript III reverse transcriptase was used for first-strand cDNA synthesis, essentially to the manufacturers' protocol (Invitrogen). Gene-specific and control primers were designed with the Beacon designer 4.02 program (PREMIER Biosoft International; Table 1). As an internal control for template amount and quality, we amplified cDNA derived from a cAMP-dependent protein kinase catalytic subunit encoding gene. Primers were designed to specifically anneal on the *G. rostochiensis* EST BM343563, which shows significant sequence similarity to *Ancylostoma caninum* cAMP-dependent protein kinase catalytic subunit, genbank accession number U15983 (Trivedi & Arasu, 2005).

Table 1: Primers used in RT-PCR and for *in-situ* probes

Amplicon	Exp.	Primer name		Sequence (5'→3')
*SECPEP*1	RT-PCR	Dgl-1-(110-134)FW		AAAGAAAATAAATCGAATTCAGGTG
		Dgl-1-(165-184)RV		TAGGTTTGTTCCTCCAGCAG
*SECPEP*2	RT-PCR	Dgl-1a-(134-153)FW		AAACTTTTGATGCCCACTGG
		Dgl-1a-(210-234)RV		TTAGTTGATTGATTGTTTGACTTTC
	In-situ	Sense-probe	411 (76-99) FW	GGACCCGAAAGCCCAAATGGTTAT
		Anti-sense probe	BM345411 3'RV	GTCGCCATCTTTAGCCTCCACG
*SECPEP*3	RT-PCR	Dgl-1b-(125-149)FW		TACAGACAGGAAAACATAAAGAACC
		Dgl-1b-(196-218)RV		GTCCCACTACCCTTAAAGAATCC
	In-situ	Sense-probe	BM344094 5'FW	GTGTTCCGGCTGCGGTGACGGTG
		Anti-sense probe	1.2.2 (170-194) RV	CAATGCCTTCAACATAATGATGTCA
*SECPEP*4	RT-PCR	Dgl-1c-(214-235)FW		ATTGGAATGATTGCGATGAAGC
		Dgl-1c-(318-342)RV		ACCATACATAAAAGGCACAATAAGG
	In-situ	Sense probe	BM344094 5'FW	GTGTTCCGGCTGCGGTGACGGTG
		Anti-sense probe	094 (196-219) RV	GAAAATCAAAGCACCCAAAAATGC
*SECPEP*5	RT-PCR	Dgl-1d-(185-208)FW		CCAGAACTCAATTTATTACGATGC
		Dgl-1d-(330-354)RV		CAAATTTAACGGCATTGAAAGAATC
	In-situ	Sense probe	2.1.2 (34-56) FW	GGTGGTTGTATTGGGAAAGACTC
		Anti-sense probe	2.1.2 (379-401) RV	CCTCGTTCAGCAGAATGGCGTCA
*SECPEP*6	RT-PCR	Dgl-1e-(107-126)FW		GAAAGAAAAGCCTCCACCGA
		Dgl-1e-(164-181)RV		GGTTAGTTGGGGCACCTA
	In-situ	Sense probe	(88-100)FW	AGAAAGAAAAGCCTCCACCGAG
		Anti-sense probe	3'UTR (271-293) RV	CAAAATGCCTTCAGCAAAATGAC
*SECPEP*7	RT-PCR	Dgl-1f-(92-111)FW		TGGCGGTGATGGAAAGAAAG
		Dgl-1f-(144-163)RV		GGGTCGTGTCTTGATCTCTC
	In-situ	Same as *SECPEP6* (different template)		
*SECPEP*8	RT-PCR	Dgl-1g-(29-51)FW		GTAATTGTTTTGGAAAGAAAGCA
		Dgl-1g-(91-110)RV		GCCTCTCTGATGTATGGTTT
*SECPEP*9	RT-PCR	Dgl-1h-(46-65)FW		GTGCGACCGACTGAATAAAA
		Dgl-1h(95-116)RV		TTTTCAGTGACCCCTCTATTTT
cAMP dep. protein kinase cat. subun.	RT-PCR	Kin-1-F		ATCAGCCCATTCAAATCTACG
		Kin-1-R		TTCTTCAGCAAGTCCTTCAAC
ATPase transportin g protein	RT-PCR	Rt-ATP-F		GATATTGACAGTTTCTGTGAG
		Rt-ATP-R		ATCAATTCACATTCATCTAAA

In situ hybridisation microscopy

In situ hybridisation microscopy was done on preparasitic juveniles (J2). The nematode fixation, hybridization and detection steps were essentially as described by (de Boer *et al.*, 1998), except for the use of single stranded DNA probes, which were synthesized by linear PCR as described by (Patel & Goodman, 1992). The primers used to generate the probes are listed in Table 1, plasmid containing cDNA was used as template in the PCR.

Antibody production and immunolocalization

Protein size is an important parameter determining immunogenicity (Abbas & Lichtman, 2003). That is why, for antibody production, SECPEP1 was produced as a fusion with glutathione-S-transferase (GST). Primers A4pGEXF (5' agctggatcctgcggtggcggtgatgga 3') and A4pGEXR (5'tttgcgtggccttcttgt 3') were used to amplify the coding region of the mature SECPEP1 peptide from *G. rostochiensis* cDNA. The amplified fragment was first made blunt ended, and subsequently it was digested with BamHI (restriction site underlined in primer). The EcoRI and BamHI sites in the pGEX2T vector (GE Healthcare) were used for cloning. The *SECPEP1* coding sequence includes an EcoRI site, which makes that plasamid digested with this enzyme could not be used directly for cloning. After digestion with EcoRI, the plasmid was made blunt ended. Subsequently, the vector was digested with BamHI after which it was ready for ligation of the PCR product. Recombinant protein was produced by the *E. coli* BL21 strain, and purified with the GST purification module (GE Healthcare). Hens were immunized with purified recombinant protein, and the resulting chicken IgY antiserum was isolated from eggs (as described by (Polson *et al.*, 1980) and further purified (Harlowe & Lane, 1997). Immunofluorescence was performed on pre-parasitic J2 as described by (de Boer *et al.*, 1996).

Sequence analysis

DNA and amino-acid sequences were analysed using the Informax Vector NTI Advance software (Invitrogen). The computer algorithm SignalP 3.0 (Bendtsen *et al.*, 2004) was used to predict the presence of a signal peptide for secretion and the corresponding putative cleavage site (Neural Networks). Blast at NCBI (www.ncbi.nlm.nih.gov) or at nemaBLAST (www.nematode.net) was used to search for matching sequences in the database.

ProtScale and ProtParam (ExPASY web server) were used for the generation of respectively the hydropathy plots (window size 9) and the calculation of the Grand Average of Hydropathy (GRAVY) score. The sequences of the antimicrobial peptides HNP-1, LL37 and VIP were as described in (Boman, 2003). The contoxin GVIA (accession number gi 461869) was described previously by (Olivera *et al.*, 1985).

Sequence alignments were generated in the ClustalW 1.83, using default settings (Thompson *et al.*, 1994). Subsequently, Jalview (v. 2.3(Clamp *et al.*, 2004) was used for manual editing of the sequence alignment. A consensus sequence of the *SECPEP* genomic DNA sequences was generated in AlignX (VNTI software), after which gaps were manually removed. The consensus sequence was used as the local subject database for a BLAST search in Bioedit (Isis Pharacauticals, CA) with the *SECPEP*1 blocks as input sequences. This BLAST assigns all the blocks with a bit-score (S'), which is calculated from the raw score S (Altschul *et al.*, 1990, Karlin & Altschul, 1990). The bit score S' correlates with the length of the input sequence. Because the *SECPEP* blocks differ in length, we standarized S' by dividing by the query length l (in bp), so that $S^*=S'/l$. The corrected bit-score S^* can range between 0, assigned to sequences for which no similarity is found in the database, to 2.0, which is assigned to sequences that exactly match a sequence in the database.

PAML analysis

Seven sufficiently matching *SECPEP* sequences were tested for positive selection;

*SECPEP*1—3 and *SECPEP*5—8. The ratio ω was estimated with the codeml program of PAML (phylogenetic analysis by maximum likelihood) (Yang, 1997, Yang & Bielawski, 2000). Two models of fitting codon substitution were used to calculate likelihood ratio statistics (LR), twice the log-likelihood between models is compared with the value of a χ^2 distribution with branches-1 degrees of freedom. Model M7 (β distributed variable selection pressure) has a ω for each site drawn from a β distribution with parameters p and q. Model M8 (β plus ω > 1) uses the M7 recipe for a fraction p_0 of the sites and assigns another ω to the remaining fraction. M7 and M8 are nested models, so they can be compared using a likelihood ratio test (LRT) which is generally robust to the assumed distribution of ω over sites. When M8 fits the data significantly better than M7 and the ω ratio estimated under model M8 is greater than 1, we assumed evidence of positive selection. To check whether it is significantly greater than 1 the log-likelihood value in M8 is recalculated while fixing ω to be 1 (model M8A from ((Wong *et al.*, 2004) and compared to the change in likelihood with a χ^2 distribution with 1 degree of freedom. Likewise the less complicated models M0 (uniform selective pressure among sites) with M3 (variable selective pressure among sites) were calculated and the results were found to give less conservative estimates than M7/M8.

Next, positive selection was tested for by studying variation among sites identifying amino acids under diversifying selection. This variation is tested with an additional LR test between M7 and M8 (Yang & Nielsen, 2000) using the empirical Bayes theorem as implemented in PAML to calculate the posterior probability that a particular amino acid belongs to a particular class (neutral, negative or positive). A particular site that belongs to the class ω>1 with a posterior probability > 95% is most likely under positive selection. This approach makes it possible to detect positive selection and identify sites under positive selection even if the average ω ratio over all sites is less than 1 (Yang, 2006). Meanwhile, for this type of study it is important to note three test characteristics. First, detection of positive selection requires significant differences between M7 and M8 and estimates of ratio that exceed 1. Second, under

M8 it is possible to estimate the proportion of sites that are under positive selection, and this proportion is denoted P1. Third, the application of these models requires a topological, or phylogenetic, assumption. For each sequence group, PAML analyses were applied using the M0 generated phylogenetic tree. The aminoacid sequence alignment was executed by ClustalX (v1.83) (Chenna *et al.*, 2003). Pal2nal (v11) was used to relate the sequences back to a nucleotide alignment by converting a multiple sequence alignment of proteins and the corresponding DNA (or mRNA) sequences into a codon-based DNA (nucleotide) alignment. The results, including gaps, were directly used in PAML.

Phylogenetic analysis

The cDNA sequences of *SECPEP*s showing >60% sequence identity to at least one of the other family members were aligned using the ClustalW algorithm (v1.83, default settings). The alignment was manually adjusted in Jalview. Treepuzzle 5.2 and Treeview (v. 1.6.6) were used to generate an unrooted phylogenetic tree, in which nodes with a posterior probablilty lower than 0.95 are considered to be unresolved.

Results

***SECPEP*1 encodes a 5 kDa peptide that localizes specifically to the lobe and the extension of the dorsal esophageal gland of pre-parasitic juveniles**

Previously, we showed that the *SECPEP*1 gene is specifically expressed in the dorsal esophageal gland of preparasitic second stage juveniles of *Globodera rostochiensis* (Qin *et al.*, 2000). This gland is believed to be an important source of secretory molecules involved in feeding site induction and maintenance in the host (Wyss, 1992). *SECPEP*1 codes for a 5 kDa peptide of which the first 18 N-terminal amino acids are predicted to act as a signal peptide for secretion.

To further investigate whether the SECPEP1 protein is indeed secreted, we studied its subcellular localisation within the esophageal cells by use of immuno-

localisation in pre-parasitic J2 nematodes. Antiserum raised to SECPEP1 genetically fused to GST (α-SECPEP1-GST) reacted specifically with secretory granules in the lobe and in the extension of the dorsal esophageal gland of pre-parasitic juveniles (Fig. 1). An antiserum against GST alone did not result in specific labeling of tissues in the nematode (data not shown). Next, we tried to conclusively show the presence of SECPEP1 in concentrated collected stylet secretions of pre-parasitic J2 in a dot-blot assay. Blots of collected stylet secretions incubated with the α-SECPEP1-GST antiserum resulted in a signal (data not shown). However, the antiserum against bacterially produced GST alone also gave a, albeit weaker, signal on the dot-blot.

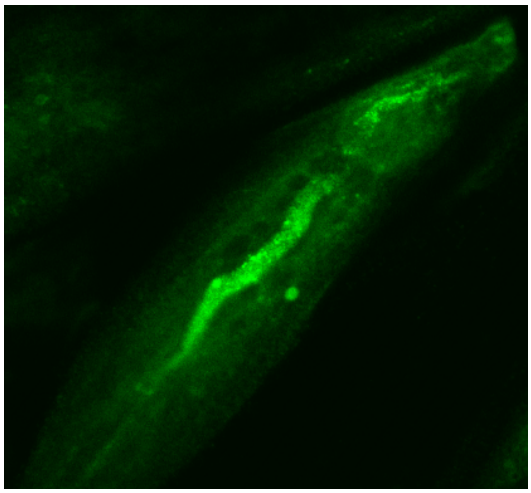


Figure 1: Immuno-labeling of pre-parasitic J2 juveniles with antiserum against SECPEP1-GST. Fluorescence is observed in granular structures in the lobe and extension of the dorsal esophageal gland.

SECPEP1 shows sequence similarity to at least 8 other genes

A BLAST-X analysis with the translated *SECPEP1* cDNA sequence on the EST database (NCBI) resulted in the finding of ten ESTs from *Globodera rostochiensis* with varying degrees of sequence similarity. Four ESTs originated from a cDNA library generated from preparasitic juveniles of a population of *G. rostochiensis* pathotype Ro1 Mierenbos, which is a mixed population of virulent and avirulent nematodes on the nematode resistance gene *H1* (reviewed by (Bakker *et al.*, 1993). These four matching

ESTs are BM345965 (E-value of BLAST result is 2e-74), BM345411 (E-value of 4e-18), BM345465 (E-value is 0.013), and BM344094 (E-value is 0.026). BM345965 differs at only 1 nucleotide position from *SECPEP1*. We found this single-nucleotide polymorphism in several independent experiments making it unlikely that it reflects an artifact and, therefore, BM345965 may represent an allelic variant or a recently duplicated copy of *SECPEP1*. BM345411 contains a 207 basepair (bp) open reading frame and represents a novel *G. rostochiensis* gene, which is named *SECPEP2*. The gene corresponding to BM344094 is named *SECPEP4*. Several attempts to re-amplify cDNA corresponding to BM345465 failed. The six other significantly matching ESTs represent 3 clones from a cDNA library generated from parasitic J2 of the single-female line 19 of *G. rostochiensis*, which was selected for avirulence on the nematode resistance gene *H1* (Janssen *et al.*, 1990). These library clones are GRAA-aaa40e10 (*SECPEP3*, E-value is 6e-09), GRAA-aaa50g05 (*SECPEP7*, E-value is 1e-6) and GRAA-aaa67a01 (*SECPEP10*, E-value is 1e-20). None of the sequences shows significant sequence similarity to other accessions in the current sequence databases.

ESTs may include only part of the corresponding transcript. To test for more sequence information up and down-stream of the EST, we used RACE for the amplification of the full-length transcripts of *SECPEP2*, *SECPEP3*, *SECPEP4* and *SECPEP7*. Interestingly, sequences representing a further four novel *SECPEPs* were obtained when we sequenced a small library of PCR amplified *SECPEP1* (chapter 5 of this thesis). These novel *SECPEPs* were named *SECPEP5*, *SECPEP6*, *SECPEP8* and *SECPEP9* (Table 1).

Nearly all of the *SECPEP* transcripts start with a 22-23 nucleotide long spliced leader (SL), which is spliced onto the 5'-end of mRNA in a process called *trans*-splicing (Blumenthal, 1995). Only for *SECPEP3* and *SECPEP7* we consistently find the same leader sequence for each transcript (after sequencing of at least 5 clones). This leader sequence, SL1, complies with the canonical leader sequence found on 50% of transcripts in *Caenorhabditis elegans* (Zorio *et al.*, 1994). For the other *SECPEPs*

transcripts with either a SL1 or novel spliced leader sequences were obtained (chapter 5 of this thesis).

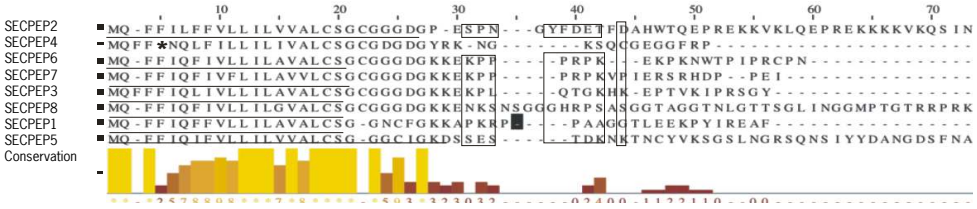


Figure 2: Sequence alignment of the SECPEP peptides. The signal peptide for secretion, as predicted by the Neural Networks algoritm in SignalP, is underlined. Residues predicted to be under positive selection are boxed.

The *SECPEP* genes code for peptides with structural similarity to host defense effector polypeptides

SECPEP1 encodes a small protein that is likely secreted from the dorsal esophageal gland cell. Sequence analysis showed that most of the *SECPEP* genes code for putatively secreted peptides ranging in molecular weight from 3.2 to 12.4 kDa (Table 1). Strikingly, the sequence of the signal peptides (SP) is conserved between the *SECPEPs*, while the predicted cleavage site of the signal peptides is not (Fig. 2). In case of *SECPEP4* and *SECPEP9* a stop-codon at respectively the 4th and 1st codon position after the start-codon prematurely terminates translation, for which we classified these genes as pseudogenes.

Hydropathy plots are used to visualize hydrophobic and hydrophilic regions in a protein sequence. In these plots, an average score, numerically expressing the overall level of hydrophilicity, can be depicted. This score is called GRAVY (Grand Average of hydropathy) score. A mean GRAVY score of -0.4 kcal/mol was reported for soluble (hydrophilic) proteins and values exceeding -0.4 kcal/mol were reported for membrane-spanning (hydrophobic) proteins (Kyte & Doolittle, 1982).

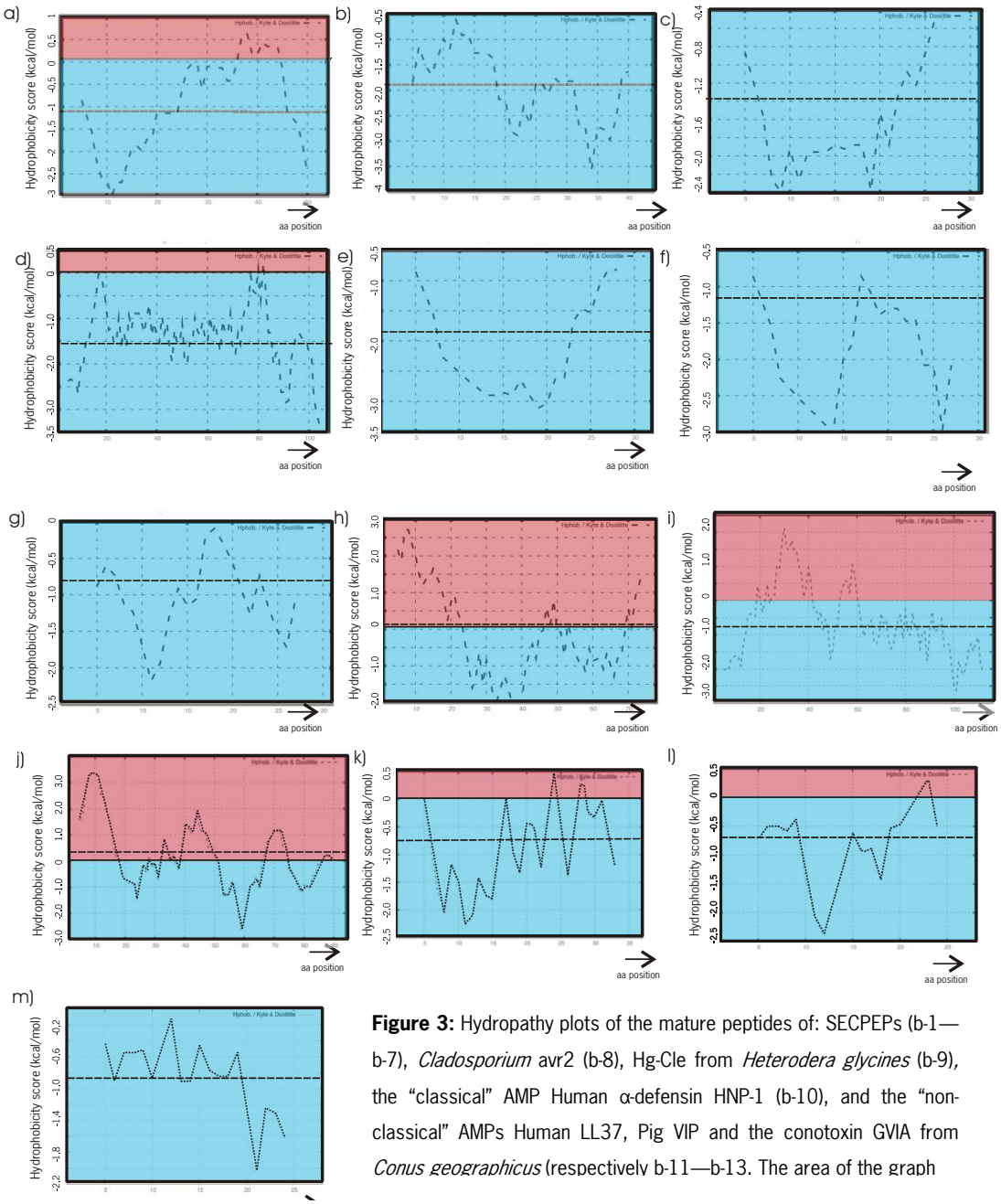


Figure 3: Hydropathy plots of the mature peptides of: SECPEPs (b-1—b-7), *Cladosporium* avr2 (b-8), Hg-Cle from *Heterodera glycines* (b-9), the "classical" AMP Human α-defensin HNP-1 (b-10), and the "non-classical" AMPs Human LL37, Pig VIP and the conotoxin GVIA from *Conus geographicus* (respectively b-11—b-13. The area of the graph corresponding to positive hydrophobicity score values is shaded relatively darker than the area corresponding to a negative hydrophobicity score. The position of the GRAVY score is indicated by a dotted line and the position of the 0 kcal/mol value is indicated by –. The plots are generated by the Kyte Doolit tle algoritm in Protscale ({Kyte, 1982 #1602}).

We generated hydropathy plots for the SECPEPs, and compared these to hydropathy plots of several other peptides of the same size range involved in interspecies interactions. The *Cladosporium fulvum* avirulence effector protein avr2, the beet cyst nematode effector Hg-Cle and the mammalian antimicrobial peptide HNP1 are partly hydrophilic and partly hydrophobic (Fig. 3h-j), while the mammalian antimicrobial peptides LL37 and VIP and the omega conotoxin peptide GVIA from the fish hunting cone snail *Conus geographicus* are predominantly hydrophilic, which is illustrated by the negative values in the hydrophobicity plots and negative GRAVY scores (Fig. 3k-m), and (Boman, 2003).

We found that the mature SECPEP peptides are also hydrophilic (Fig. 3a-g). Several classes of host defense peptides, *e.g.* antimicrobial peptides and conotoxins, are characterized by a net positive charge, and cysteine containing host defence polypeptides contain a C-X-G or G-X-C γ-core structural motif (reviewed in Yount, 2004 #695} and (Yeaman & Yount, 2007). Interestingly, also all the SECPEP peptides are positively charged, with a net charge ranging from +1 to +6 at pH 7 (Table 1) and contain a C-X-G γ-core motif near their N-terminus (Fig. 2).

Table 1: Characteristics of the SECPEPs

Gene	Trans-cript size (bp)	Size of mature peptide	Net charge at pH7
SECPEP1	389	5.0	+6
SECPEP2	397	5.2	+1
SECPEP3	352	3.2	+4
SECPEP4	378	-	-
SECPEP5	573	12.4	+6
SECPEP6	350	3.5	+5
SECPEP7	346	3.3	+2
SECPEP8	376	3.3	+3
SECPEP9	309	-	-

The *SECPEP* genes are specifically expressed in the dorsal esophageal gland

To investigate the site of expression of the *SECPEP*s in nematodes we conducted whole mount *in-situ* hybridization microscopy on infective J2s. Figures 4a-f show that transcripts corresponding to *SECPEP2* to *7* specifically localize to the dorsal esophageal gland. Except for *SECPEP5*, all the *SECPEP* transcripts were detected in the cytoplasm around the nucleus of the gland cell. *SECPEP5* transcripts localize at the site of the nucleus of the dorsal esophageal gland (Fig. 4d). Furthermore, for *SECPEP5* two different variants of transcripts have been detected; the full length transcript and an alternatively spliced variant in which exon 4 is lacking (*SECPEP5*-exon4). We investigated whether a probe specific for *SECPEP5*-exon4 shows a different hybridisation pattern as was observed for *SECPEP5*. We found that the *SECPEP5*-exon4 probe also hybridizes to the nuclear region of the dorsal esophageal gland (data not shown). For each gene, a probe corresponding to the sense strand was included to asses for non-specific hybridisation to gDNA. No hybridisation was observed for any of the sense probes, confirming that all the probes specifically hybridized to mRNA.

In addition, we further investigated the subcellular localisation of the SECPEP3, SECPEP4 and SECPEP5 peptides with immuno-localisation in pre-parasitic J2. However, none of the antisera raised to these proteins resulted in specific labeling in the nematodes.

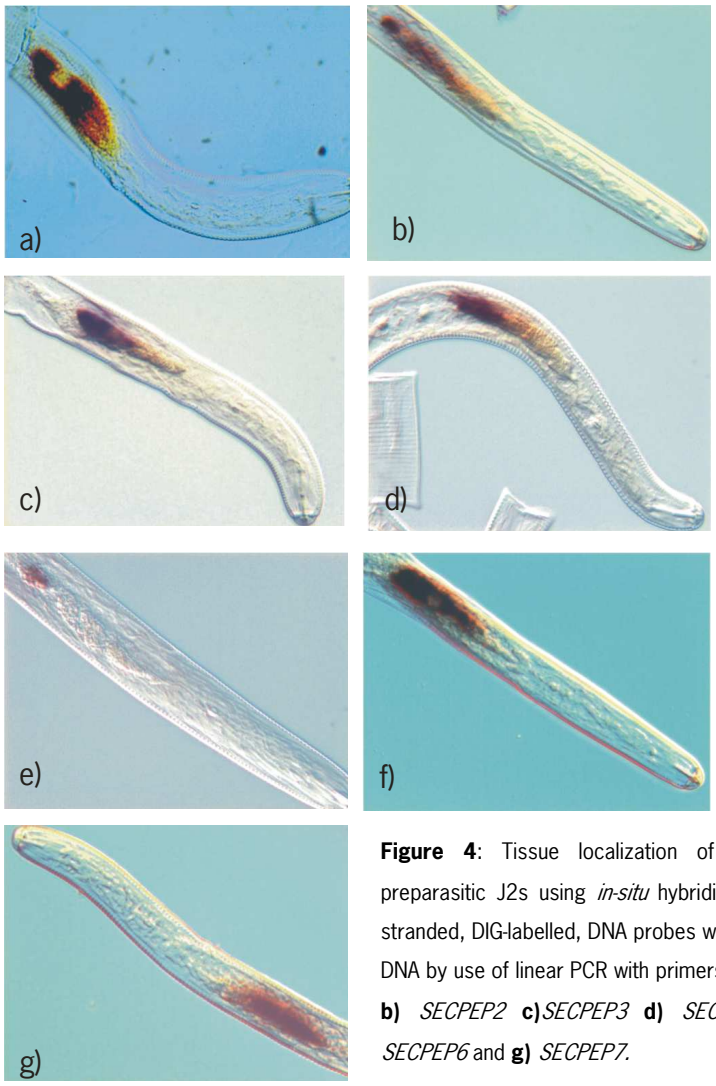


Figure 4: Tissue localization of *SECPEP* transcripts in preparasitic J2s using *in-situ* hybridization microscopy. Single stranded, DIG-labelled, DNA probes were amplified from plasmid DNA by use of linear PCR with primers specific for: **a)** *SECPEP1* **b)** *SECPEP2* **c)** *SECPEP3* **d)** *SECPEP4* **e)** *SECPEP5* **f)** *SECPEP6* and **g)** *SECPEP7.*

The *SECPEP* genes differ in expression throughout development

It is believed that there are distinct stages in parasitism of nematodes in plants (viz. migration, feeding, and reproduction) each requiring a different set of effector molecules. We investigated if expression of the *SECPEP*s is correlated with specific events in the parasitic cycle. Reverse Transcription (RT)-PCR showed a pronounced

band of the expected size (74bp) for *SECPEP1* in pre-parasitic and parasitic J2 and a fainter band of the same size is observed in later parasitic stages and in the adult female (Fig. 5). *SECPEP1* expression could not be detected in adult males (Fig. 5). *SECPEP4* follows the same expression pattern as *SECPEP1*, while the other *SECPEP* genes showed a surprising variety in expression profiles. A band of the size expected for *SECPEP2* (185 bp) and *SECPEP5* (169 bp) were predominantly found in pre-parasitic juveniles, while expression of *SECPEP3* (93 bp), *SECPEP6* (75 bp), *SECPEP7* (71 bp) and *SECPEP8* (81 bp) are mainly found in parasitic J2.

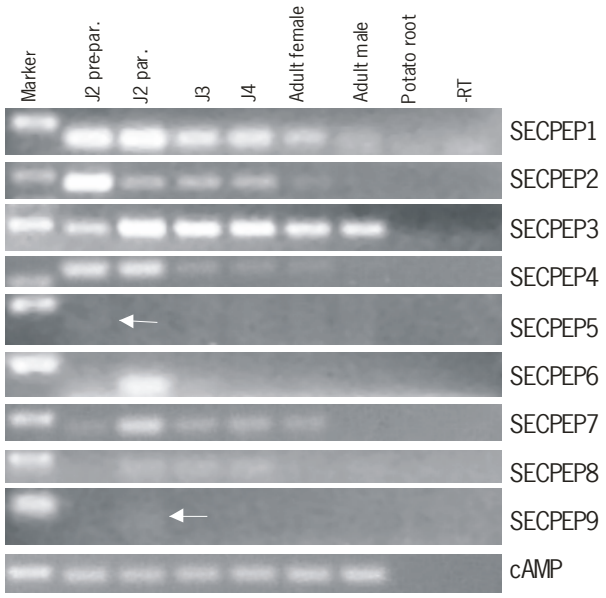


Figure 5: Developmental expression patterns as determined by semi-quantitative RT-PCR of *SECPEP* cDNAs from different life stages of *G. rostochiensis.* The samples were analysed on gel after 30 PCR cycles. J2 parasitic juveniles are collected 13-14 days post inoculation (dpi) of plants with cysts. J3 parasitic juveniles are collected at 19 dpi, J4 at 23 dpi, adult females at 34 dpi and adult males at 27 dpi.

Potato root cDNA was included to confirm primer specifity and the sample in which reverse transcriptase was omitted (-RT) allows investigation of the presence of contaminating genomic DNA. cAMP dependent protein kinase catalytic subunit cDNA (cAMP)(EST accession nr. BM343563) was included as an internal control for equal template conditions. All the bands observed are of the expected sizes. Arrows indicate the position of the faint band observed for *SECPEP5* and *SECPEP9*.

The constitutively expressed cAMP-dependent protein kinase catalytic subunit encoding gene (Trivedi & Arasu, 2005) and ATP-ase transporting protein (data not shown) were used as reference genes for constitutive expression showing that an equal amount of template was used in all the samples (Fig. 5). For all the primer sets a control PCR reaction was performed on a pre-parasitic J2 sample in which reverse transcriptase was omitted (Fig. 5). No product was detected in these samples for any of the primer combinations, which indicates the absence of genomic DNA contamination. The specificity of the primer sets was further confirmed in a reaction including potato root cDNA.

The *SECPEP* genes consist of blocks showing different levels of sequence divergence

A protein sequence alignment of the *SECPEP*s shows considerable sequence divergence in the mature peptide, as compared to the SP domain (Fig. 2). This led us to examine whether the observed divergence extends into the non-coding regions of the genes. In addition, we compared the genomic organisation of the *SECPEP*s to get more insight into their ancestry and mode of evolution. The genomic sequence of *SECPEP1—8* was amplified from the first nucleotide of the start codon to the poly-A adenylation site. The amplified sequences were divided into blocks including the exons, with exon 1 divided into a SP region and a remaining part (SP-i1), and the introns (Fig. 6). To investigate sequence divergence among the distinct blocks, we calculated the divergence relative to the consensus sequence derived from the alignment of the *SECPEP*s. *SECPEP5* is more than three times larger than the average *SECPEP* sequence. In an alignment, this size difference results in extensive gap formation, and therefore *SECPEP5* was excluded from the alignment used for the assembly of the consensus sequence. The consensus sequence is used as a local subject database for a BLAST search with the *SECPEP* blocks. This results in corrected bit-score values (S*) that can range between 0, assigned to sequences for which no similarity is found in the

database, to 2.0, which is assigned to sequences that exactly match a sequence in the database.

The SP-encoding region (marked by an asterisk) is highly conserved between the family members, with S* values higher than 1.5 for most SPs. This finding is striking, as the SP encoding region is often a relatively less conserved part of a gene, because its functionality depends on hydrophobicity rather than on sequence ((Martoglio & Dobberstein, 1998)). *SECPEP4* is a pseudogene, which is not expected to be translated, so no SP could be assigned to *SECPEP4*. The region from the end of the signal peptide until the start of intron 1 (SP-i1) shows in most *SECPEP*s a lower level of sequence conservation than is observed for the SP region. For SP-i1 of *SECPEP2* and *SECPEP5* no similarity was found in the BLAST search (S*=0), and SP-i1 of *SECPEP8* and *SECPEP1* have S* values of 0.46 and 0.78, respectively. On the contrary, SP-i1 of *SECPEP3*, *SECPEP6* and *SECPEP7* do show sequence conservation, as indicated by S* values of 1.4-1.54. The latter is also reflected in the phylogenetic tree of the *SECPEP* family which reveals a more recent divergence of *SECPEP3*, *SECPEP6*, *SECPEP7* and *SECPEP2* (Fig. 7).

Most of the *SECPEP* genes contain two introns, with borders that all comply with the "GU-AG"-rule (Blumenthal, #268). The position of the two introns is conserved between all the *SECPEP* genes with intron 1 (i1) being located within the mature peptide encoding region, and intron 2 (i2) locating (in most cases) in the 3'UTR (Fig. 6). For *SECPEP1, SECPEP6*, *SECPEP7*, *SECPEP8* and to a lesser extent *SECPEP4* are also the sequence and the size (71-74 nt) of intron 1 conserved. Also intron 2 of *SECPEP6*, *SECPEP7* and *SECPEP8* shows some sequence conservation (S^*=0.59-0.83). This sequence conservation is, however, totally absent in exon 2 for which in most *SECPEP*s no similarity to the consensus sequence was found (S^*=0). Exon 3 (largely corresponding to the 3'UTR) is relatively conserved between most of the family members (S*=0.5-1.4).

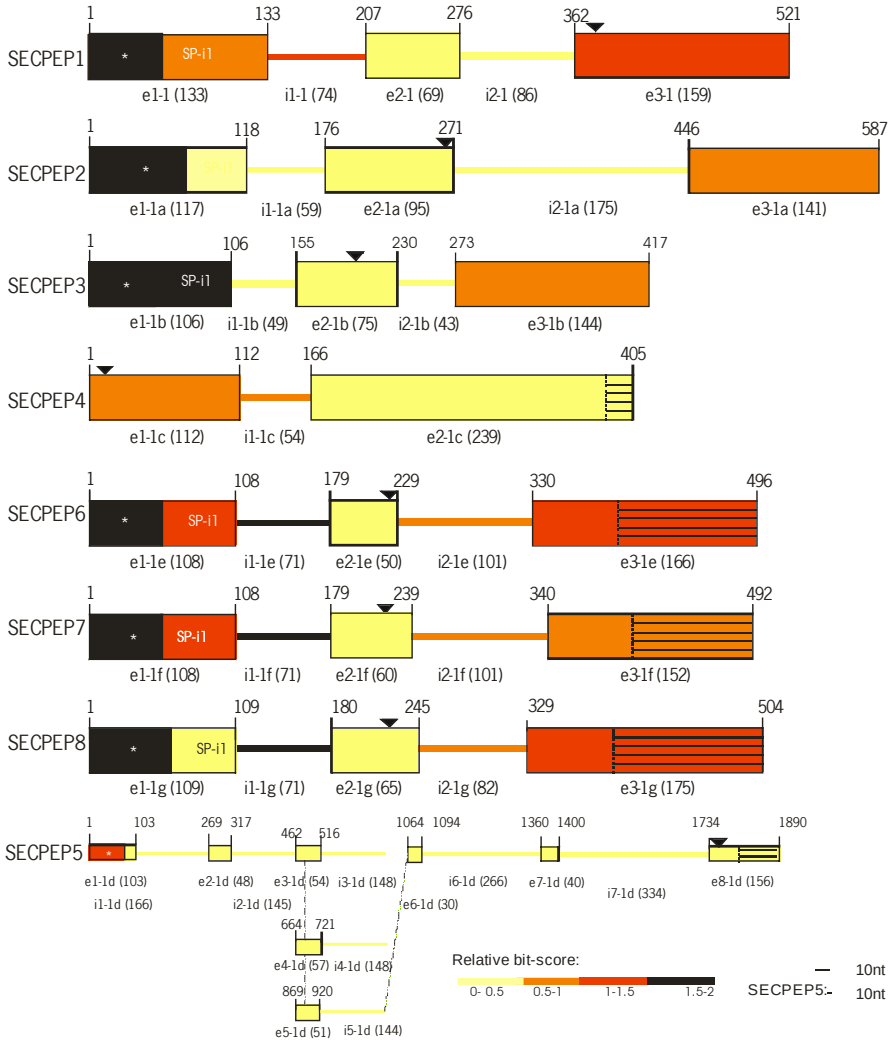


Figure 6: Exon/intron organisation of the *SECPEP* family. Sequences are shown from the start-codon (A of ATG is numbered 1) until the poly-A tail addition site in the mRNA. Exons (boxes) and introns (lines) are drawn proportionally to size. Sizes (nt) are depicted between brackets. The greyscale corresponds to the relative bit-score derived from a BLAST bit-score. The signal peptide region is marked with *, the position of the stop-codon is marked by a triangle, and shaded areas have only been amplified from cDNA. E3-1d/i3-1d, e4-1d/i4-1d and e5-1d/i5-1d are the result of a duplication event, and are positioned under eachother.

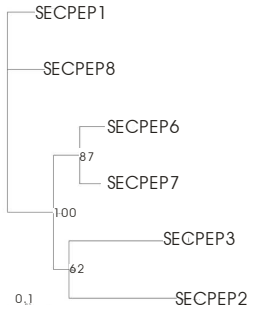


Figure 7: The cDNA sequences of *SECPEP*s showing >60% sequence identity to at least one of the other family members were aligned and were used to generate an unrooted phylogenetic tree, in which nodes with a posterior probablilty lower than 0.95 are considered to be unresolved.

Residues in the SECPEP peptides are undergoing positive diversifying selection

The ratio (ω) between non-synonymous (d_n) and synonymous changes (d_s) is a measure for the selection pressure acting on a codon ($\omega = d_n / d_s$). Codons with $\omega<1$ are under negative selection, while $\omega=1$ is found for codons under neutral selective pressure. Codons under positive selection pressure are characterized by $\omega>1$ (Li & Graur, 1991, Yang, 2006). Exon 1 of the *SECPEPs* shows sequence conservation for the signal peptide encoding region while diversification is observed for the remaining part of exon 1 which encodes (a part of) the mature peptide (Fig. 6). We used the PAML algorithm to investigate whether positive selection acts on sites in exon 1. Most commonly, sequence data from sibling species are used to detect positive selection. However, in the absence of these data, recent paralogous gene duplicates can be analysed for evidence of positive selection (Thomas, 2006, Thomas *et al.*, 2005, Win *et al.*, 2007). Our analysis of *SECPEP1—3* and *SECPEP5—8* indicates that multiple sites in exon 1 (residue 31-33, and residue 39-42) are subject to positive selection ($p<0.01$)(Fig. 2). PAML analysis was not conclusive for exon 2, as the diversity in exon 2 results in an alignment in extensive gap formation, to which no ω value is assigned.

SECPEP5 consists of duplicated introns and exons

SECPEP5 is more than three times larger in size than most of the other *SECPEP* genes.

It consists of (at least) 8 exons, separated by 7 large introns (average size 193 nucleotides). This intron size is in contrast to what has been observed for the introns of the other *SECPEP* genes and for *C. elegans* genes, where the size of most introns ranges from 45-54 nucleotides (Blumenthal & Thomas, 1988; Blumenthal & Steward, 1997). Exon3 (e3), exon4 (e4) and exon5 (e5) of *SECPEP5* show a striking level of sequence conservation. Exons 4 and 5 are identical, apart from a stretch of 6 additional nucleotides at the 3' end of exon 4. These 6 nucleotides are also present in exon 3 of *SECPEP5*. Exon 3 of *SECPEP5* is 81% identical to exon 4 and 72% identical to exon 5. Differences are mainly located at the 5' end of the exons, and to a stretch of 3 nucleotides found in the central part of exon 4 and exon 5 but which is absent from exon 3. The similarity observed in exon 3—5 is reflected in the neighboring intron 3, intron 4 and intron 5. Intron 4 and intron 5 are 100% identical, and intron 3 is 77% identical to intron 4 and intron 5. These regions may have originated from an (recent) intron and exon duplication event (Fig. 6).

Discussion

In this chapter, we report on the molecular characterisation of a family of novel, hypervariable, effector genes from the potato cyst nematode *Globodera rostochiensis*. *SECPEP1* was first identified by cDNA-AFLP as the transcript derived fragment A4 which expression was found to be up-regulated during hatching of infective juveniles of *Globodera rostochiensis* (Qin et al., 2000). Here we show that *G. rostochiensis* has at least eight genes coding for small proteins with sequence similarity to *SECPEP1*. Although their primary sequence similarity is limited to the SP and non-coding regions, we consider the *SECPEP*s as a gene family based on similarities in structure and expression.

The *SECPEP* genes all code for highly divergent dorsal esophageal gland specific secretory peptides

No sequence similarity was found for the *SECPEP*s with ESTs from other plant parasitic nematodes, including other related cyst nematodes. Our data suggest that the SECPEP gene family exhibits hypervariability even within *G. rostochiensis*: a feature that renders them hard to find in other nematode species. The lack of homologs in other species in the databases may also partly be attributed to the small sizes of the transcripts. These short transcripts may have been removed in the size fractionation step during the construction of diverse cDNA libraries. Furthermore, most of the more recently generated ESTs from plant parasitic nematodes are constructed with the SMART technology. A bias towards reverse transcription of transcripts that are not *trans*-spliced seems to exist in libraries generated with this technology (Vanholme *et al.*, 2006)(van Bers *et al.,* chapter 5 of this thesis). Amplification of the 5'-end of the *SECPEP* transcripts shows that all the *SECPEP* genes undergo *trans*-splicing. However, the ESTs of only one out of the seven *SECPEP* transcripts in the Genbank database includes a partial SL-sequence, which supports the idea that a bias against *trans*-spliced transcripts exists.

In situ hybridisation microscopy with gene specific probes shows that all *SECPEP*s investigated are specifically expressed in the dorsal esophageal gland of preparasitic juveniles of *G. rostochiensis*. All antisense probes labeled the cytoplasm of the dorsal esophageal gland, except for the *SECPEP5* probe, which reacts with the nuclear region of this gland. Specific localization of a cyst or root knot nematode transcript in or around the nuclear region of the dorsal gland has also been reported for transcripts of the family of SPRYSEC genes in *G. rostochiensis* (Rehman *et al.*, 2008).

Mutations resulting in aberrant splicing are known to hamper export of transcripts from the nucleus (Boelens *et al.*, 1995). Two different splice forms have been found for *SECPEP5*, either including or excluding exon 4. However, two *SECPEP5* specific probes, one including the exon 4 sequence and the other one without, show

the same, nuclear, localisation. *SECPEP5* may occur in a different, yet unknown, splice variant which represents a version that is transported to the cytoplasm. Alternatively, the nuclear retention of *SECPEP5* transcripts may be required in preparasitic juveniles, and export may be limited to parasitic stages of the nematode life cycle when immediate secretion into the plant cell can take place. The dorsal gland of *G. rostochiensis* consists of only a single cell, whick makes it likely that all secretory products are released simultaneously. We postulate that differential regulation of the nature of secretory effectors released may take place by nuclear sequesteration of mRNAs. A shift from a nuclear localisation in preparasitic juveniles to cytoplasmic localisation in parasitic juveniles was observed upon hybridisation of the parasitsm gene clone 28B03 from the cyst nematode *Heterodera glycines* (Gao *et al.*, 2003).

The type of alternative splicing observed for *SECPEP5* is known as exon skipping, and it is the main form of alternative splicing found in humans (Ast, 2004). An estimated 40-60% of the human genes occur in alternatively spliced forms (Hanke *et al.*, 1999, Modrek & Lee, 2002). Alternative splicing has also been observed in the free living nematode *C. elegans* and in the animal pathogen *Haemonchus contortus* (Nikolaou *et al.*, 2004, Rukov *et al.*, 2007). As far as we know, *SECPEP5* is the first gene from a plant parasitic nematode reported to undergo alternative splicing of exons other than the spliced leader exon. Alternative splicing may be a common mechanism for generating functional and/or structural diversity in gene products of plant parasitic nematodes. Also for *SECPEP9* an alternatively spliced variant has been found. In this case the splice variants differ in the retention or exonization of intron 1 (data not shown).

The dorsal esophageal gland is believed to be a major source for effectors involved nematode parasitism. Secretion of the SECPEPs is indicated by the presence of a predicted N-terminal signal peptide (SP) for secretion. Besides the presence of a signal peptide, the likely secretion of the SECPEPs is further supported by the binding of the α-SECPEP1-GST antiserum to granular structures in the cytoplasm and extension

of the dorsal esophageal gland of preparasitic juveniles of *G. rostochiensis*. Immuno-localisation studies with antiserum against SECPEP3 and SECPEP5 did not result in specific labeling of the glands in preparasitic juveniles. For SECPEP3, this may be explained by its developmental expression predominantly in later parasitic stages of the nematode life cycle. The absence of specific binding of SECPEP5 antiserum may be due to a low amount of protein in preparasitic stages, which is supported by the low expression levels we observed in the RT-PCR and by the relatively faint signal observed upon *in-situ* hybridisation.

The *SECPEP* genes show a variety in expression patterns, with *SECPEP2* and *5* being predominantly expressed in the preparasitic, invasive stage of the nematode life cycle, which may point to a role very early in the parasitic interaction, e.g. in migration or in evading the host defense response during migration. *SECPEP1* and *4* are also expressed in preparasitic stages, but their expression remains high throughout the J2 parasitic stage. *SECPEP3* and *SECPEP6—9* are predominantly expressed in the J2 parasitic stage with *SECPEP3* expression remaining high throughout the later J3 and J4 parasitic stages, and lowering in the adult stages.

The *SECPEP*s encode host defense effector molecules

The *SECPEP* genes show a striking pattern of sequence divergence, with sequence conservation in signal peptide (SP) and non-coding regions, which is remarkably contrasted by a highly diverged mature peptide. Interestingly, the sequence conservation of the SP-region extends into the 8-10 residues C-terminal of the predicted SP. Our first hypothesis is that the SECPEPs have an alternative SP cleavage site, which is located further C-terminal than the predicted one. Close inspection of the SignalP score-list indicates the presence of a second putative SP cleavage site (CxG/xD) in most of the SECPEPs. The alternative hypothesis proposes the presence of a conserved domain or motif just C-terminal of the SP. Interestingly, the tri-residue C-X-G motif found in all the SECPEPs is identical to the γ-core motif found in host defence

effector polypeptides, e.g. antimicrobial peptides (AMPs), venom and spider toxins (Yeaman & Yount, 2007, Yount & Yeaman, 2004). In addition to the γ-core motif, the SECPEPs share their small size, positive charge and sequence diversity with host defense effector polypeptides (Andreu & Rivas, 1998, Yount & Yeaman, 2004, Bulet & Sto?cklin, 2005). The γ-core motif may reflect membrane-targeting or receptor activation activities, as most host defence effector molecules function by interaction with negatively charged microbial membranes or by alteralting ion-fluxes over eukaryotic membranes. Additionally, the γ-core motif may play a role in the molecular diversification of host defence polypeptides (see below)(Yeaman & Yount, 2007). Gene duplication followed by positive Darwinian selection is one of the main evolutionary mechanism involved in the generation of the sequence divergence observed for several host defence polypeptides (Kordis & Gubens?ek, 2000; Silverstein *et al.*, 2005; Tennessen, 2005) see also chapter 6 of this thesis).

It is all in the family: mechanisms for molecular diversification of the *SECPEP*s

We showed that positive selection is acting on some of the residues of the mature peptide residing in exon 1. The majority of residues constituting the mature peptide resides in exon 2. However, the sequence divergence of this exon is so extreme that alignment and PAML analysis was not possible. In addition to positive selection we postulate four other mechanisms that may contribute to *SECPEP* molecular diversification: 1) exon shuffling 2) alternative splicing (described above) 3) hypermutation and 4) RNA editing.

Exon shuffling is one of the primary mechanisms resulting in molecular diversification of host defence effector polypeptides (reviewed by (Yeaman & Yount, 2007). Shuffled exons tend to correspond to functional domains or modules and therefore often constitute of a multitude of 3 basepairs (Gilbert *et al.*, 1997, Patthy, 1987, Patthy, 1999). An additional explanation for this observation is the fact that only insertion of these symmetrical exons will retain the existing reading frame in the case

that insertion takes place into a previously existing intron. The observation that all *SECPEP* exons completely residing within the coding area, are a multitude of 3 bp supports the hypothesis that exon-shuffling has contributed to *SECPEP* diversity. One could envision that the *SECPEPs* originate from a shuffling event between an AMP-gene and a parasitism gene specifically expressed in the dorsal esophageal gland. This would provide the nematode with a quick route for the adaptation of antimicrobial peptides in parasitism.

In order to retain the original reading frame, the shuffled exon should be of the same phase-class as the intron it is inserted into. Analysis of the intron phase of 11,117 eukaryotic introns shows that the three intron phase classes are far from equally distributed, with an excess of phase 0 introns (48%), while 30% of the introns are of phase 1 (Long *et al.*, 1995). However, an excess (74%) of single exon domains are flanked by phase 1 introns (1-1 exons), which is regarded as a signature of exon shuffling (Kaessmann *et al.*, 2002). Nearly all *SECPEP* introns are of phase 1, which leads us to believe that exon shuffling of the 1-1 exons may have happened in *SECPEP1* and *SECPEP5*. For *SECPEP2—3* and *SECPEP8* the termination codon resides in exon 2, so only the phase of the upstream flanking intron determines the reading frame of the shuffled exon. The upstream intron is in phase 1, which may indicate that also exon 2 of *SECPEP2—3* and *SECPEP8* are the result of an exon shuffling event. For *SECPEP3* the sequence similarity with the other *SECPEPs* is found in exon 3 rather than for the 3'UTR. This pattern can partly be explained by exon shuffling of exon 2. On the other hand, a recent shift of the *SECPEP3* stop-codon may be an alternative explanation. Two types of exon shuffling have been reported, exon insertion and exon duplication (Li & Graur, 1991). The sequence similarity of exon 3, 4 and 5 of *SECPEP5* suggest that they may result from recent exon duplication event. Exon 4 of *SECPEP5* is alternatively spliced, which is common for duplicated exons (Kondrashov & Koonin, 2001, Letunic *et al.*, 2002).

In analogy to the mechanism that contributes to immunoglobulin genes diversification in mammals (Papavasiliou & Schatz, 2002, Maizels, 2005), focal hypermutation has been postulated to result in hyperdivergence of conus peptides and preprodermaseptins, a class of host defence effector molecules found in the skin of frogs belonging to the families *Ranidae* and *Hylidae* (Conticello *et al.*, 2001)(Nicolas *et al.*, 2003), (Olivera *et al.*, 1999). For the frog preprodermaseptins, hypermutation is hypothesized to result from replication by a mutagenic DNA-polymerase, similar to *Escherichia coli* pol V (Nicolas et al., 2003). The conserved cysteine residues and the SP and prodomain are believed to be protected from the mutational action of the polymerase. Recently, Yeaman and Yount suggested a role for the γ-core motif in hypermutability tolerance (Yeaman & Yount, 2007). Interestingly, the γ-core motif separates the conserved SP region from the hypervariable mature peptide region in the *SECPEP*s.

Sequence analysis of a small library of *SECPEP1* cDNA and gDNA reveals a striking pattern of nucleotide changes (van Bers *et al.,* chapter 5 of this thesis). In *SECPEP1* cDNA significantly more non-synonymous mutations are found, as compared to *SECPEP1* gDNA or *Gr-gpd* cDNA. Additionally, none of the non-synonymous changes observed in *SECPEP1* gDNA or in *Gr-gpd* cDNA are radical, while more than half of the nucleotide changes observed in the *SECPEP1* cDNA clones result in a charge or polarity change of the residue. This indicates that *SECPEP1* diversification mainly takes place at transcript level, either during or after transcription. This observation can be explained by a process called RNA editing. RNA editing is, besides alternative splicing, a RNA processing event resulting in the generation of different mRNAs from a single gene (Herbert & Rich, 1999). The requirement for RNA polymerase fidelity is less stringent than for DNA polymerases, where misincorporation errors can lead to permanent genetic changes (Nesser *et al.*, 2006). Alternatively or additionally to RNA editing, a RNA polymerase II complex with decreased fidelity may have contributed to *SECPEP1* transcript diversity.

Two models explaining the effector function of the *SECPEP*s: host defense or signalling molecule?

A genomic organisation characterized by a conserved SP, and non-coding regions, and a highly divergent mature peptide is found for a large number of genes coding for small, positively charged proteins involved in host-parasite interactions (reviewed in chapter 6 of this thesis). For the cognate effector encoding gene *avrL567* from the flax rust fungus *Melampsora lini*, the pattern of extreme sequence divergence has been suggested to result from the direct interaction between *avrL567* and the corresponding resistance gene product (Dodds *et al.*, 2006), and also for the product of the avirulence gene *ATR13* from *Hyaloperonospora parasitica* a direct interaction with its corresponding resistance gene product RRP13 has been hypothesized (Allen *et al.*, 2004). In this line of thought, the *SECPEPs* may directly interact with plant (defense) factors, and as such may have (a)virulence activity on yet unknown resistant genotypes.

The net positive charge of several host defence polypeptides is, along with clusters of hydrophobic amino-acids, believed to be required for the interaction with the negatively charged microbial membranes (Yeaman & Yount, 2003). The *SECPEPs* are all positively charged, but are overall hydrophilic proteins, which lack such clusters of hydrophobic amino-acids. This may indicate that the *SECPEPs* interact with negatively charged ligands other than microbial membranes. An increasing number of AMP-like molecules are emerging to have functions other than protection against microbes, e.g. having anti-inflammatory, cell proliferative or protease inhibitory effects (Scott *et al.*, 2007, Brogden *et al.*, 2005, Wong *et al.*, 2006, Silverstein *et al.*, 2007). Additionally, venom toxins and AMP-like molecules have been shown to inhibit the formation of reactive oxygen species (Pukala *et al.*, 2007, Zeng *et al.*, 2005, Brown & Hancock, 2006).

Boman (Boman, 2003) used the average hydrophilicity of the amino-acid side chains (Radzicka & Wolfenden, 1988) to distinguish the “classical” antimicrobial

peptides that only show antimicrobial activity from antimicrobial peptides having an additional or alternative function as peptide hormones. He showed that AMPs functioning as peptide hormones have overall more hydrophilic amino-acid side chains than "classical" AMPs. The hydrophilicity of amino-acid side chains is also expressed by the GRAVY score which is deduced from the Kyte-Doolittle hydropathy plots (Kyte & Doolittle, 1982). The GRAVY score for the "classical" AMP HNP-1 is +0.2. This score is indeed higher than the values of -0.72 and -0.64 found for respectively LL37 and VIP, which are non-classical AMPs acting as respectively a peptide hormone and a neurotransmitter (Boman, 2003). The GRAVY scores calculated for the *SECPEPs* range from -0.77 to -1.83, which may suggest a role as peptide hormones for these peptides.

In animals, peptides are the most common mediators of cell-to-cell interactions, which may be attributed to their variation in size, sequence and post-translational modifications Alberts, 1997; (Matsubayashi & Sakagami, 2006). In plants, biologically active peptides are currently regarded as a new class of plant hormones (Matsubayashi & Sakagami, 2006). To date, a few hundred putative signalling peptides have been discovered in plants (Germain *et al.*, 2006), and this number is believed to be only the tip of the ice-berg. For pathogens, the use of peptides like the SECPEPs would be a sophisticated way of interfering with plant defense and development.

In conclusion, expression of the *SECPEPs* in the dorsal esophageal gland of *Globodera rostochiensis*, and the presence of a signal peptide makes it likely that they are secreted into the host cell cytoplasm, and as such play a role in parasitism. To the best of our knowledge, this paper is the first to report on peptides from a *parasite* showing AMP-like characteristics. There are two possible models for the role of AMP-like peptides for plant-parasitic nematodes. The SECPEPs may function as classical AMPs, and in that way could protect the host plant from secondary infections from other pathogens. *Globodera rostochiensis* is a biotrophic nematode, and is for the completion of its life cycle over six weeks utterly dependent on a viable host (Turner & Evans, 1998). The migratory tract of damaged host cells caused by the invading

nematode creates opportunities for other pathogens to invade the host tissue. These secondary infections pose a threat to the nematodes' food source, which makes it likely that nematodes have evolved mechanisms to secure host plants against opportunistic microbes. Currently, we are investigating if (some of) the SECPEPs have antimicrobial activity.

The expression of a subset of the *SECPEPs* in later parasitic stages of the nematode life cycle, and the hydrophilic nature of the SECPEP peptides leads us to propose an alternative model in which the SECPEPs function as peptide hormones. Interestingly, the fraction of small proteins (<3 kDa) in the secretions of *G. rostochiensis* has been shown to co-stimulate cell proliferation of tobacco protoplasts (Goverse *et al.*, 1999). Additionally, SECPEP3, SECPEP6 and SECPEP7 share sequence motifs with cyclin dependent kinase inhibitors (chapter 4 of this thesis). Currently, we are investigating if the SECPEPs have cell proliferative activity.

References

Abbas, A. K. and Lichtman, A. H. (2003) *Cellular and Molecular Immunology*. Philadelphia: Saunders.

Allen, R. L., Bittner-Eddy, P. D., Grenville-Briggs, L. J., Meitz, J. C., Rehmany, A. P., Rose, L. E.*, et al.* (2004) Host-parasite coevolutionary conflict between Arabidopsis and downy mildew. *Science,* **306,** 1957-1960.

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

Andreu, D. and Rivas, L. (1998) Animal antimicrobial peptides: an overview. *Biopolymers,* **47,** 415-433.

Ast, G. (2004) How did alternative splicing evolve? *Nature Reviews Genetics,* **5,** 773-782.

Bakker, J., Folkertsma, R. T., Rouppe van der Voort, J. N. A. M., de Boer, J. M. and Gommers, F. J. (1993) Changing concepts and molecular approaches in the

management of virulence genes in potato cyst nematodes. *Annual Review of Phytopathology,* **31,** 169-190.

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

Blumenthal, T. and Thomas, J. (1988) Cis and trans mRNA splicing in C. elegans. *Trends in Genetics,* **4,** 305-308.

Blumenthal, T. (1995) Trans-splicing and polycistronic transcription in Caenorhabditis elegans. *Trends Genet,* **11,** 132-136.

Blumenthal, T. and Steward, K. (1997) RNA Processing and gene structure. In: *C. elegans II.* (Riddle, D. L., Blumenthal, T., Meyer, B. J. and Priess, J. R., ed.^eds.). Cold Spring Harbor Laboratory Press, pp. 117-146.

Boelens, W. C., Dargemont, C. and Mattaj, I. W. (1995) Export of mRNA through the nuclear pore complex. In: *Pre-mRNA processing.* (Lamond, A. I., ed.^eds.). Heidelberg: Springer-Verlag, pp. 173-186.

Boman, H. G. (2003) Antibacterial peptides: Basic facts and emerging concepts. *Journal of Internal Medicine,* **254,** 197-215.

Brogden, K. A., Guthmiller, J. M., Salzet, M. and Zasloff, M. (2005) The nervous system and innate immunity: The neuropeptide connection. *Nature Immunology,* **6,** 558-564.

Brown, K. L. and Hancock, R. E. W. (2006) Cationic host defense (antimicrobial) peptides. *Current Opinion in Immunology,* **18,** 24-30.

Bulet, P. and Sto?cklin, R. (2005) Insect antimicrobial peptides: Structures, properties and gene regulation. *Protein and Peptide Letters,* **12,** 3-11.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T., Higgins, D.*, et al.* (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res,* **31,** 3497-3500.

Clamp, M., Cuff, J., Searle, S. M. and Barton, G. J. (2004) The Jalview Java Alignment Editor. *Bioinformatics,* **20,** 426-427.

Conticello, S. G., Gilad, Y., Avidan, N., Ben-Asher, E., Levy, Z. and Fainzilber, M. (2001) Mechanisms for evolving hypervariability: the case of conopeptides. *Mol Biol Evol,* **18,** 120-131.

Curran, J., Baillie, D. L. and Webster, J. M. (1985) Use of genomic DNA restriction fragment length differences to identify nematode species. *Parasitology,* **90,** 137-144.

De Almeida Engler, J., De Vleesschauwer, V., Burssens, S., Celenza, J. L., Jr., Inze, D., Van Montagu, M.*, et al.* (1999) Molecular markers and cell cycle inhibitors show the importance of cell cycle progression in nematode-induced galls and syncytia. *Plant Cell,* **11,** 793-808.

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.

de Boer, J. M., Smant, G., Goverse, A., Davis, E. L., Overmars, H. A., Pomp, H.*, et al.* (1996) Secretory granule proteins from the subventral esophageal glands of the potato cyst nematode identified by monoclonal antibodies to a protein fraction from second-stage juveniles. *Mol Plant Microbe Interact,* **9,** 39-46.

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

Dodds, P. N., Lawrence, G. J., Catanzariti, A. M., Teh, T., Wang, C. I. A., Ayliffe, M. A.*, et al.* (2006) Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences of the United States of America,* **103,** 8888-8893.

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

Germain, H., Chevalier, E. and Matton, D. P. (2006) Plant bioactive peptides: An expanding class of signaling molecules. *Canadian Journal of Botany,* **84,** 1-19.

Gilbert, W., De Souza, S. J. and Long, M. (1997) Origin of genes. *Proceedings of the National Academy of Sciences of the United States of America,* **94,** 7698-7703.

Goverse, A., de Engler, J. A., Verhees, J., van der Krol, S., Helder, J. H. and Gheysen, G. (2000) Cell cycle activation by plant parasitic nematodes. *Plant Mol Biol,* **43,** 747-761.

Goverse, A., Rouppe van der Voort, J., Roppe van der Voort, C., Kavelaars, A., Smant, G., Schots, A.*, et al.* (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.

Hanke, J., Brett, D., Zastrow, I., Aydin, A., Delbru?ck, S., Lehmann, G.*, et al.* (1999) Alternative splicing of human genes: More the rule than the exception? *Trends in Genetics,* **15,** 389-390.

Harlowe, E. and Lane, D. (1997) *Antibodies- A laboratory manual*. New York.

Herbert, A. and Rich, A. (1999) RNA processing and the evolution of eukaryotes. *Nature Genetics,* **21,** 265-269.

Hussey, R. S. (1989) *Disease-inducing secretions of plant-parasitic nematodes*.

Ithal, N., Recknor, J., Nettleton, D., Hearne, L., Maier, T., Baum, T. J.*, et al.* (2007) Parallel genome-wide expression profiling of host and pathogen during soybean cyst nematode infection of soybean. *Molecular Plant-Microbe Interactions,* **20,** 293-305.

Janssen, R., Bakker, J. and Gommers, F. J. (1990) Selection of virulent and avirulent lines of *Globodera rostochiensis* for the *H1* resistance gene in *Solanum tuberosum* ssp. *andigena* CPC 1673. *Revue de Nematologie,* **13,** 265-268.

Jones, J. T. and Robertson, W. M. (1997) Nematode secretions. In: *Cellular and*

molecular aspects of plant-nematode interactions. (Fenoll, C., Grundler, F. M. W. and Ohl, S. A., ed.^eds.). Kluwer Academic Publishers, pp. 98-106.

Kaessmann, H., Zo?llner, S., Nekrutenko, A. and Li, W. H. (2002) Signatures of domain shuffling in the human genome. *Genome Research,* **12,** 1642-1650.

Karczmarek, A., Overmars, H., Helder, J. and Goverse, A. (2004) Feeding cell development by cyst and root-knot nematodes involves a similar early, local and transient activation of a specific auxin-inducible promoter element. *Molecular Plant Pathology,* **5,** 343-346.

Karlin, S. and Altschul, S. F. (1990) Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proceedings of the National Academy of Sciences of the United States of America,* **87,** 2264-2268.

Kondrashov, F. A. and Koonin, E. V. (2001) Origin of alternative splicing by tandem exon duplication. *Human Molecular Genetics,* **10,** 2661-2669.

Kordis, D. and Gubens?ek, F. (2000) Adaptive evolution of animal toxin multigene families. *Gene,* **261,** 43-52.

Kyte, J. and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology,* **157,** 105-132.

Letunic, I., Copley, R. R. and Bork, P. (2002) Common exon duplication in animals and its role in alternative splicing. *Hum Mol Genet,* **11,** 1561-1567.

Li, W.-H. and Graur, D. (1991) *Fundamentals of molecular evolution*. Sunderland, Massachusetts: Sinauer Associates, Inc.

Long, M., Rosenberg, C. and Gilbert, W. (1995) Intron phase correlations and the evolution of the intron/exon structure of genes. *Proceedings of the National Academy of Sciences of the United States of America,* **92,** 12495-12499.

Maizels, N. (2005) Immunoglobulin gene diversification. *Annu Rev Genet,* **39,** 23-46.

Martoglio, B. and Dobberstein, B. (1998) Signal sequences: more than just greasy peptides. *Trends Cell Biol,* **8,** 410-415.

Matsubayashi, Y. and Sakagami, Y. (2006) Peptide Hormones in Plants. *Annu Rev Plant Biol*.

Modrek, B. and Lee, C. (2002) A genomic view of alternative splicing. *Nature Genetics,* **30,** 13-19.

Nesser, N. K., Peterson, D. O. and Hawley, D. K. (2006) RNA polymerase II subunit Rpb9 is important for transcriptional fidelity in vivo. *Proceedings of the National Academy of Sciences of the United States of America,* **103,** 3268-3273.

Nicolas, P., Vanhoye, D. and Amiche, M. (2003) Molecular strategies in biological evolution of antimicrobial peptides. *Peptides,* **24,** 1669-1680.

Niebel, A., de Almeida Engler, J., Hemerly, A., Ferreira, P., Inze, D., Van Montagu, M.*, et al.* (1996) Induction of cdc2a and cyc1At expression in Arabidopsis thaliana during early phases of nematode-induced feeding cell formation. *Plant J,* **10,** 1037-1043.

Nikolaou, S., Hartman, D., Nisbet, A. J., Presidente, P. J. A. and Gasser, R. B. (2004) Genomic organization and expression analysis for hcstk, a serine/threonine protein kinase gene of Haemonchus contortus, and comparison with Caenorhabditis elegans par-1. *Gene,* **343,** 313-322.

Olivera, B. M., Gray, W. R. and Zeikus, R. (1985) Peptide neurotoxins from fish-hunting cone snails. *Science,* **230,** 1338-1343.

Olivera, B. M., Walker, C., Cartier, G. E., Hooper, D., Santos, A. D., Schoenfeld, R.*, et al.* (1999) Speciation of cone snails and interspecific hyperdivergence of their venom peptides. Potential evolutionary significance of introns. *Ann N Y Acad Sci,* **870,** 223-237.

Papavasiliou, F. N. and Schatz, D. G. (2002) Somatic hypermutation of immunoglobulin genes: merging mechanisms for genetic diversity. *Cell,* **109 Suppl,** S35-44.

Patel, N. and Goodman, C. (1992) *Nonradioactive In Situ Hybridization Application Manual*. Boehringer Mannheim Biochemica, Mannheim, Germany.

Patthy, L. (1987) Intron-dependent evolution: preferred types of exons and introns.

FEBS Letters, **214,** 1-7.

Patthy, L. (1999) Genome evolution and the evolution of exon-shuffling - A review. *Gene,* **238,** 103-114.

Polson, A., Von Wechmar, M. B. and Van Regenmortel, M. H. V. (1980) Isolation of viral IgY antibodies from yolks of immunized hens. *Immunological Communications,* **9,** 475-493.

Pukala, T. L., Doyle, J. R., Llewellyn, L. E., Kuhn-Nentwig, L., Apponyi, M. A., Separovic, F.*, et al.* (2007) Cupiennin 1a, an antimicrobial peptide from the venom of the neotropical wandering spider Cupiennius salei, also inhibits the formation of nitric oxide by neuronal nitric oxide synthase. *Febs J.*

Qin, L., Overmars, H., Helder, J., Popeijus, H., van der Voort, J. R., Groenink, W.*, et al.* (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.

Qin, L., Prins, P., Jones, J. T., Popeijus, H., Smant, G., Bakker, J.*, et al.* (2001) GenEST, a powerful bidirectional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP. *Nucleic Acids Research,* **29,** 1616-1622.

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.

Radzicka, A. and Wolfenden, R. (1988) Comparing the polarities of the amino acids side-chain distribution coefficients between the vapor phase cyclohexane 1 octanol and neutral aqueous solution. *Biochemistry . 1988;,* **27,** 1664-1670.

Rehman, S., Qin, L., Overmars, H., Petrescu, A. J., Thomma, B. P. H. J., Helder, J.*, et al.* (2008) Overexpression of a nematode SPRYSEC-gene in potato promotes

the virulence of a nematode, a fungus and a virus. *in PhD thesis, Wageningen University, the Netherlands*, **74**-107.

Rukov, J. L., Irimia, M., Mørk, S., Lund, V. K., Vinther, J. and Arctander, P. (2007) High qualitative and quantitative conservation of alternative splicing in Caenorhabditis elegans and Caenorhabditis briggsae. *Molecular Biology and Evolution,* **24,** 909-917.

Scott, M. G., Dullaghan, E., Mookherjee, N., Glavas, N., Waldbrook, M., Thompson, A.*, et al.* (2007) An anti-infective peptide that selectively modulates the innate immune response. *Nat Biotechnol,* **25,** 465-472.

Silverstein, K. A. T., Graham, M. A., Paape, T. D. and Vandenbosch, K. A. (2005) Genome organization of more than 300 defensin-like genes in Arabidopsis. *Plant Physiology,* **138,** 600-610.

Silverstein, K. A. T., Moskal Jr, W. A., Wu, H. C., Underwood, B. A., Graham, M. A., Town, C. D.*, et al.* (2007) Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants. *Plant Journal,* **51,** 262-280.

Tennessen, J. A. (2005) Enhanced synonymous site divergence in positively selected vertebrate antimicrobial peptide genes. *Journal of Molecular Evolution,* **61,** 445-455.

Thomas, J. H. (2006) Adaptive evolution in two large families of ubiquitin-ligase adapters in nematodes and plants. *Genome Research,* **16,** 1017-1030.

Thomas, J. H., Kelly, J. L., Robertson, H. M., Ly, K. and Swanson, W. J. (2005) Adaptive evolution in the SRZ chemoreceptor families of Caenorhabditis elegans and Caenorhabditis briggsae. *Proceedings of the National Academy of Sciences of the United States of America,* **102,** 4476-4481.

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

Trivedi, S. and Arasu, P. (2005) Evaluation of endogenous reference genes for real-time PCR quantification of gene expression in Ancylostoma caninum. *Molecular and Biochemical Parasitology,* **143,** 241-244.

Turner, S. J. and Evans, K. (1998) The origins, global distribution and biology of potato cyst nematodes (*Globodera rostochiensis* (Woll.) and *Globodera pallida* Stone). In: *Potato cyst nematodes: biology, distribution and control.* (Marks, R. J. and Brodie, B. B., ed.^eds.). Wallingford: CAB International, pp. 7-26.

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

von Mende, N., Gravato-Nobre, M. J. and Perry, R. N. (1998) Host finding, invasion and feeding. In: *The cyst nematodes.* (Sharma, S., ed.^eds.). Dordrecht: Kluwer Academic Publishers.

Win, J., Morgan, W., Bos, J., Krasileva, K. V., Cano, L. M., Chaparro-Garcia, A.*, et al.* (2007) Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Cell,* **19,** 2349-2369.

Wong, J. H., Zhang, X. Q., Wang, H. X. and Ng, T. B. (2006) A mitogenic defensin from white cloud beans (Phaseolus vulgaris). *Peptides,* **27,** 2075-2081.

Wong, W. S., Yang, Z., Goldman, N. and R., N. (2004) Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. *Genetics,* **168,** 1041-1051.

Wyss, U. (1992) Observations on the feeding behaviour of *Heterodera schachtii* throughout development, including events during moulting. *Fundam. appl. Nematol.,* **15,** 75-89.

Yang, Z. (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci,* **13,** 555-556.

Yang, Z. (2006) *Computational Molecular Evolution*. Oxford: oxford University Press.

Yang, Z. and Bielawski, J. P. (2000) Statistical methods for detecting molecular

adaptation. *Trends in Ecology and Evolution,* **15,** 496-503.

Yang, Z. and Nielsen, R. (2000) Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Molecular Biology and Evolution,* **17,** 32-43.

Yeaman, M. R. and Yount, N. Y. (2003) Mechanisms of antimicrobial peptide action and resistance. *Pharmacological Reviews,* **55,** 27-55.

Yeaman, M. R. and Yount, N. Y. (2007) Unifying themes in host defence effector polypeptides. *Nature Reviews Microbiology,* **5,** 727-740.

Yount, N. Y. and Yeaman, M. R. (2004) Multidimensional signatures in antimicrobial peptides. *Proceedings of the National Academy of Sciences of the United States of America,* **101,** 7363-7368.

Zeng, X. C., Corzo, G. and Hahin, R. (2005) Scorpion venom peptides without disulfide bridges. *IUBMB Life. 2005;;* **57,** 13-21.

Zorio, D. A., Cheng, N. N., Blumenthal, T. and Spieth, J. (1994) Operons as a common form of chromosomal organization in C. elegans. *Nature,* **372,** 270-272.

Chapter 3

NEMPEP, a γ-core effector polypeptide from the potato cyst nematode *Globodera rostochiensis*, accumulates in the nucleolus of tobacco cells

Nikkie van Bers, Hein Overmars, Erik Slootweg, Ling Qin, Aska Goverse, Jaap Bakker and Geert Smant

Abstract

Secreted effectors from cyst nematodes are essential for succesfull feeding site induction and maintenance. In the past decade, research has mainly focussed on the identification of novel nematode effectors, of which a major subset codes for pioneering proteins. In this paper, we describe the cloning and functional characterisation of the novel peptide NEMPEP. NEMPEP is a positively charged, hydrophilic peptide with a N-terminal γ-core (G-X-C) motif. This motif of unknown function is found in a wide range of host defense effector polypeptides showing antimicrobial activity. To provide insight in the functional role of NEMPEP in parasitism, we expressed *NEMPEP* in plants. Subcellular localization studies of NEMPEP-GFP fusion proteins shows that NEMPEP accumulates in the nucleolus of tobacco cells. Based on the phenotype of some of the potato and *Arabidopsis* plants transformed with a *35S::NEMPEP* overexpression construct (with or without GFP) we hypothesize intereference of NEMPEP with cytokinin signaling.

Introduction

The potato cyst nematode (*Globodera rostochiensis* and *G. pallida*) parasitizes on a small range of Solanaceous plants, which includes the economically important crops potato, tomato and eggplant. The life cycle of cyst nematodes consists of four juvenile stages followed by the adult stage (Lee, 2002, Turner & Evans, 1998, Zuncke & Eisenback, 1998). Upon hatching, the pre-parasitic J2 migrates towards and subsequently invades the plant root in the cell elongation zone (von Mende *et al.*, 1998). Soon after feeding cell induction by parasitic J2 juveniles, the nematode loses its ability to move and becomes sedentary, while feeding from the plant. During this feeding process the nematode undergoes three additional moults, into J3, J4 and finally into the adult stage.

Cyst nematodes have enlarged dorsal and subventral esophageal glands, producing secretory effector molecules. Nematode effectors facilitate migration, and may play a role in evasion or suppression of host defense. This is important, as the nematode resides for an extended period of time in extremely close contact with its host (Williamson & Kumar, 2006). Additionally, nematode effectors are essential for feeding cell induction and maintenance. To that purpose, the oral stylet delivers the effector molecules into the initial feeding cell, usually a procambium or pericycle cell (Doncaster & Seymour, 1973, Endo, 1964). Subsequent progressive cell wall dissolution and protoplast fusion results in a multinuclear cell complex, which can consist of up to 200 cells (Golinowski *et al.*, 1997, Grundler & Böckenhoff, 1997, Zuncke & Eisenback, 1998). Enlarged nuclei are embedded in the dense cytoplasm of the syncytium. Feeding cell formation is accompanied by cell cycle reactivation, and the enlarged nuclei are the result of endoreduplication cycles in which DNA is replicated in absence of mitosis (Goverse *et al.*, 2000a, Niebel *et al.*, 1996, Wyss, 1992).

Many of the changes that occur when the nematode becomes sedentary, e.g. cell wall breakdown, endoreduplication and lateral root initiation, can be mimicked by exogenous application of the plant hormone auxin (reviewed by (Gheysen & Jones,

2006). Experimental evidence further confirmed the involvement of auxins and ethylene in feeding cell formation, e.g. inhibition of syncytium development on auxin-insensitive tomato and *Arabidopsis* mutants, a local accumulation of auxin in and around developing syncytia and hyperinfection of roots of ethylene overproducing mutants (Goverse *et al.*, 2000b), (Karczmarek *et al.*, 2004).

Peptides are currently regarded as a novel class of plant hormones (Matsubayashi & Sakagami, 2006). Several peptides have been described as effector molecules in cyst nematode parasitism. A polypeptide from the soybean cyst nematode shares sequence similarity with the CLAVATA3/ESR-related (CLE) protein family in *Arabidopsis*, and has been shown to rescue the mutant phenotype of the *clv3-1* mutant (Wang *et al.*, 2005). Recently, we characterized the *SECPEP* gene family from *Globodera rostochiensis*. This family consists of nine members, *SECPEP1 —SECPEP9*, coding for hypervariable peptides. Currently there are two models regarding the functional role of the SECPEP peptides in parasitism. The first one regards a role as defense molecules in protecting the host plant against opportunistic microbes. The second hypothesis concerns a functional role as peptide hormones interfering with hormone signaling (van Bers *et al.*, chapter 2 of this thesis). In this paper, we report on the cloning, characterization, and functional analysis of the novel effector peptide NEMPEP, specifically found in the potato cyst nematode *Globodera rostochiensis*.

Materials and methods

Nematode collection, RNA isolation and cDNA-AFLP

Preparasitic juveniles of *G. rostochiensis* pathotype Ro1-Mierenbos were collected as described previously (De Boer *et al.*, 1992). Parasitic second stage juvenile (J2) nematodes, third stage juveniles (J3) and fourth stage juveniles (J4) were collected respectively 13/14, 19 and 23 days post inoculation (dpi). Adult female and adult male nematodes were collected at respectively 34 and 27 days post inoculation (dpi). Extraction of the nematodes from plant material was performed as described

previously for adult males and females by (De Boer et al., 1992). Total RNA was isolated from preparasitic juveniles using TRIzol reagent (Invitrogen, Breda, the Netherlands). cDNA-AFLP was performed as described previously by (Qin *et al.*, 2000).

Cloning of full-length cDNA, amplification of genomic DNA and southern blotting

Five µg of total RNA was used as the starting material for full-length, RNA ligase-mediated rapid amplification of 5' and 3' cDNA ends (RLM-RACE) with the GeneRacer Kit (Invitrogen). Super-script III reverse transcriptase (Invitrogen) was used for cDNA synthesis, according to the manufacturer protocol. Full-length cDNA was amplified using gene specific primers (5' CCCAGAACAATCACCACAAGCATC 3' for 5'RACE and 5' TCGATGCTTGTGGTGATTG 3' for 3'RACE) in combination with adaptor specific primers, and was subsequently cloned into the pCR4-TOPO vector (Invitrogen). Transformed *E. coli* colonies were checked for presence of the expected insert by PCR with the same primers as used for the amplification. Plasmids were purified using the Wizard *Plus* Miniprep DNA Purification System (Promega Corporation, Madison, WI, USA) and sequenced at the Sequence Facility Wageningen (Wageningen, the Netherlands) or at BaseClear (Leiden, the Netherlands).

Genomic DNA was extracted from preparasitic J2 as described by (Curran *et al.*, 1985). Genomic DNA fragments corresponding to the *NEMPEP* were amplified with the gene specific A42 5' FW (5' GATGCTTGTGGTGATTGTTCTGGG 3') and the

A42 3' RV (5' CATTTTCGTCTTATGAGCTTGCTTCC 3') primer. Amplification products were cloned into the pCR4-TOPO vector (Invitrogen). For southern blot analysis, about 2.5 µg of genomic DNA was digested by respectively *EcoRI*, *BamHI*, *KpnI* and *BglII* and separated on a 0.8% agarose gel. The southern blot was performed as described by (Sambrook *et al.*, 1989). The 178 bp A42 TDF sequence was re-amplified from plasmid DNA with DIG-labelled dUTP (Roche Diagnostics), and used as a probe. Hybridisation was performed at 40°C overnight in DIG Easy Hyb solution (Roche Diagnostics). The

filter was washed twice in 0.1xSSC/0.1% SDS solution before immuno-detection. Alkaline phosphatase activity was detected using DIG Luminiscent Detection Kit (Roche Diagnostics) and the chemiluminiscent signal was exposed on X-ray films.

Semi-quantitive RT-PCR analysis

Transcript expression of *NEMPEP* was analysed in different parasitic stages by reverse transcription (RT)-PCR. Total RNA was extracted using TRIzol Reagent (Invitrogen). For cDNA synthesis, total RNA was treated with Turbo DNA-free (Ambion) to degrade possible contaminating genomic DNA. Superscript III reverse transcriptase was used for first-strand cDNA synthesis, essentially to the manufacturers' protocol (Invitrogen). Gene-specific and control primers were designed with the Beacon designer 4.02 program (PREMIER Biosoft International; Table 1). As an internal control for template amount and quality, we amplified cDNA derived from a cAMP-dependent protein kinase catalytic subunit encoding gene. Primers were designed to specifically anneal on the *G. rostochiensis* EST BM343563, which shows significant sequence similarity to *Ancylostoma caninum* cAMP-dependent protein kinase catalytic subunit, Genbank accession number U15983 (Trivedi & Arasu, 2005).

***In situ* hybridisation microscopy**

In situ hybridisation microscopy was done on preparasitic juveniles (J2). The nematode fixation, hybridization and detection steps were essentially as described by (de Boer *et al.*, 1998), except for the use of single stranded DNA probes, which were synthesized by linear PCR as described by (Patel & Goodman, 1992). The sequences of the primers used to generate the probes are *taq* FW (5'TCGATGCTTGTGGTGATTG 3') and *EcoRI* RV(5' GAATTCTAAAGTTTGTC 3').

Sequence analyses

DNA and amino-acid sequences were analysed using the Informax Vector NTI Advance software (Invitrogen). The computer algorithm SignalP 3.0 (Bendtsen *et al.*, 2004) was used to predict the presence of a signal peptide for secretion and the corresponding putative cleavage site (Neural Networks). BLAST at NCBI (www.ncbi.nlm.nih.gov) or at nemaBLAST (www.nematode.net) was used to search for matching sequences in the database. PSORT (Nakai & Horton, 1999) was used to predict the subcellular localization of NEMPEP, and Prosite was used to predict putative phosphorylation sites in the NEMPEP peptide sequence (Sigrist *et al.*, 2002). ProtScale and ProtParam (ExPASY web server) were used for the generation of respectively the hydropathy plots and the calculation of the Grand Average of Hydropathy (GRAVY) score.

The on-line available cysteine bonding prediction software Dlpro and Disulfind were used to investigate the presence of cysteine-bonds in NEMPEP (Cheng *et al.*, 2006, Ceroni *et al.*, 2006).

In-planta localization studies, protein isolation and immuno-blotting

The mature peptide encoding sequence of NEMPEP was PCR amplified from a cDNA library using primers incorporating *AscI* and *PacI* restriction sites (Table 1). The resulting PCR product was cloned into TOPO-pCR4 (Invitrogen), which was subsequently digested with the enzymes *Asc I* and *PacI.* The digested fragment was gel-purified and cloned into a *AscI/PacI* digested pMDC vector (Table 1)(Curtis & Grossniklaus, 2003). Error-free clones were confirmed by sequence analysis.

The constructs were transformed into *Agrobacterium tumefaciens*, strain LBA4404 (Hoekema *et al.*, 1983) by electroporation (Shen & Forde, 1989). Transient expression in *Nicotiana benthamiana* epidermal cells was performed as described previously by Kudla *et al.* (2007)(Kudla *et al.*, 2007). A Zeiss LSM 510 confocal microscope was used to visualize GFP.

Table 1: Primers used for the generation over NEMPEP over-expression constructs

Construct	Primer 1(5'→3')[a]	Primer 2(5'→3')[a]	Vector
35S::NEMPEP	GGCGCGCCAAATATGtgtggtgattgttctgggaaaagc	TTAATTAAtcattttcgtcttatgagcttgcttcc	pMDC32
35S::GFP-NEMPEP	GGCGCGCCAAATATGtgtggtgattgttctgggaaaagc	TTAATTAAtcattttcgtcttatgagcttgcttcc	pMDC45
35S::NEMPEP-GFP	TTAATTAAAATATGtgtggtgattgttctgggaaaagc	GGCGCGCCAttttcgtcttatgagcttgcttcc	pMDC201
35S::GFP	GGCGCGCCAAATATGagtaaaggagaagaacttttc	TTAATTAAtcatttgtatagttcatccatgcc	pMDC32

[a] The restriction sites, kozak sequence and start codon are depicted in capitals.

For protein isolation, leaves were harvested three days after infiltration. The leaf material was grinded in liquid nitrogen and transferred to a pre-cooled tube. Protein was extracted by vigorously mixing the tissue with 8 M ureum (100 µl/ 23 mg tissue) and protein sample buffer (6X stock contains 0.35 M Tris-HCL, 10.3% SDS (w/v), 36% glycerol, 5% ß-mercapto-EtOH and 0.012% bromophenol blue) in the presence of the serine protease inhibitor pefabloc SC Plus (Roche diagnostics). The protein sample was boiled for ten minutes, and the supernatant was separated on 12% NuPAGE NOVEX Bis-Tris gels (Invitrogen) in MES-buffer (Invitrogen) and subsequently transferred to PVDF membrane (Millipore) as described by (de Boer *et al.*, 1996). Membranes were blocked in 10% BSA in PBS supplied with 0.1% Tween20 (PBS-T) and incubated with 1:5000 rabbit polyclonal GFP antibody (Abcam) in PBS-T containing 3% BSA.

Generation of transgenic plants

Stem segments (0.5-1 cm) from *in-vitro* grown potato line V (genotype 6487-9, (Horsman *et al.*, 2001) plants were used for *Agrobacterium* mediated plant transformation with pMDC32::*NEMPEP* and pMDC201::*NEMPEP* (see Table 1) and regenerated as described by (van Engelen *et al.*, 1994). For construct preparation and *Agrobacterium* transformation, see previous section. One cutting of the each of the transformed lines was used for phenotypic analysis. Root morphology was monitored in

vitro, and for analysis of the leaf morphology the plants were transferred to pots and were studied under greenhouse conditions.

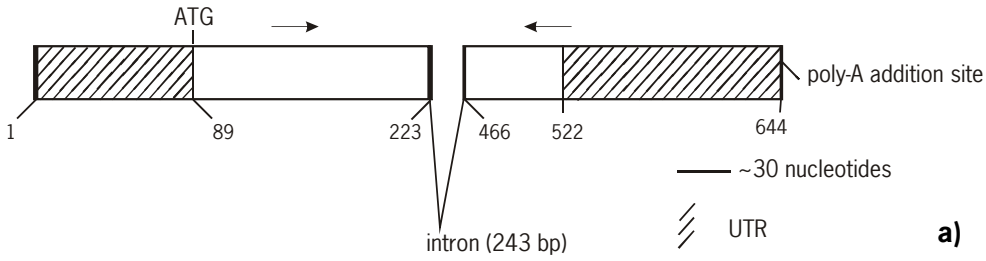
Arabidopsis thaliana was transformed with *Agrobacterium tumefaciens* harbouring the pMDC32::*NEMPEP*, pMDC32 empty vector control or pMDC45::*NEMPEP* construct by use of the floral dip method as described by (Clough & Bent, 1998). Transformed seeds were surface sterilized as described by (Karczmarek et al., 2004) and selected for transgenicity by germination on basic agar medium containing hygromycin (30 mg/L) as described by (Nakazawa & Matsui, 2003). The first generation of transformed seeds (T1) was used for phenotypic analysis. For phenotypic analysis of root growth, freshly germinated seedlings were transferred to knop medium (Sijmons *et al.*, 1991). After analysis of the root phenotype, the plants were transferred to pots and leaf morphology was studied under greenhouse conditions.

Results

Developmental expression pattern of *NEMPEP*

cDNA-Amplified Fragment Length Polymorphism (cDNA-AFLP) is a powerful tool to identify genes specifically expressed in response to abiotic or biotic stress (e.g.(Qin et al., 2000, Sarosh & Meijer, 2007, Wang & Bughrara, 2007). Previously, we reported on the use of cDNA-AFLP to acquire developmental expression profiles of *Globodera rostochiensis* genes in five successive developmental stages in order to identify genes putatively involved in parasitism (Qin et al., 2000), Qin, unpublished data). Here, we focus on the functional characterization of the Transcript Derived Fragment (TDF) with code A42, which was one of the 216 TDFs upregulated at the onset of parasitism (preparasitic J2 stage)(data not shown). Specific expression of a gene in this stage of the life cycle may indicate a functional role in parasitism. We re-amplified the TDF A42 from cDNA isolated from pre-parasitic juveniles, and primers designed on the TDF sequence were used for rapid amplification of the cDNA ends (RACE). This resulted in the amplification of a full-length transcript containing a 189 bp open reading frame (ORF)(Fig. 1a).

This ORF corresponds to a polypeptide of 62 amino-acids, which we named NEMPEP (Fig. 1b). Sequence similarity search did not reveal any significantly matching sequences outside *Globodera rostochiensis*. Within *Globodera spp.* two significantly matching ESTs were found (E- value 3e-20). These represent one clone (GRAA-aaa41c11) from a cDNA library generated from parasitic J2 of the single female line 19 of *G. rostochiensis*, which was selected for avirulence on the nematode resistance gene *H1* (Janssen *et al.*, 1990). The sequence similarity with *NEMPEP* is confined to only 176 basepairs (bp) of the 412 bp GRAA-aaa41c11 EST sequence. The region of similarity is preceeded by an intron-like sequence bordered by a 5' GT and 3' AG dinucleotide and an A/T percentage of 76%. To further investigate if an intron is retained in the GRAA-aaa41c11 sequence, we amplified the *NEMPEP* genomic sequence. Indeed, we found a 243 bp intron interrupting the *NEMPEP* ORF (Fig. 1a).



mnafiifvifivffsqiyeidacgdcsgksekkksekikhddikaflktketgggskliirrk **b)**

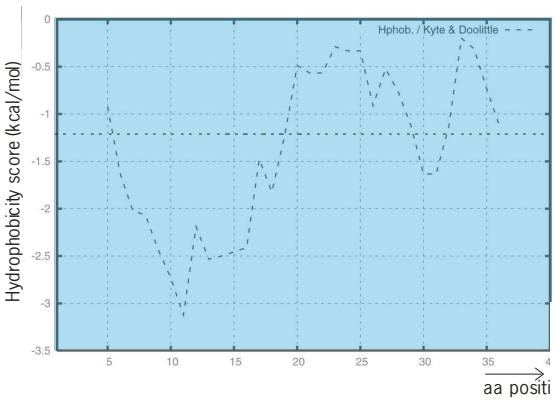


Figure 1: a) Schematic overview of the *NEMPEP* sequence. The 401 bp of the *NEMPEP* transcript was not found to be *trans*-spliced. The *NEMPEP* coding sequence is interspersed by at least one intron of 243 bp. Arrows indicate the annealing position of the primers used for the amplification of the *NEMPEP* genomic sequence. The UTR regions are shaded.

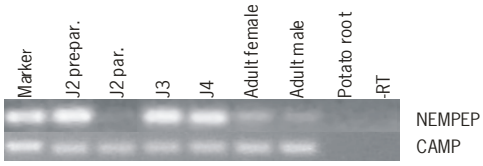
b) The sequence of the NEMPEP peptide, which includes a 62 amino-acids (aa) peptide, of which the first 22 aa correspond to a predicted signal peptide for secretion (underlined). NEMPEP contains a γ-core motif (indicated by arrows), two putative phosphorylation sites (indicated by +++ and ****) and two stretches of positively charged amino-acids (boxed), which may be involved in nuclear and/or nucleolar targeting of the protein. **c)** Hydropathy plot of NEMPEP. The position of the GRAVY score is indicated by a dotted line. The plots are generated by the Kyte-Doolittle algoritm in Protscale according to (Kyte & Doolittle, 1982)

The *NEMPEP* genomic sequence nearly exactly matches with the GRAA-aaa41c11 EST sequence (1 nucleotide difference). To investigate whether *NEMPEP* is a member of a gene family, we performed southern blot hybridisation on *G. rostochiensis* genomic

DNA digested with four different enzymes. The 178 bp TDF sequence was used as a probe, and hybridisation resulted in one single band on the blot for each of the enzymes used (data not shown).

To investigate secretion of NEMPEP, we analysed the subcellular localisation of its transcripts by *in-situ* hybridisation microscopy (Fig. 2a). This showed that *NEMPEP* is specifically expressed in the dorsal esophageal gland of pre-parasitic juveniles.The esophageal glands of nematodes are the main sites for production of secreted effector molecules involved in parasitism. The 22 N-terminal amino-acids of NEMPEP are predicted to act as a cleavable signal peptide for secretion (Fig. 1b).

a)



b)

Figure 2: a) Tissue localization of *NEMPEP* transcripts in preparasitic J2s using *in-situ* hybridization microscopy. Single stranded, DIG-labelled, DNA probes were amplified from plasmid DNA by use of linear PCR. DIG-labelling was specifically observed in the dorsal esophageal gland (arrowhead). **b)** Developmental expression pattern as determined by semi-quantitative RT-PCR of *NEMPEP* cDNA from different life stages of *G. rostochiensis.* J2 parasitic juveniles are collected 13-14 days post inoculation (dpi) of plants with cysts, J3 parasitic juveniles are collected at 19 dpi, J4 at 23 dpi, adult females at 34 dpi and adult males at 27 dpi. Potato root cDNA was included to confirm primer specificity and the sample in which reverse transcriptase was omitted (-RT) allows investigation of the presence of contaminating genomic DNA. cAMP dependent protein kinase catalytic subunit cDNA (cAMP)(EST accession nr. BM343563) was included as an internal control for equal template conditions. The samples were analysed on gel after 30 PCR cycles. All the bands observed are of the expected size.

It is believed that there are distinct stages in parasitism of nematodes in plants (*viz.* migration, feeding, and reproduction), each requiring a different set of effector

molecules. We investigated if expression of *NEMPEP* is correlated with specific events in the life cycle. Reverse-Transcription PCR showed a pronounced band of the expected size (101 bp) upon analysis of cDNA from pre-parasitic juveniles, in J3, and in J4 parasitic juveniles (Fig. 2b). We repeatedly found only a very weak expression of *NEMPEP* in parasitic J2 juveniles.

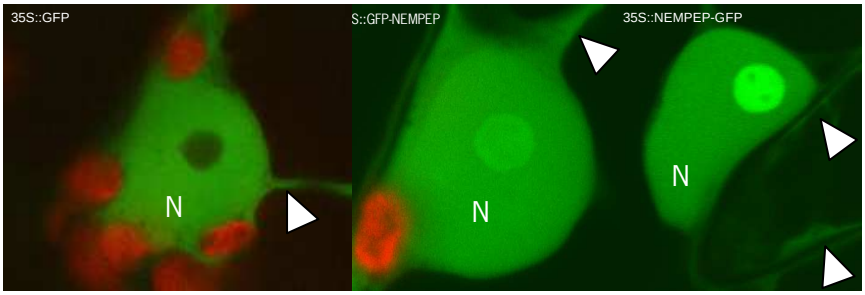
NEMPEP encodes a secretory, highly charged, γ-core polypeptide

NEMPEP contains a 62 AA ORF, which encodes a 4 kDa mature polypeptide preceeded by a signal peptide for secretion. Although the mature peptide contains two cysteine residues, no disulfide bonding is predicted with high probability. The NEMPEP mature peptide sequence is characterized by a highly positive net charge (+7) at pH 7, and contains two putative phosphorylation sites, of which the most N-terminal is located adjacent to a so-called γ-core motif (G-X-C)(Glycine-X-Cysteine)(Fig. 1b)(Yount & Yeaman, 2004, Yeaman & Yount, 2007). This motif is found in diverse classes of host defence effector proteins with antimicrobial activity, although its exact function is unknown. However, also several peptide hormones belong to this class of peptides, although their primary function does not seem to be related to antimicrobial activity. Boman (2003) used the level of hydrophilicity to discriminate between classical antimicrobial peptides (AMPs) that exhibit antimicrobial activity by interacting with microbial membranes and hydrophilic peptides having an additional or alternative function as peptide hormones. NEMPEP has a GRAVY score (Kyte & Doolittle, 1982) of -1.3 kcal/mol (Fig. 1c), which classes NEMPEP as a hydrophilic protein. The GRAVY score of NEMPEP is in the same range as found for the non-classical AMPs VIP and LL37, functioning as peptide hormones, and for a family of γ-core cyst nematode peptides, the SECPEPs (-0.64 to -1.83 kcal/mol)(van Bers et al., chapter 2). As a comparison, the classical AMP magainin 2 has a GRAVY score of +0.08 kcal/mol (van Bers et al., chapter 2).

Accumulation of NEMPEP-GFP in the nucleolus of tobacco epidermal cells

NEMPEP is rich in the positively charged amino-acids lysine and arginine (33% in frequency), and is predicted to be localized in the nucleus. No typical nuclear localisation signal (NLS) is found in the NEMPEP mature peptide sequence, but stretches of positively charged residues may act as NLS (Kalderon *et al.*, 1984). In order to investigate whether NEMPEP indeed accumulates in the nucleus of host cells, we expressed *NEMPEP* under control of the 35S promoter, either as a N- or C-terminal fusion protein with GFP (Fig. 3a). Free GFP (35S::GFP, Fig. 3a) localizes in the cytoplasm and the nucleus of tobacco cells. The plant cell nucleus is surrounded by a membrane and transport over this membrane takes place through the Nuclear Pore Complex (NPC). The diffusion limit of the NPC is about 40 kDa, depending on the biochemical and structural character of the molecule (Jans & Hu?bner, 1996). The molecular weight of GFP is 26 kDa, which allows its' free diffusion into the nucleus of tobacco cells. The nucleolus is a highly dynamic nuclear compartment with a denser structure than the nucleus (Handwerger *et al.*, 2005). Nucleolar fluorescence is hardly (data not shown) or not observed in leaves infiltrated with the 35S::GFP construct (Fig. 3a). This is in sharp contrast to leaves infiltrated with 35S::*GFP-NEMPEP* or 35S::*NEMPEP-GFP*, in which strong accumulation of fluorescence in the nucleoli was observed (Fig. 3a). Nucleolar accumulation seems to be strongest upon fusion of GFP at the C-terminus of NEMPEP. In addition to nucleolar accumulation, the NEMPEP fusion proteins are also detected in the nucleus and cytoplasm of tobacco cells (Fig. 3a).

To confirm the expected size of the fusion proteins, we performed western blot analysis of grinded leaf material overexpressing NEMPEP-GFP fusions (Figure 3b). Fusion of GFP to NEMPEP is expected to result in a molecule with a mass of 30 kDa. A band of this size was found in the lane with protein extracted from leaves infiltrated with 35S::*GFP-NEMPEP*. Immunoblotting of protein extracted from leaves expressing 35S::*NEMPEP-GFP* gives rise to a band with a slightly higher molecular weight (about 32 kDa)(Fig. 3b).



a)

b)

Figure 3: a) Tissue localization of GFP-NEMPEP (middle), NEMPEP-GFP (right), and GFP (control, left) overexpression in tobacco leaf epidermal cells, visualized by Confocal Laser Scanning Microscopy. GFP is detected in the cytoplasm (arrowheads) and nucleus of the cells (N), while GFP-NEMPEP and NEMPEP-GFP also accumulate in the nucleolus of these cells.

b) Immunoblot analysis of blotted protein extracted from leaves infiltrated with 35S::GFP (left lane), 35S::GFP-NEMPEP (middle) and 35S::NEMPEP-GFP (right) detected with an anti-GFP antibody. The molecular size (in kDa) of the corresponding marker bands is depicted at the left side.

Overexpression of *NEMPEP* lowers the transformation efficiency in both *Arabidopsis* and potato

To further investigate the functional role of NEMPEP, we generated transgenic *Arabidopsis thaliana* and *Solanum tuberosum* (potato) plants overexpressing the mature NEMPEP peptide, either alone (35S*::NEMPEP*) or as a fusion with GFP (35S::*GFP-NEMPEP* for *Arabidopsis* and 35S::*NEMPEP-GFP* for potato).

Upon transformation with 35S::*NEMPEP*, we observed for both *Arabidopsis* and potato a lower transformation efficiency (of respectively 0.005% and 1.9%) as compared to either empty vector, 35S::*NEMPEP-GFP* or 35::*GFP-NEMPEP* transformed plants (Table 2). The transformation efficiency was measured as the ratio of shoot-forming calli for potato and as percentage of hygromycin resistant seeds for *Arabidopsis*.

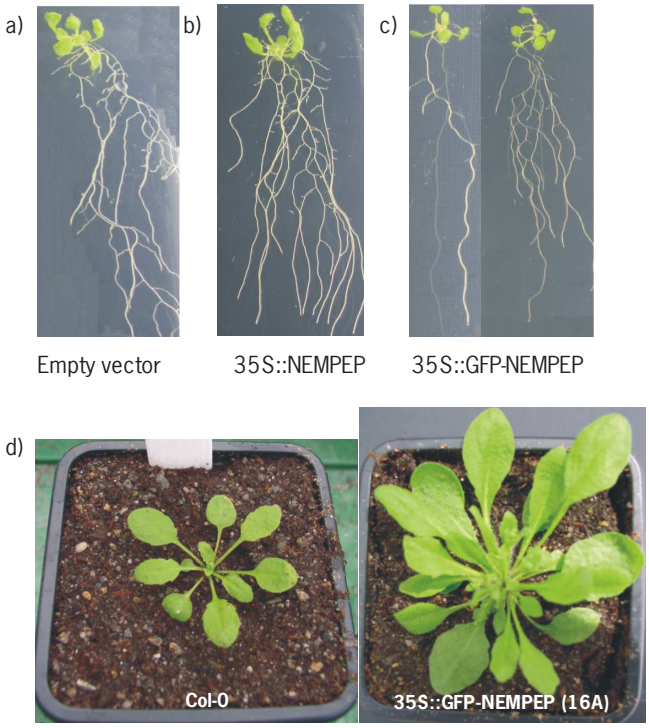
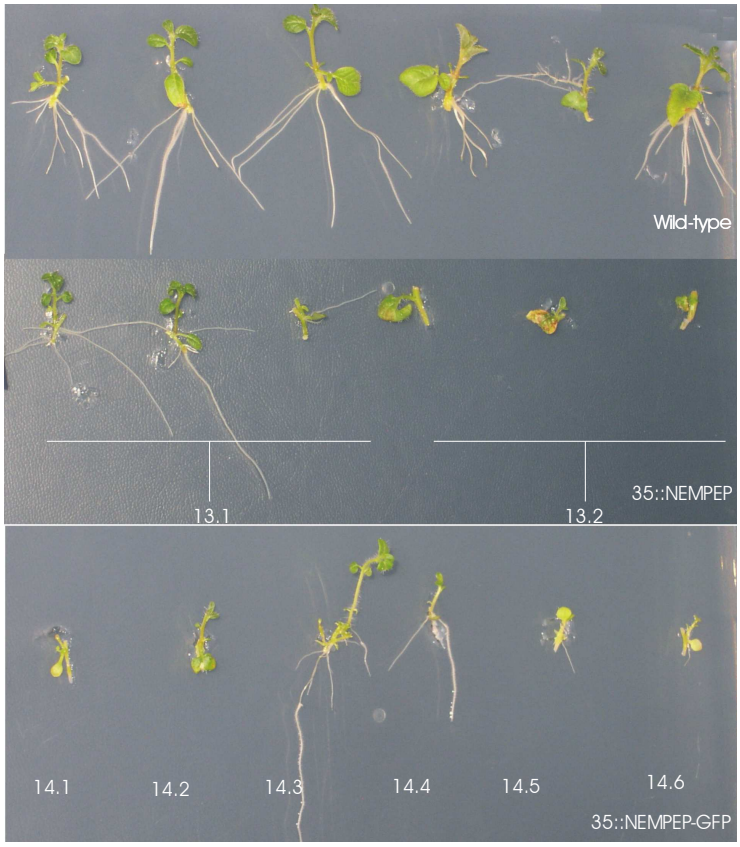


Figure 4: Phenotype of *Arabidopsis thaliana* plants transformed with a *NEMPEP* overexpression construct **a)-c)** Root system of three weeks old Arabidopsis seedlings: wild type Col-0 **(a)**, 35S::*NEMPEP* **(b)** and 35S::*GFP-NEMPEP* **(c)**. **d)** Rosette leaves of wild type Col-0 (left) and 35S::*GFP-NEMPEP* (line 16A; right) at the onset of flowering

Transformation with 35S::*GFP-NEMPEP* retards root growth in *Arabidopsis*

The single 35S::*NEMPEP* *Arabidopsis* transformant obtained showed no aberrant root or shoot phenotype (respectively Fig. 4b and data not shown), as compared to wild-type Col-0 or empty vector transformed plants (Table 2). The majority of the 15 *Arabidopsis* plants independently transformed with the 35S::*GFP-NEMPEP* construct seemed to show no aberrant phenotype, although four transformants had a smaller root system with shorter lateral roots (Fig. 4c and Table 2). One of these transformants (16A) seemed to lack lateral roots alltogether. The aerial parts of 10 of the 35S::*GFP-NEMPEP* plants were analysed under greenhouse conditions.



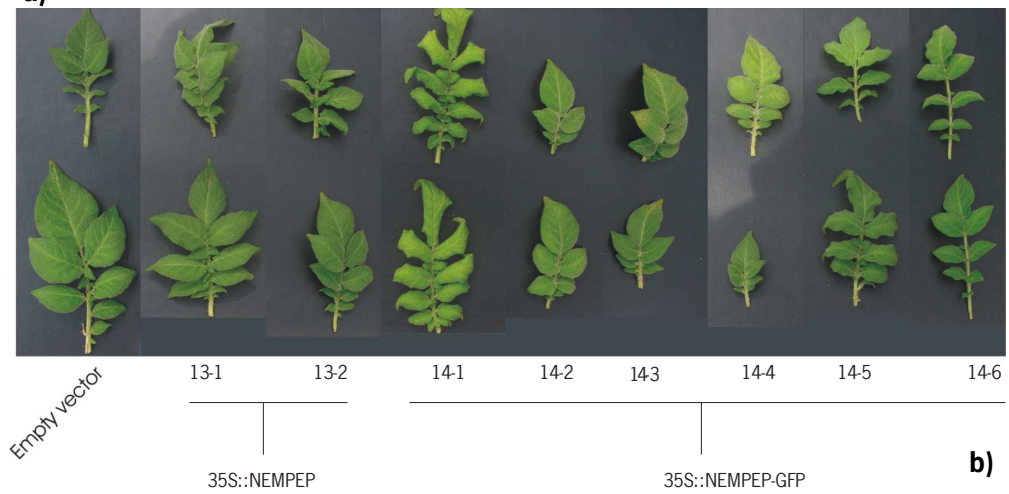


Figure 5: Phenotype of potato plants transformed *NEMPEP* over-expression constructs. **a)** Root sytem of two weeks old cuttings of wild type potato line V (top),35S::*NEMPEP* transformed plants (middle) and six independently transformed 35S:: *NEMPEP-GFP* plants (bottom) **b)** Overview of potato leaves transformed with respectively empty vector (EV),35S:: *NEMPEP* and 35S::*NEMPEP-GFP*. Most of the leaves transformed with a NEMPEP overexpression construct have shorter internodes and some have wrinkled leaves (eg. 14-1 and 14-5).

Interestingly, two of the transformants, 16A and 16D2, had more rosette leaves (respectively 18 and 12) at the onset of flowering (bolt) than wild type Col-0 plants (Fig. 4d and Table 2), which in our analysis had on average 8 rosette leaves. Transformant 16A also formed multiple inflorescences at the onset of flowering.

Table 2: Transformation efficiency, root development and shoot phenotype of *Arabidopsis* and potato plants upon transformation with NEMPEP overexpression constructs

		Empty vector	**35S::GFP-NEMPEP**	**35S::NEMPEP-GFP**	**35S::NEMPEP**
Arabidopsis thaliana	**Transformation efficiency**	0.06% (8/14365)	0.04% (15/36245)	NA	0.005% (1/17205)
	Retarded root development	0% (0/8)	27% (4/15)	NA	0% (0/1)
	Increase in number of rosette leaves[a]	13% (1/8)	20% (2/10)	NA	0% (0/1)
Potato line V	**Transformation efficiency**	4.9% (4/81)	NA	4.5% (6/132)	1.9% (2/104)
	Retarded root development[b]	NA	NA	17% (1/6)	100% (2/2)
	Altered shoot phenotype[c]	-	NA	83% (5/6)	100% (2/2)

[a] as compared to wild type Col-o plants. The number of rosette leaves was determined at the onset of bolting
[b] as compared to potato line V plants
[c] as compared to empty vector transformed plants
NA: Not analysed

Also the two potato plants independently transformed with the 35S::NEMPEP construct had a smaller root system than wild type potato line V plants (Table 2). One of the plants had shorter primary and lateral roots as compared to wild type plants, while the second line seemed to lack lateral roots (Fig. 5a). Also one of the six 35S::*GFP-NEMPEP* lines analysed (number 14-2), seemed to lack lateral roots (Fig. 5a and Table 2). The other five lines had a root system that looked comparable to what we observed for wild type plants (Fig. 5a and Table 2). The shoot of the two potato 35S::*NEMPEP* lines is shorter than observed for wild type line V or empty vector transformed plants (data not shown), and the leaves had shorter internodes (Fig. 5b). Shorter leaf internodes were also observed in five of the 35S::*NEMPEP-GFP* transformed lines (Fig.

5b). Two of these lines had wrinkled leaves (14-1 and 14-5)(Fig 5b), of which one showed some chlorosis (14-1).

Discussion

In this paper, we characterized the novel 4 kDa polypeptide NEMPEP from the potato cyst nematode. NEMPEP is predicted to be secreted from the dorsal esophageal gland, from the preparasitic juvenile to late parasitic stage of the nematode life-cycle. Based on these characteristics, it is likely that NEMPEP acts as an effector peptide.

NEMPEP is a positively charged, hydrophilic peptide, containing a γ-core structural G-X-C motif found in a wide range of host defence effector polypeptides showing antimicrobial activity (Yeaman & Yount, 2007, Yount & Yeaman, 2004). Up to now, no function has been ascribed to the γ-core G-X-C motif, although a function has been suggested in protecting a part of the primary amino-acid sequence from hyperdivergence (Yeaman & Yount, 2007). A possible antimicrobial activity of NEMPEP could protect the host plant from further secondary, opportunistic microbes upon nematode infection. However, the hydrophilicity of NEMPEP makes direct interaction with hydrophobic microbial membranes unlikely. This and the remarkable developmental expression pattern of *NEMPEP*, which shows absence of expression in migratory J2 juveniles, argue against a role as antimicrobial peptide, because antimicrobial activity is of particular importance in the migratory phase of parasitism.

NEMPEP is highly expressed in preparasitic J2 juveniles and in later J3 and J4 parasitic juveniles. The timing of host defense responses against nematodes suggests that nematode effectors are recognized early in the interaction (Cabrera Poch *et al.*, 2006, Grundler *et al.*, 1997). NEMPEP may be a cognate effector for a plant disease resistance protein, and its non-expression in migratory juveniles may evade recognition in the early stages of parasitism. At later stages, the nematode may have achieved sufficient suppression of host defense, which allows further upregulation of *NEMPEP* again.

Upon secretion, NEMPEP accumulates in the nucleolus of the plant cell. No consensus nucleolar targeting signals have been described, but stretches of basic amino-acids and binding to nucleolar factors seem to be the major mechanisms of nucleolar localization (Jans & Hübner, 1996, Misteli, 2005). The NEMPEP mature peptide sequence is characterized by two stretches of basic residues, one located at the ultimate C-terminus and one located near the N-terminus of the peptide. The most N-terminal stretch of basic amino-acids partly overlaps with a putative cAMP dependent protein kinase phosphorylation site (KKKS). Phosphorylation plays a central role in nuclear import and nuclear accumulation of proteins (Jans & Hübner, 1996), and phosphorylation sites close to nuclear localization signals often modify nuclear transport (Birbach *et al.*, 2004). We observed increased nucleolar accumulation of NEMPEP-GFP as compared to fusion of GFP at the NEMPEP N-terminus (GFP-NEMPEP). Interference of GFP with phosphorylation of NEMPEP may hamper nucleolar accumulation of GFP-NEMPEP. This hypothesis is supported by the observed size difference of about 2 kDa between NEMPEP-GFP and GFP-NEMPEP isolated from tobacco leaves. A similar phosphorylation related mobility shift on SDS-PAGE was reported for the mammalian membrane protein phrogrin (Wasmeier & Hutton, 2001).

Nucleoli are build around multiple copies of ribosomal DNA (rDNA), and are primarily known for their role in ribosomal gene expression and in ribosome assembly (Olson *et al.*, 2002). The nematode feeding cell is a highly metabolically active cell (Gheysen *et al.*, 1997). This activity may require enhanced ribosome supply, in which NEMPEP may play a functional role. New roles have recently been attributed to the nucleolus, eg. in cell stress sensing and signaling, in viral infection and in cell cycle control (reviewed in (Dundr & Misteli, 2002, Raŝka *et al.*, 2006, Olson et al., 2002, Pederson, 1998, Boisvert *et al.*, 2007) and(Hiscox, 2002). The nucleolus serves as a sequestration site for cell cycle regulators, like the cyclin dependent kinase cdc14 in yeast and retinoblastoma (Rb) protein in humans (Shou *et al.*, 1999, Takemura *et al.*, 2002), and in that way contributes to cell cycle progression.

Several cyst nematode effectors have been shown to localize to the nucleolus of plant cells, eg. SPRYSEC19 which is a member of the *SPRYSEC* gene family of *G. rostochiensis* (Rehman, 2008), the C-terminal domain of the ubiquitin extension protein Hs-Ubi1 from the beet cyst nematode *Heterodera schachtii* (Tytgat *et al.*, 2004), the pioneer parasitism gene 6E07 from the soybean cyst nematode (Elling *et al.*, 2007) and several members of the *SECPEP* family of γ-core peptides from *G. rostochiensis* (van Bers et al., chapter 2). Syncytium formation is accompanied by reactivation of the cell cycle resulting in a cell with enlarged nuclei. Interestingly, the fraction of small proteins (<3 kDa) present in the secretions of *Globodera rostochiensis* have been shown to co-stimulate cell proliferation in the presence of auxin and cytokinin (Goverse et al., 2000a, Goverse *et al.*, 1999). We are currently investigating if NEMPEP alters the promoter activity of a subset of cell cycle genes.

Plant transformation with a *NEMPEP* overexpression construct resulted in reduced primary root elongation and lateral root formation in the two *35S::NEMPEP* potato transformants, in twenty percent of the 35S::*NEMPEP-GFP* transformed potato lines and in twenty-five percent of the *35S::GFP-NEMPEP* transformed plants. Additionally, nearly all the potato transformants (88%) seemed to have shorter internodes between the leaflets of their compound leaves. Interestingly, overexpression of the cytokinin biosynthetic isopentenyltransferase (*IPT*) gene from *Agrobacterium tumefaciens* in potato resulted in increased endogenous cytokinin levels and inhibited primary root elongation, lateral root formation and reduced internode size (Ivana *et al.*, 1997). Also *Arabidopsis* plants overexpressing *PGA22*, an *IPT* homolog from Arabidopsis, had shorter roots (Sun *et al.*, 2003), and plants overexpressing the *Arabidopsis SOB5* gene, involved in cytokinin biosynthesis, showed reduced responsiveness to exogenous cytokinin, had an underdeveloped root system, a longer life cycle and curled down, smaller dark green leaves (Zhang *et al.*, 2006). In the near future we will investigate if the expression level of *NEMPEP* correlates with the severity of the phenotype observed. Additionally, the phenotypic analysis described here will be

repeated with fresh cuttings of the potato transformants, and with the T2 generation of the *Arabidopsis* plants. This is important in order to exclude phenotypic effects as a result of the transformation procedure, rather than expression of the transgene. However, none of the empty vector control plants showed reduction in root growth, and an altered shoot phenotype was not observed in potato control plants, and in only one of the *Arabidopsis* empty vector control plants (Table 2).

We hypothesize that NEMPEP interferes with cytokinin signaling, resulting in a cytokinin overproducing related phenotype in some of the *NEMPEP* overexpressing plants. Possible interference with cytokinin signaling is supported by the observed reduction in transformation efficiency upon expression of 35S*::NEMPEP* in potato and *Arabidopsis*. The observed reduction of shoot formation in callus tissue in potato may be due to a decreased perception of exogenous cytokinin or to decreased cytokinin signaling, as *Arabidopsis* mutants with enhanced levels of endogenous cytokinin are able to form shoots from callus tissue in the absence of exogenous cytokinin (Sun et al., 2003). A phenotype suggesting a reduced cytokinin response during the regeneration procedure after transformation may seem contradictory to the phenotype linked to enhanced cytokinin signaling observed in root growth and internode elongation of the NEMPEP transformed plants. However, a model for cytokinin signaling suggests a bell-shape dose response curve, in which upon crossing a certain treshold level an opposing effect is observed, eg. a very strong reduction in cytokinin signaling would result in root growth inhibition, even as a slight increase in cytokinin signaling (Ferreira & Kieber, 2005).

In *Arabidopsis*, transformation efficiency was measured as the percentage of hygromycin resistant seeds. The observed decrease in transformation efficiency of *35S::NEMPEP* transformed *Arabidopsis* plants may be attributed to embryo lethality caused by the transgene, to a role of the transgene in seed germination, or to inhibition in the release of seed dormancy. The plant hormones abscisic acid (ABA) and gibberellins are probably the most well known players in controlling seed dormancy

release and germination, although cross-talk of these hormones with cytokinin has been proposed to play a role in seed germination in *Sorghum* (reviewed by Kucera *et al.*, 2005).

Additionally, triple cytokinin receptor mutants of *Arabidopsis* showed that a decrease in cytokinin signaling resulted in more rapid seed germination, supporting the suggested functional role of cytokinin in seed germination (Riefler *et al.*, 2006). As such, an increase in cytokinin signaling in *35S::NEMPEP* overexpressing plants may severely hamper seed germination, resulting in a decreased percentage of hygromycin resistant seeds. The constructs resulting in expression of NEMPEP as a fusion protein with GFP did not result in a decreased transformation efficiency, which may indicate interference of GFP with NEMPEP functionality.

Cyst nematodes were shown to secrete low levels of cytokinin (Meutter *et al.*, 2003), which was suggested to play a role in manipulating host defense. On the other hand, transcript profiling of nematode feeding cells in *Arabidopsis* showed upregulation of ARR9, which is a negative regulator of cytokinin response, and downregulation of histidine kinases that function as cytokinin receptors (Ithal *et al.*, 2007). This effect was prominent early in feeding cell formation (2 dpi in this study) and was less prominent or even reversed in later stages (eg. downregulation of ARR9 at 5dpi). Interestingly, NEMPEP expression was hardly detectable in parasitic J2 juveniles, responsible for early feeding cell formation, so NEMPEP may be involved in increasing cytokinin signaling later in feeding cell formation.

In this paper we showed that the novel cyst nematode effector NEMPEP accumulates in the nucleolus of tobacco cells upon fusion with GFP. A N-terminally located phosphorylation site may be involved in this accumulation. Additionally, we show that transformation of *Arabidopsis* and potato with a NEMPEP overexpression construct (either with or without GFP) results in inhibition of root formation. We hypothesize a funtional role for NEMPEP in interference with cytokinin signaling involved in feeding cell formation. Complementation studies with *Arabidopsis* mutants showing

altered cytokinin responses may shed more light on this. In the near future, the identification of NEMPEP interactors will further guide the unraveling of the functional role of NEMPEP in parasitism. Furthermore, we are currently investigating if NEMPEP exerts antimicrobial activity.

References

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

Birbach, A., Bailey, S. T., Ghosh, S. and Schmid, J. A. (2004) Cystosolic, nuclear and nucleolar localization signals determine subcellular distribution and activity of the NF-?B inducing kinase NIK. *Journal of Cell Science,* **117,** 3615-3624.

Boisvert, F. M., Van Koningsbruggen, S., Navascue?s, J. and Lamond, A. I. (2007) The multifunctional nucleolus. *Nature Reviews Molecular Cell Biology,* **8,** 574-585.

Boman, H. G. (2003) Antibacterial peptides: Basic facts and emerging concepts. *Journal of Internal Medicine,* **254,** 197-215.

Cabrera Poch, H. L., Manzanilla López, R. H. and Kanyuka, K. (2006) Functionality of resistance gene Hero, which controls plant root-infecting potato cyst nematodes, in leaves of tomato. *Plant, Cell and Environment,* **29,** 1372-1378.

Ceroni, A., Passerini, A., Vullo, A. and Frasconi, P. (2006) Disulfind: A disulfide bonding state and cysteine connectivity prediction server. *Nucleic Acids Research,* **34**.

Cheng, J., Saigo, H. and Baldi, P. (2006) Large-scale prediction of disulphide bridges using kernel methods, two-dimensional recursive neural networks, and weighted graph matching. *Proteins: Structure, Function and Genetics,* **62,** 617-629.

Clough, S. J. and Bent, A. F. (1998) Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant Journal,* **16,** 735-743.

Curran, J., Baillie, D. L. and Webster, J. M. (1985) Use of genomic DNA restriction fragment length differences to identify nematode species. *Parasitology,* **90,**

137-144.

Curtis, M. D. and Grossniklaus, U. (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol,* **133,** 462-469.

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.

de Boer, J. M., Smant, G., Goverse, A., Davis, E. L., Overmars, H. A., Pomp, H.*, et al.* (1996) Secretory granule proteins from the subventral esophageal glands of the potato cyst nematode identified by monoclonal antibodies to a protein fraction from second-stage juveniles. *Mol Plant Microbe Interact,* **9,** 39-46.

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

Doncaster, C. C. and Seymour, M. K. (1973) Exploration and selection of penetration site by tylenchida. *Nematologica,* **19,** 137-145.

Dundr, M. and Misteli, T. (2002) Nucleolomics: An inventory of the nucleolus. *Molecular Cell,* **9,** 5-7.

Elling, A. A., Davis, E. L., Hussey, R. S. and Baum, T. J. (2007) Active uptake of cyst nematode parasitism proteins into the plant cell nucleus. *International Journal for Parasitology,* **37,** 1269-1279.

Endo, B. Y. (1964) Penetration and development of Heterodera glycines in soybean roots and related anatomical changes. *Phytopathology,* **54,** 79-88.

Ferreira, F. J. and Kieber, J. J. (2005) Cytokinin signaling. *Current Opinion in Plant Biology,* **8,** 518-525.

Gheysen, G., Engler, J. D. A. and Montagu, M. V. (1997) Cell cycle regulation in nematode feeding sites. In: *Cellular and molecular aspects of plant-nematode interactions.* (Fenoll, C., Grundler, F. M. W. and Ohl, S. A., ed.^eds.). Kluwer Academic Publishers, pp. 120-132.

Gheysen, G. and Jones, J. T. (2006) Molecular Aspects of Plant-Nematode Interactions. In: *Plant Nematology.* (Perry, R. N. and Moens, M., ed.^eds.). Wallingford: CABI, pp. 234-254.

Golinowski, W., Sobczak, M., Kurek, W. and Grymaszewska, G. (1997) The structure of syncytia. In: *Cellular and molecular aspects of plant-nematode interactions.* (Fenoll, C., Grundler, F. M. W. and Ohl, S. A., ed.^eds.). Kluwer Academic Publishers, pp. 80-97.

Goverse, A., de Engler, J. A., Verhees, J., van der Krol, S., Helder, J. H. and Gheysen, G. (2000a) Cell cycle activation by plant parasitic nematodes. *Plant Mol Biol,* **43,** 747-761.

Goverse, A., Overmars, H., Engelbertink, J., Schots, A., Bakker, J. and Helder, J. (2000b) Both induction and morphogenesis of cyst nematode feeding cells are mediated by auxin. *Mol Plant Microbe Interact,* **13,** 1121-1129.

Goverse, A., Rouppe van der Voort, J., Roppe van der Voort, C., Kavelaars, A., Smant, G., Schots, A.*, et al.* (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.

Grundler, F. M. W. and Böckenhoff, A. (1997) Physiology of nematode feeding and feeding sites. In: *Cellular and molecular aspects of plant-nematode interactions.* (*al.,* C. F. *e.*, ed.^eds.). Kluwer academic publishers, pp. 107-119.

Grundler, F. M. W., Sobczak, M. and Lange, S. (1997) Defence responses of Arabidopsis thaliana during invasion and feeding site induction by the plant-parasitic nematode Heterodera glycines. *Physiological and Molecular Plant Pathology,* **50,** 419-429.

Handwerger, K. E., Cordero, J. A. and Gall, J. G. (2005) Cajal bodies, nucleoli, and speckles in the Xenopus oocyte nucleus have a low-density, sponge-like

structure. *Molecular Biology of the Cell,* **16,** 202-211.

Hiscox, J. A. (2002) The nucleolus - A gateway to viral infection? *Archives of Virology,* **147,** 1077-1089.

Hoekema, A., Hirsch, P. R., Hooykaas, P. J. J. and Schilperoort, R. A. (1983) A binary plant vector strategy based on separation of vir- and T-region of the Agrobacterium tumefaciens Ti-plasmid. *Nature,* **303,** 179-180.

Horsman, K., Gavrilenko, T., Bergervoet, M., Huigen, D. J., Joe, A. T. W. and Jacobsen, E. (2001) Alteration of the genomic composition of Solanum nigrum (+) potato backcross derivatives by somatic hybridization: Selection of fusion hybrids by DNA measurements and GISH. *Plant Breeding,* **120,** 201-207.

Ithal, N., Recknor, J., Nettleton, D., Maier, T., Baum, T. J. and Mitchum, M. G. (2007) Developmental transcript profiling of cyst nematode feeding cells in soybean roots. *Molecular Plant-Microbe Interactions,* **20,** 510-525.

Ivana, M., Lidiya, S., Milos, O., Oksana, Z., Tatyana, K., Josef, E.*, et al.* (1997) Growth pattern, tuber formation and hormonal balance in in vitro potato plants carrying ipt gene. *Plant Growth Regulation,* **21,** 27-36.

Jans, D. A. and Hübner, S. (1996) Regulation of protein transport to the nucleus: Central role of phosphorylation. *Physiological Reviews,* **76,** 651-685.

Janssen, R., Bakker, J. and Gommers, F. J. (1990) Selection of virulent and avirulent lines of *Globodera rostochiensis* for the *H1* resistance gene in *Solanum tuberosum* ssp. *andigena* CPC 1673. *Revue de Nematologie,* **13,** 265-268.

Kalderon, D., Roberts, B. L., Richardson, W. D. and Smith, A. E. (1984) A short amino acid sequence able to specify nuclear location. *Cell,* **39,** 499-509.

Karczmarek, A., Overmars, H., Helder, J. and Goverse, A. (2004) Feeding cell development by cyst and root-knot nematodes involves a similar early, local and transient activation of a specific auxin-inducible promoter element. *Molecular Plant Pathology,* **5,** 343-346.

Kucera, B., Cohn, M. A. and Leubner-Metzger, G. (2005) Plant hormone interactions

during seed dormancy release and germination. *Seed Science Research,* **15,** 281-307.

Kudla, U., Milac, A. L., Qin, L., Overmars, H., Roze, E., Holterman, M.*, et al.* (2007) Structural and functional characterization of a novel, host penetration-related pectate lyase from the potato cyst nematode Globodera rostochiensis. *Molecular Plant Pathology,* **8,** 293-305.

Kyte, J. and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology,* **157,** 105-132.

Lee, D. L. (2002) Life cycles. In: *The Biology of Nematodes.* (Lee, D. L., ed.^eds.). London: Taylor and Francis, pp. 61-72.

Matsubayashi, Y. and Sakagami, Y. (2006) Peptide Hormones in Plants. *Annu Rev Plant Biol*.

Meutter, J. d., Tytgat, T., Witters, E., Gheysen, G., Onckelen, H. v. and Gheysen, G. (2003) Identification of cytokinins produced by the plant parasitic nematodes *Heterodera schachtii* and *Meloidogyne incognita*. *Molecular Plant Pathology,* **4,** 271-277.

Misteli, T. (2005) Going in GTP cycles in the nucleolus. *Journal of Cell Biology,* **168,** 177-178.

Nakai, K. and Horton, P. (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends in Biochemical Sciences,* **24,** 34-35.

Nakazawa, M. and Matsui, M. (2003) Selection of hygromycin-resistant Arabidopsis seedlings. *BioTechniques,* **34,** 28-30.

Niebel, A., de Almeida Engler, J., Hemerly, A., Ferreira, P., Inze, D., Van Montagu, M.*, et al.* (1996) Induction of cdc2a and cyc1At expression in Arabidopsis thaliana during early phases of nematode-induced feeding cell formation. *Plant J,* **10,** 1037-1043.

Olson, M. O. J., Hingorani, K. and Szebeni, A. (2002) Conventional and nonconventional

roles of the nucleolus. *International Review of Cytology,* **219,** 199-266.

Patel, N. and Goodman, C. (1992) *Nonradioactive In Situ Hybridization Application Manual*. Boehringer Mannheim Biochemica, Mannheim, Germany.

Pederson, T. (1998) The plurifunctional nucleolus. *Nucl. Acids Res.,* **26,** 3871-3876.

Qin, L., Overmars, H., Helder, J., Popeijus, H., van der Voort, J. R., Groenink, W.*, et al.* (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.

Raŝka, I., Shaw, P. J. and Cmarko, D. (2006) New Insights into Nucleolar Architecture and Activity. In: *International Review of Cytology.* (ed.^eds.). pp. 177-235.

Rehman, S., Qin, L., Overmars, H., Petrescu, A. J., Thomma, B. P. H. J., Helder, J.*, et al.* (2008) Overexpression of a nematode SPRYSEC-gene in potato promotes the virulence of a nematode, a fungus and a virus. *in PhD thesis, Wageningen University, the Netherlands***,** 74-107.

Riefler, M., Novak, O., Strnad, M. and Schmülling, T. (2006) Arabidopsis cytokinin receptors mutants reveal functions in shoot growth, leaf senescence, seed size, germination, root development, and cytokinin metabolism. *Plant Cell,* **18,** 40-54.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*. New York.

Sarosh, B. R. and Meijer, J. (2007) Transcriptional profiling by cDNA-AFLP reveals novel insights during methyl jasmonate, wounding and insect attack in Brassica napus. *Plant Molecular Biology,* **64,** 425-438.

Shen, W. J. and Forde, B. G. (1989) Efficient transformation of Agrobacterium spp. by high voltage electroporation. *Nucleic Acids Research,* **17,** 8385.

Shou, W., Seol, J. H., Shevchenko, A., Baskerville, C., Moazed, D., Susan Chen, Z. W.*, et al.* (1999) Exit from mitosis is triggered by Tem1-dependent release of the protein phosphatase Cdc14 from nucleolar RENT complex. *Cell,* **97,** 233-244.

Sigrist, C. J., Cerutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M., *et al.* (2002) PROSITE: a documented database using patterns and profiles as motif descriptors. *Briefings in bioinformatics,* **3,** 265-274.

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

Sun, J., Niu, Q. W., Tarkowski, P., Zheng, B., Tarkowska, D., Sandberg, G., *et al.* (2003) The Arabidopsis AtIPT8/PGA22 gene encodes an isopentenyl transferase that is involved in de novo cytokinin biosynthesis. *Plant Physiology,* **131,** 167-176.

Takemura, M., Ohoka, F., Perpelescu, M., Ogawa, M., Matsushita, H., Takaba, T., *et al.* (2002) Phosphorylation-dependent migration of retinoblastoma protein into the nucleolus triggered by binding to nucleophosmin/B23. *Experimental Cell Research,* **276,** 233-241.

Trivedi, S. and Arasu, P. (2005) Evaluation of endogenous reference genes for real-time PCR quantification of gene expression in Ancylostoma caninum. *Molecular and Biochemical Parasitology,* **143,** 241-244.

Turner, S. J. and Evans, K. (1998) The origins, global distribution and biology of potato cyst nematodes (*Globodera rostochiensis* (Woll.) and *Globodera pallida* Stone). In: *Potato cyst nematodes: biology, distribution and control.* (Marks, R. J. and Brodie, B. B., ed.^eds.). Wallingford: CAB International, pp. 7-26.

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

van Engelen, F. A., Schouten, A., Molthoff, J. W., Roosien, J., Salinas, J., Dirkse, W. G., *et al.* (1994) Coordinate expression of antibody subunit genes yields high levels of functional antibodies in roots of transgenic tobacco. *Plant Molecular*

Biology, **26,** 1701-1710.

von Mende, N., Gravato-Nobre, M. J. and Perry, R. N. (1998) Host finding, invasion and feeding. In: *The cyst nematodes.* (Sharma, S., ed.^eds.). Dordrecht: Kluwer Academic Publishers.

Wang, J. P. and Bughrara, S. S. (2007) Monitoring of gene expression profiles and identification of candidate genes involved in drought responses in Festuca mairei. *Molecular Genetics and Genomics,* **277,** 571-587.

Wang, X., Mitchum, M. G., Gao, B., Li, C., Diab, H., Baum, T. J.*, et al.* (2005) A parasitism gene from a plant-parasitic nematode with function similar to CLAVATA3/ESR (CLE) of Arabidopsis thaliana. *Molecular Plant Pathology,* **6,** 187-191.

Wasmeier, C. and Hutton, J. C. (2001) Secretagogue-dependent phosphorylation of the insulin granule membrane protein phogrin is mediated by cAMP-dependent protein kinase. *J Biol Chem,* **276,** 31919-31928.

Williamson, V. M. and Kumar, A. (2006) Nematode resistance in plants: the battle underground. *Trends in Genetics,* **22,** 396-403.

Wyss, U. (1992) Observations on the feeding behaviour of *Heterodera schachtii* throughout development, including events during moulting. *Fundam. appl. Nematol.,* **15,** 75-89.

Yeaman, M. R. and Yount, N. Y. (2007) Unifying themes in host defence effector polypeptides. *Nature Reviews Microbiology,* **5,** 727-740.

Yount, N. Y. and Yeaman, M. R. (2004) Multidimensional signatures in antimicrobial peptides. *Proceedings of the National Academy of Sciences of the United States of America,* **101,** 7363-7368.

Zhang, J., Wrage, E. L., Vankova, R., Malbeck, J. and Neff, M. M. (2006) Over-expression of SOB5 suggests the involvement of a novel plant protein in cytokinin-mediated development. *Plant Journal,* **46,** 834-848.

Zuncke, U. and Eisenback, J. (1998) Morphology and ultrastructure. In: *The cyst*

nematodes. (Sharma, S., ed.^eds.). Dordrecht: Kluwer Academic Publishers, pp. 31-56.

Chapter 4

Overexpression of the *SECPEP3* peptide from the potato cyst nematode *Globodera rostochiensis* alters development of potato and *Arabidopsis*

Nikkie van Bers, Hein Overmars, Jessica Nouws, Erik Slootweg, Aska Goverse, Jaap Bakker and Geert Smant

Abstract

Secreted nematode effectors are essential for host invasion and for the succesfull induction and maintance of a feeding cell. The *SECPEP* gene family of the potato cyst nematode *Globodera rostochiensis* codes for novel nematode effector peptides. Here we describe the *in-planta* expression of one of the members of the *SECPEP* gene family, *SECPEP3*, in order to provide insight into its functional role during parasitism. Translational fusion of SECPEP3 to GFP results in fluorescence in the nucleolus of tobacco epidermal cells, besides fluorescence in the nucleoplasm and in the cytoplasm. This is in contrast to free GFP, which localizes to the nucleoplasm and cytoplasm of plants cells, but is largely excluded from the nucleolus. Expression of *35S::SECPEP3-GFP* resulted in enhanced lateral root formation in potato hairy roots, while transformation with *35S::GFP-SECPEP3* or *35S::SECPEP3* reduced root growth and altered the morphology of leaves of potato and *Arabidopsis*. We provide evidence for the hypothesis that expression of *SECPEP3* interferes with cell cycle regulation in a way similar to cyclin-dependent kinase inhibitors, which may result in an alteration in ploidy levels due to endoreduplication cycles.

Introduction

The potato cyst nematode *Globodera rostochiensis* is an obligate biotrophic nematode, parasitizing on a limited range of Solanaceous plants, e.g. potato, tomato and egg plant (Evans & Stone, 1977). Second stage juveniles (J2) penetrate the plant root in the zone of elongation and migrate intracellularly towards the area of cell differentiation in order to find a cell suitable for the initiation of a feeding cell or syncytium. Procambial, pericycle, cortex or endodermal cells are preferably chosen as initial syncytium cells (ISCs) (Hussey & Grundler, 1998, Sobczak *et al.*, 1997, von Mende *et al.*, 1998).

Feeding site formation is accompanied by local cell wall dissolution which results in the partial fusion of neighbouring cells (Grundler *et al.*, 1998), and the large central vacuole of the ISC is replaced by numerous small vacuoles (Golinowski *et al.*, 1996). Additionally, the number of organelles increases and the multiple nuclei enlarge (Endo, 1964), due to a process called endoreduplication. The endoreduplication cycle, which is common in both animals and plants, is a cell cycle mode in which iterative rounds of DNA replication take place in the absence of subsequent mitosis and cytokinesis (Inze, 2005).

A “normal” cell cycle is characterized by a round of DNA replication (S-phase) followed by mitosis and cytokinesis (M phase). The S and M phase are separated by two gap phases (G1 and G2) during which cells monitor whether all conditions are favourable to proceed replication or mitosis. Phosphorylation by kinase complexes, which in minimal configuration consist of a kinase (cyclin dependent kinase or CDK) and a regulatory cyclin subunit, play a central role in cell cycle progression (reviewed by (Inze, 2005)). Proteolysis and cyclin-dependent kinase inhibitors (ICKs) are the major regulators of CDK activity. ICKs prevent the formation of a complex between cyclin-dependent kinase A (CDKA) and cyclin D (CYCD)((De Veylder *et al.*, 2001)). Cell cycle activation is essential for feeding site formation, which is indicated by the observation that blockage of DNA synthesis severely hampered feeding cell formation (de Almeida Engler *et al.*, 1999). Additionally, CDKA;1 (formerly known as cdc2a) and CYCB1;1

(formerly known as cyc1) promoter activity was observed inside and around young syncytia (Niebel *et al.*, 1996).

Previously, we showed that small proteins present in the secretions of the potato cyst nematode co-stimulate proliferation of tobacco protoplasts and human blood cells. The large dorsal and subventral esophageal glands are major sites for production of secreted nematode effectors essential for feeding cell formation. The *SECPEP* gene family consists of nine members, *SECPEP1*—SECPEP9, that all code for positively charged nematode effector peptides. The members of the *SECPEP* gene family do not show significant sequence similarity to sequences outside *Globodera rostochiensis* or to sequences to which a function has been assigned (*van Bers et al.,* chapter 2). However, all members of the *SECPEP* gene family share characteristics with several classes of peptides involved in defense, e.g. antimicrobial peptides found in the skin of ranid frogs, toxins isolated from scorpion venom and effector proteins recognized by plant disease resistance proteins (van Bers *et al.,* chapter 2). These characteristics are their overall positive charge at pH 7 and an extremely high level of sequence diversity, which is largely restricted to the mature peptide encoding region. Additionally, nearly all the SECPEPs harbour a γ-core C-X-G motif found in several classes of host defense peptides (Yeaman & Yount, 2007, Yount & Yeaman, 2004).

Currently, there are two models regarding the functional role of the SECPEP peptides in parasitism (van Bers *et al.,* chapter 2). The first model regards a role as defense molecules in protecting the host plants against opportunistic microbes upon nematode infection, and the second one concerns a role as peptide hormones interfering with signaling in plants. As such, the SECPEP peptides may be involved in cell-cycle reactivation occurring upon syncytium formation. SECPEP3 is a hydrophilic peptide (van Bers *et al.*, chapter 2), which makes a direct interaction with hydrophobic microbial membranes less likely and may indicate that SECPEP3 functions as a peptide hormone. In this paper, we provide the first insight into the functional role of SECPEP3. As a first step, we analysed the subcellular localization of SECPEP3 (*in-planta*) upon

overexpression as a fusion to GFP. Subsequently, we investigated the effect of expression of *SECPEP3* in potato hairy roots, and on development of potato and *Arabidopsis* plants.

Material and methods

Sequence analysis

The computer algorithm SignalP 3.0 (Bendtsen *et al.*, 2004)) was used to predict the presence of a signal peptide for secretion and the corresponding putative cleavage site (Neural Networks) in the SECPEP3 protein sequence. Blast at NCBI (www.ncbi.nlm.nih.gov) or at nemaBLAST (www.nematode.net) was used to search for matching sequences in the database. PSORT (Nakai & Horton, 1999)) was used to predict the subcellular localization of SECPEP3, and Prosite was used to predict putative phosphorylation sites in the SECPEP3 peptide sequence (Sigrist *et al.*, 2002). ProtParam (ExPASY web server) was used for the calculation of the Grand Average of Hydropathy (GRAVY) score.

In-planta localization studies, protein isolation and immuno-blotting

The mature peptide encoding sequence of SECPEP3 was PCR amplified from cDNA isolated from preparasitic *Globodera rostochiensis* juveniles, using primers incorporating *AscI* and *PacI* restriction sites (Table 1). The resulting PCR product was cloned into pCR4-TOPO (Invitrogen), which was subsequently digested with the enzymes *AscI* and *PacI.* The digested fragment was gel-purified and cloned into a *AscI/PacI* digested pMDC vector (Table 1)(Curtis & Grossniklaus, 2003). Error-free clones were confirmed by sequence analysis.

The constructs were transformed into *Agrobacterium tumefaciens*, strain LBA4404 (Hoekema *et al.*, 1983)) by electroporation (Shen & Forde, 1989)). Transient expression in *Nicotiana benthamiana* epidermal cells was performed as described previously by

Kudla *et al.* (2007)(Kudla *et al.*, 2007)). A Zeiss LSM 510 confocal microscope was used to visualize GFP.

Table 1: Primers used for the generation over SECPEP3 overexpression constructs

Construct	Primer 1(5'→3')[a]	Primer 2(5'→3')[a]	Vector
35S::SECPEP3	GGCGCGCCAAATATGtcTggctgcggtggcg gtgatgg	TTAATTAAtcagtaaccagaCcttgg gat	pMDC32
35S::GFP-SECPEP3	GGCGCGCCAAATATGtcTggctgcggtggcg gtgatgg	TTAATTAAtcagtaaccagaCcttgg gat	pMDC45
35S::SECPEP3-GFP	TTAATTAAAATATGtcTggctgcggtggcggt gatgg	GGCGCGCCAgtaaccagaccttgg gat	pMDC201
35S::GFP	GGCGCGCCAAATATGagtaaaggagaagaa cttttc	TTAATTAAtcatttgtatagttcatccat gcc	pMDC32

[a] The restriction sites, kozak sequence, start codon and nucleotide mutations inserted to avoid primer dimerization are depicted in capitals.

For protein isolation, leaves were harvested three days after infiltration. The leaf material was grinded in liquid nitrogen and transferred to a pre-cooled tube. Protein was extracted by vigorously mixing the tissue with 8 M ureum (100 µl/ 23 mg tissue) and protein sample buffer (6X stock contains 0.35 M Tris-HCL, 10.3% SDS (w/v), 36% glycerol, 5% ß-mercapto-EtOH and 0.012% bromophenol blue) in the presence of the serine protease inhibitor pefabloc SC Plus (Roche diagnostics). The protein sample was boiled for ten minutes, and the supernatant was separated on 12% NuPAGE NOVEX Bis-Tris gels (Invitrogen) in MES-buffer (Invitrogen) and subsequently transferred to PVDF membrane (Millipore) as described by (de Boer *et al.*, 1996)). Membranes were blocked in 10% BSA in PBS supplied with 0.1% Tween20 (PBS-T) and incubated with 1:5000 rabbit polyclonal GFP antibody (Abcam) in PBS-T containing 3% BSA.

Expression of SECPEP3 in potato hairy roots

The pMDC201-*SECPEP*3 plasmid (*35S::SECPEP3-GFP*) and the empty vector control pMDC201 (Curtis & Grossniklaus, 2003) were transferred into *Agrobacterium*

rhizogenes ATCC 15834 by electroporation (Shen & Forde, 1989)) and transformed into potato (*Solanum tuberosum* 6487-9) using stem piece transformation. Briefly, 0.5-1 cm stem pieces were incubated for four days on CIM medium (MS20 supplemented with 0.5g/l MES, 0.2 mg/l 2,4D and 0.2 mg/l kinetin)(Visser *et al.*, 1989). After four days, inoculation was performed with liquid *A. rhizogenes* cultures followed by a 3 day incubation period on SM agar plates (MS20 supplemented with 0.5g/l MES and gelrite agar (Duchefa, the Netherlands)). The stem pieces were washed with liquid SM medium supplemented with 500 mg/l carbenicillin to eliminate *A. rhizogenes*, and roots formed on solidified SM medium (supplemented with 500 mg/l carbenicillin and 100 mg/l vancomycin) were selected for transgenicity by transfer to selective medium supplemented with 30 mg/l hygromycin, 500 mg/l carbenicillin and 100 mg/l vancomycin.

Generation of transgenic plants and phenotypic analysis

Stem segments (0.5-1 cm) from *in-vitro* grown potato plants (*Solanum tuberosum* 6487-9) (Horsman *et al.*, 2001) were used for *Agrobacterium tumefaciens* mediated plant transformation with pMDC32::*SECPEP3*, pMDC45::*SECPEP3* and pMDC201::*SECPEP3* (van Engelen *et al.*, 1994). For construct preparation and *Agrobacterium* transformation, see above. One cutting of each transgenic line was grown on MS20 medium and used for phenotypic analysis of the root morphology. After analysis, the same individual plants were transferred to soil, and the leaf morphology was monitored under greenhouse conditions.

Arabidopsis thaliana was transformed with *Agrobacterium tumefaciens* harbouring the pMDC32::*SECPEP3*, pMDC32::*GFP*, pMDC201::*SECPEP3* or pMDC45::*SECPEP3* construct by use of the floral dip method as described by (Clough & Bent, 1998). Transformed seeds were surface sterilized as described by (Karczmarek *et al.*, 2004) and selected for transgenicity by germination on basic agar medium containing hygromycin (30 mg/L) as described by (Nakazawa & Matsui, 2003). The first

generation of transformed seeds (T1) was used for phenotypic analysis. For phenotypic analysis of root growth, freshly germinated seedlings were transferred to KNOP medium (Sijmons *et al.*, 1991). After analysis of the root phenotype, the plants were transferred to pots and leaf morphology was studied under greenhouse conditions.

Results

Fusion of SECPEP3 to GFP induces a shift in fluorescence towards the nucleolus

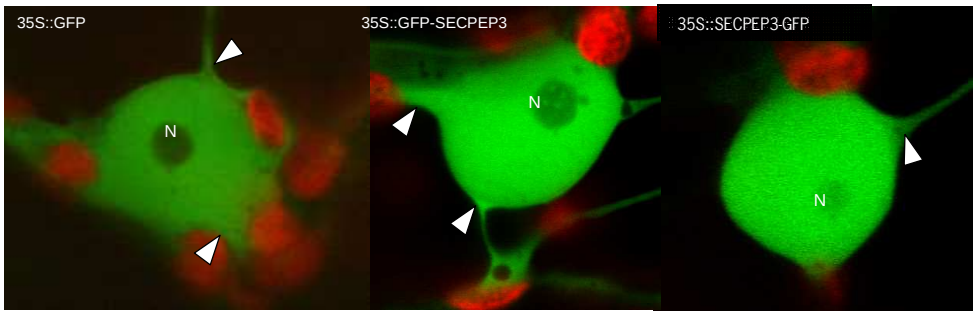
SECPEP3 encodes a 3.2 kDa peptide which is preceded by a cleavable, N-terminal signal peptide for secretion (van Bers *et al.,* chapter 2). The mature SECPEP3 peptide is rich in the positively charged amino-acids arginine and lysine (Fig. 1a), and has a positive net charge of +4 at pH7. PSORTII (Nakai & Horton, 1999) predicts a nuclear localisation for the mature SECPEP3 peptide, although no known nuclear localisation signal is found within the SECPEP3 sequence.

To test the predicted nuclear localisation of SECPEP3, we created 35S promoter-driven translational fusion constructs of a cDNA sequence encoding the mature SECPEP3 peptide with the GFP protein coding sequence. GFP was fused to either the N-terminus (*35S::GFP-SECPEP3*) or the C-terminus of SECPEP3 (*35S::SECPEP3-GFP*). Unfused GFP expressed under control of the CaMV 35S promoter was used as a control (*35S::GFP*).The constructs were transformed into *Agrobacterium tumefaciens* and used for overexpression in *Nicotiana benthamiana* leaves by use of agroinfiltration. Unfused GFP localizes to the nucleus and cytoplasm of tobacco cells (Fig. 1b). Nuclear localisation of GFP can be attributed to its small size (26 kDa), which allows free diffusion through the nuclear pore complex into the nucleus (Jans & Hübner, 1996).

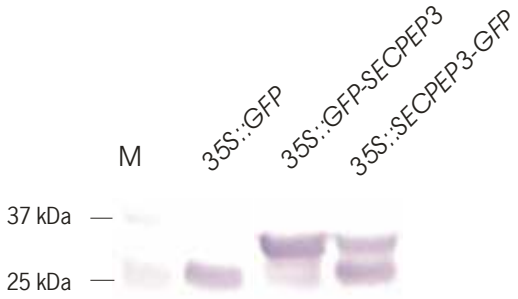
Nucleolar fluorescence was not observed in leaves infiltrated with the *35S::GFP* construct (Fig. 1b). On the other hand, nucleolar fluorescence, as well as nuclear and cytoplasmic fluorescence, was observed in leaves infiltrated with *35S::GFP-SECPEP3*

(Fig. 1b). Some nucleolar fluorescence was also observed in leaves infiltrated with the *35S::SECPEP3-GFP* construct, however, this level was lower than observed for the N-terminal GFP fusion construct (Fig. 1b).

a)



b)



c)

Figure 1a) The peptide sequence of SECPEP3. The first 19 amino-acids encode the cleavable signal peptide for secretion (underlined). The mature peptide sequence (grey font) contains a γ-core C-X-G motif (boxed) and a putative protein kinase C phosphorylation site (arrows) **b)** Tissue localization of SECPEP3-GFP by confocal laser scanning microscopy of a tobacco epidermal leaf cell transiently transformed with a plasmid harbouring respectively the *35S:: GFP* (left), *35S::GFP-SECPEP3* (middle) or *35S::SECPEP3-GFP* (right) sequence. Fluorescence of GFP is observed in the nucleus and in the cytoplasm of tobacco cells (indicated by arrowheads), while the SECPEP3-GFP fusion protein localizes to the nucleolus (N) as well. **c)** Immunoblot analysis of blotted protein extracted from leaves infiltrated with 35S::GFP (left), 35S::GFP-SECPEP3 and 35S::SECPEP3-GFP (right) detected with an anti-GFP antibody. The molecular size (in kDa) of the corresponding marker bands (M) is depicted at the left side.

Fusion of GFP to the C-terminus of SECPEP3 induces proteolytic processing

In order to confirm the presence of the GFP-SECPEP3 and SECPEP3-GFP fusion proteins, we performed western blot analysis and used an antibody recognizing GFP for subsequent immunodetection. For the protein fraction isolated from *35S::GFP-SECPEP3* infiltrated leaves we observed a single band on the western blot. This band corresponds to proteins migrating at a level expected for the GFP-SECPEP3 fusion protein (29 kDa)(Fig. 1c). For the protein sample isolated from *35S::SECPEP3-GFP* infiltrated leaves we detected two bands on the membrane after immuno-detection with α-GFP (Fig. 1c). The slowest migrating band seems to migrate at the same level as the band observed for proteins from the *35S::GFP-SECPEP3* infiltrated leaf sample (29 kDa), while the faster migrating band migrates slightly slower than free GFP (~27 kDa).

SECPEP3 shares two sequence motifs with *Arabidopsis* CDK inhibitors

SECPEP3 does not share significant sequence similarity to anything with a known function present in sequence databases. However, sequence analysis of the 31 amino-acid sequence of the mature SECPEP3 peptide reveals that SECPEP3 shares a five amino-acid motif, KEKPL, with the *Arabidopsis* cyclin-dependent kinase inhibitor ICK1/KRP1 (Fig. 2a). ICK1/KRP1 is essential for cell cycle control, and interacts with CDKA;1, CYCD1, CYCD2 and CYCD3 in *Arabidopsis* (Wang *et al.*, 1998, Zhou *et al.*, 2002, Zhou *et al.*, 2003b). A similar motif, KEKP, is found in another *Arabidopsis* protein MGOUN3/BRUSHY1/TONSOKU, and this motif is also found in the close relatives of SECPEP3, SECPEP6 and SECPEP7. SECPEP 6 and 7 also share an additional motif, (K)PRPP(P), which is positioned in reverse orientation in SECPEP 6 and 7, with three other proteins that are believed to function as cyclin dependent kinase inhibitors: the mammalian $p27^{kip}$ protein and the *Arabidopsis* proteins SIAMESE (SIM) and Smr1 (Fig. 2c). The (K)PRPP(P) motif is not found in SECPEP3, but SECPEP3 shares a three amino-acid motif (KIP) with SIM and Smr1 (Fig. 2b). Neither the KEKPL

motif of ICK1/KRP1 and SECPEP3, nor the (K)PRPP(P) or KIP motif of SIM/SMR1 show obvious similarity to a domain or motif with a known function.

Expression of *35S::SECPEP3-GFP* increases lateral root formation in potato hairy roots

As a first assay to investigate the functional role of SECPEP3 in nematode feeding site formation, the mature SECPEP3 peptide sequence was expressed as a C-terminal fusion with GFP (*35S::SECPEP3-GFP*) in potato hairy roots. Expression of *35S::SECPEP3-GFP* resulted in more and longer lateral roots in thirty-eight percent of the transformed lines (Table 2), as compared to roots transformed with an empty vector construct. Additionally, one of the transformed lines showed callus formation, which was restricted to the distal part of the primary root (line 2.1, Fig. 3a). In order to confirm the presence of the SECPEP3-GFP fusion protein, we performed western blot analysis on protein isolated from two of the hairy root lines showing enhanced lateral root formation. Immuno-detection with an antibody recognizing GFP resulted in two bands on a western blot (Fig. 3b). The lower band corresponds to a protein fraction migrating at a level corresponding to unfused GFP (~26 kDa). The other band observed migrates at a slightly slower pace, and comparison with co-migrating protein size marker bands indicates that this fraction corresponds to a proteins with a molecular weight of ~30 kDa. The expected molecular mass of the SECPEP3-GFP fusion protein is 29 kDa. Of the eight independently transformed hairy root lines examined, line 2.1 shows the most prominent lateral root formation, and this line also shows callus formation on the primary root. Hairy root line 6.1 also shows enhanced lateral root formation, as compared to empty vector transformed hairy roots, but lateral root formation was less excessive than observed for line 2.1 (Fig. 3a), and no callus formation was observed. Western blot analysis showed that in line 2.1 the 26 kDa protein fraction is more prominent than the 30 kDa protein fraction, while the 30 kDa protein fraction is most prominent in line 6.1 (Fig. 3b).

Table 2: Expression of SECPEP3-GFP in potato hairy roots

	Empty vector	35S::GFP	35S::GFP-SECPEP3
Increase in lateral root formation*	- (-/6)	0% (0/2)	38% (3/8)

*as compared to empty vector control

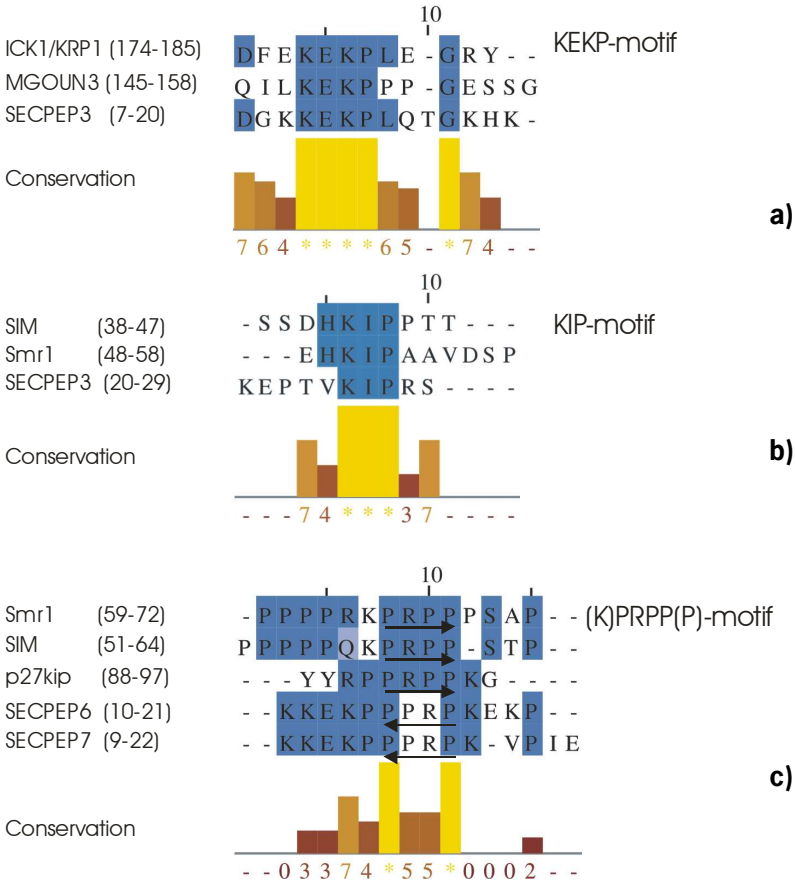


Figure 2: Sequence motifs found in the SECPEPs **a)** SECPEP3, SECPEP6 and SECPEP7 share a KEKP(L) motif with ICK1/KRP1 and with MGOUN3 **b)** a KIP motif is shared between SECPEP3 and the cyclin-dependent kinase inhibitors SIM and Smr1 and **c)** SECPEP6 and SECPEP7 share a (K)PRPP(P) motif with the cyclin-dependent kinase inhibitors Smr1, SIM and p27kip. The motif is found in C- to N-terminal orientation in SECPEP6 and SECPEP7.

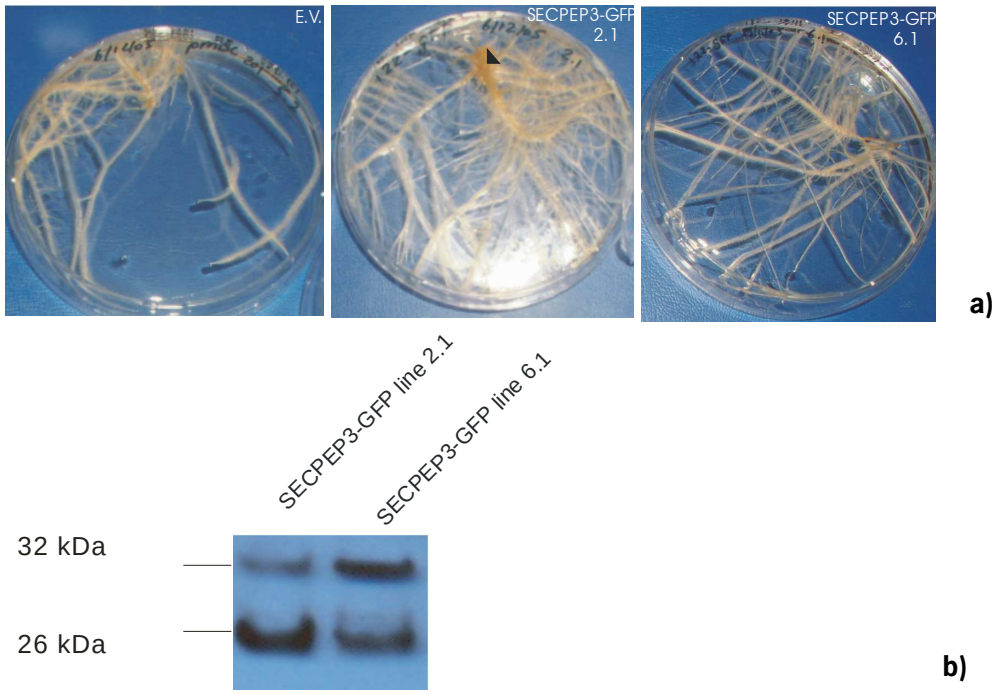


Figure 3 a) Six weeks old potato hairy roots transformed with the pMDC201 empty vector (left) and *35S::SECPEP3-GFP* transformed hairy root lines 2.1 (middle) and 6.1 (right). Roots transformed with *35S::SECPEP3-GFP* are characterized by more and longer lateral roots and in line 2.1 by callus formation on the primary root (marked by an arrowhead, middle) **b)** western blot analysis of protein isolated from an equal amount of grinded hairy root tissue expressing SECPEP3-GFP. An antibody recognizing GFP was used for the immunodetection. Lines indicating the position of the co-migrating protein size marker are shown on the left.

Efficiency of transformation is altered by transformation with a *SECPEP3* overexpression construct

To further investigate the functional role of SECPEP3, we generated transgenic potato and Arabidopsis plants, transformed with *Agrobacterium tumefaciens* harbouring a *35S::SECPEP3* construct, either with or without GFP as a fusion partner (Table 3). Transformation of potato with a *35S::SECPEP3* or *35S::GFP-SECPEP3* harbouring plasmid resulted in a shoot formation efficiency of respectively 9.7% and 9.5%, which is

nearly doubled as compared to the shoot formation efficiency of 4.9% observed upon transformation by *A. tumefaciens* harbouring a pMDC32 empty vector construct (Table 3). In contrast, transformation with a *35S::SECPEP3-GFP* containing plasmid did not result in any shoot forming callus pieces, while the total number of calli obtained was the same as observed for tissue transformed with either the empty vector control, *35S::SECPEP3* or *35S::GFP-SECPEP3* (Table 3).

Table 3: Transformation efficiency, root development and shoot phenotype of *Arabidopsis* and potato plants upon transformation with SECPEP overexpression constructs

		Empty vector	**35S::GFP-SECPEP3**	**35S::SEC-PEP3-GFP**	**35S::SECPEP3**
Potato line V	**Transformation efficiency[a]**	4.9% (4/81)	9.5% (7/74)	0% (0/83)	9.7% (7/72)
	Retarded root development[b]	NA	71% (5/7)	-	100% (7/7)
	Altered shoot phenotype[c]	-	58% (4/7)	-	14% (1/7)
Arabidopsis thaliana	**Transformation efficiency[d]**	0.06% (8/14365)	0.06% (16/25965)	NA	0.07% (17/25560)
	Retarded root development[e]	0% (0/8)	19% (3/16)	NA	18% (3/17)
	Altered number number of rosette leaves[f]	13% (1/8)	75% (12/16)	NA	24% (12% (2/17) showed an increase and 12% (2/17) showed a decrease)

[a] efficiency was calculated from the number of shoot forming calli
[b] as compared to untransformed potato line V plants
[c] as compared to empty vector transformed plants
[d] efficiency was calculated from the number of hygromycin seeds
[e] as compared to empty vector transformed plants
[f] as compared to wild type Col-o plants. The number of rosette leaves was determined at the onset of bolting
NA: Not analysed

Transformation with *35S::SECPEP3* and *35S::GFP-SECPEP3* alters the leaf morphology of potato plants and reduces root growth

Potato leaves are compound leaves consisting of an odd number of leaflets. The majority (86%) of the potato plants resulting from transformation with

35S::SECPEP3 did not show an aberrant leaf phenotype, as compared to empty vector transformed plants (Table 3). However, leaves of one of the lines (line 7.3) showed a severe reduction of the internode size between the leaflets and also the leaf area of the individual leaflets seems reduced (Fig. 4a). Analysis of the potato plants transformed with *A. tumefaciens* harbouring the *35S::GFP-SECPEP3* construct, showed that nearly sixty percent of these plants had an aberrant leaf morphology, characterized by leaves consisting of a reduced number of leaflets, fusion of some of the individual leaflets and wrinkled leaves (Fig. 4a and Table 3). Furthermore, most of the plants transformed with either *35S::SECPEP3* or *35S::GFP-SECPEP3* showed a reduction in root formation, e.g. in number of roots formed from freshly transferred cuttings (Fig. 4b and Table 3), and a delay in the timing of the formation of the first roots. Two of the *35S::GFP-SECPEP3* transformed lines with aberrant leaf morphology (line 9.6 and line 9.8) showed an increased density in root hairs and the induction of root hairs and of lateral roots occured closer to the primary root tips than observed in empty vector transformed plants (Fig. 4c).

***Arabidopsis* transformation with a *SECPEP3* overexpression construct alters the number of rosette leaves and root growth in some of the transformants**

Arabidopsis thaliana transformants were generated by use of the floral dip method ((Clough & Bent, 1998)). Efficiency of the transformation was measured as the fraction of hygromycin resistant seeds obtained upon germination of the T1 generation. The efficiency of transformation of *Arabidopsis thaliana* was not altered in case of transformation with a *SECPEP3* overexpression construct, as compared to transformation with an empty vector control (Table 3). Also, the number of rosette leaves at the onset of flowering (bolt) and the leaf morphology were not different from wild type Col-0 or empty vector transformed plants in seventy-six percent of the *Arabidopsis* plants transformed with the *35S::SECPEP3* construct (Table 3). Although the number of rosette leaves was not different from wild-type plants, the transformant

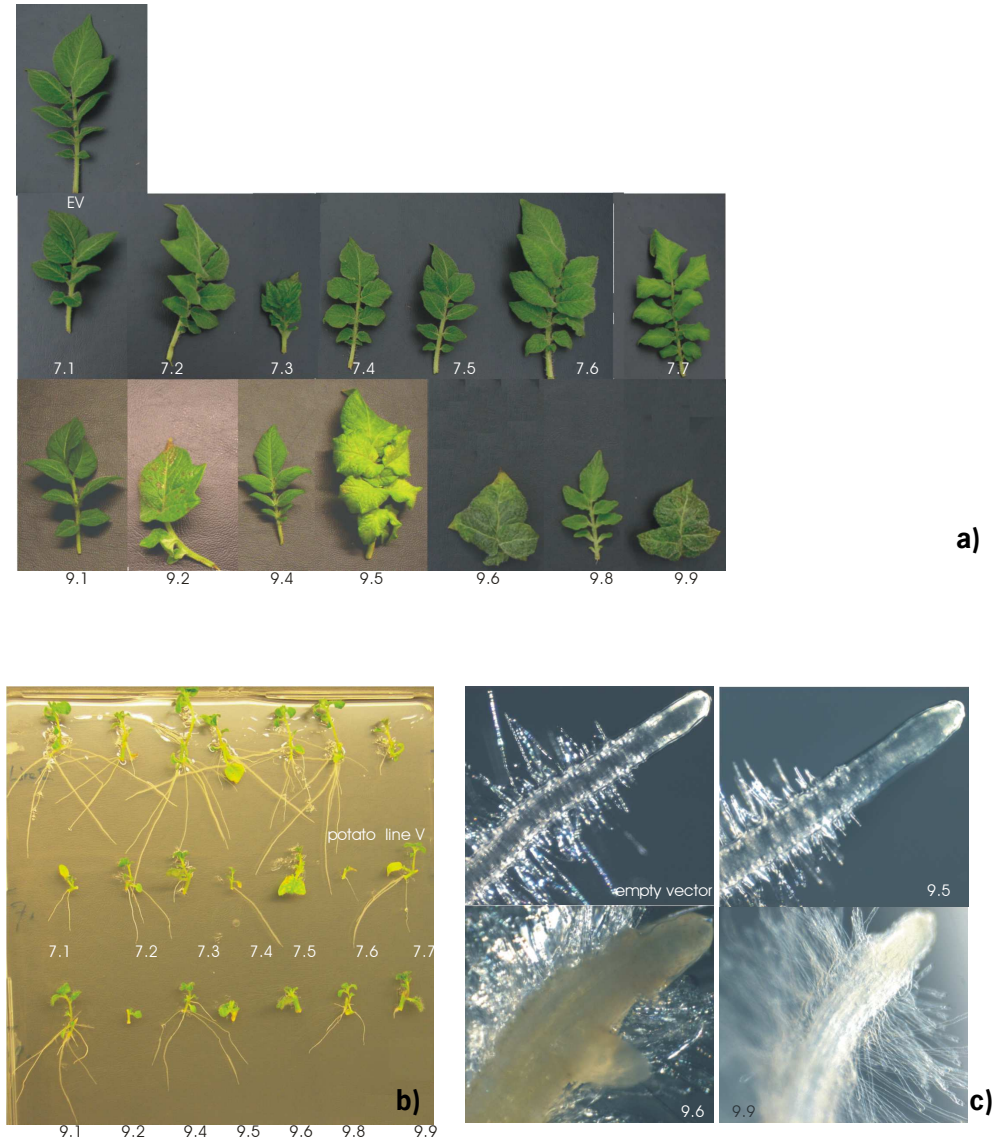


Figure 4 a) Leaf morphology of a wild type potato leaf (top), leaves from *35S::SECPEP3* transformed plants (middle) and leaves from *35S::SECPEP3-GFP* transformed plants at the bottom row **b)** two weeks old cuttings of potato line v (top row) and *35S::SECPEP3* (middle) and *35S::SECPEP3-GFP* transformed plants (bottom row) **c)** microscopy image (20x) of roots of potato plants transformed with an empty vector construct (top left) or with a *35S::GFP-SECPEP3* construct (9.5: top right (20x), 9.6: bottom left and 9.9: bottom right (35x). Both line 9.6 and line 9.9 show root hair formation closer to the root tip than observed for empty vector transformed plants, and for line 9.6 also lateral root formation was observed closer to the root tip.

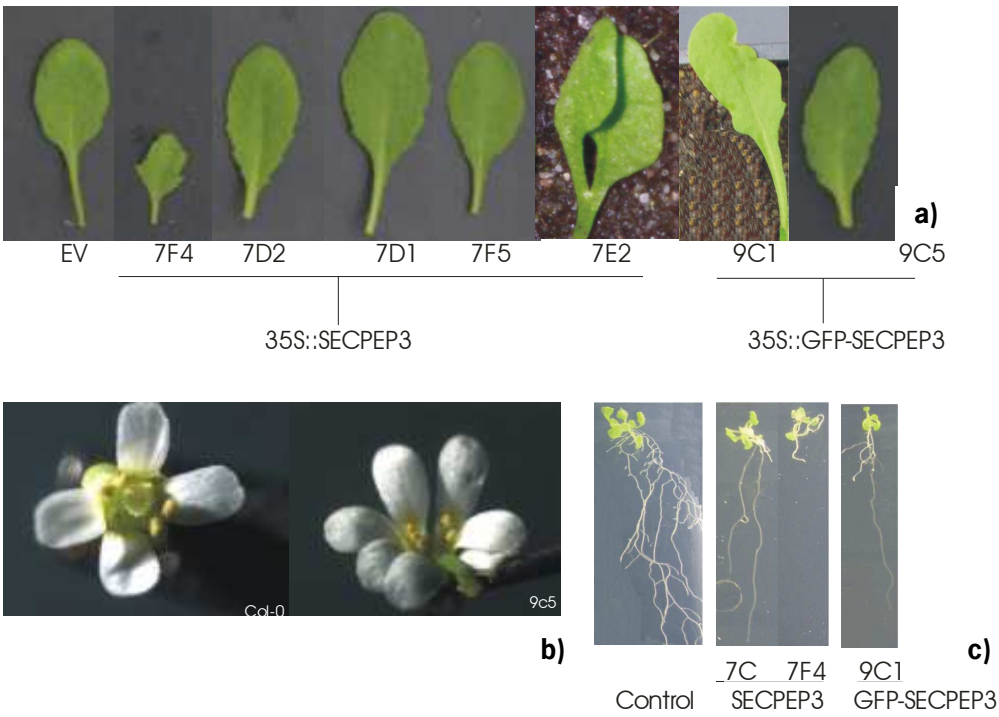


Figure 5: Phenotype of SECPEP3 overexpressing Arabidopsis plants **a)** Rosette leaf morphology of *Arabidopsis thaliana* plants transformed with an empty vector control (left) or with *35S::SECPEP3* or *35S::GFP-SECPEP3* (right) **b)** microscopy image (10x) of a wild type Col-0 flower (left) and of a flower of plant 9C5, which is transformed with *35S::GFP-SECPEP3* **c)** Root system of three weeks old Arabidopsis seedlings. The root system of a wildtype Col-0 seedling is shown on the left, and two examples of *35S::SECPEP3* transformed plants with reduced (lateral) root growth are shown in the middle. One example of a *35S::GFP-SECPEP3* transformed

with number 7F4 (*35S::SECPEP3*) had smaller rosette leaves which had an increased

sinus depth of the leaf edge that results in a “dented” appearance of the leaves (socalled serrated leaves)(Fig. 5a). Most of the rosette leaves of transformant 9C1 (*35S::GFP-SECPEP3*) had an irregular leaf edge, and one of the rosette leaves of 7E2 (*35S::SECPEP3*) seemed to have splitted (Fig. 5a). Twelve percent of the *35S::SECPEP3* transformed plants (plant number 7E1 and 7E2) had less rosette leaves at the onset of flowering (respectively 4 and 6, as compared to 8 in Col-0 plants), and

another twelve percent (plant number 7C and 7F3) had more rosette leaves (respectively 14 and 13 leaves) at the onset of flowering (Table 3). For seventy-five percent of the *35S::GFP-SECPEP3* transformed *Arabidopsis* plants we observed an increase in number of rosette leaves at the onset of flowering (ranging from 9-20 leaves as compared to 8 leaves in Col-0 plants)(Table 3). One of these plants also showed an altered flower phenotype, with 7 petals instead of the 4 petals observed for flowers of Col-0 plants (plant number 9C5, fig. 5b). However, this alteration in number of petals was only observed for the flowers of one of the two inflorescences of 9C5.

The root system of most (~80%) of the *Arabidopsis* plants transformed with either *35S::SECPEP3* or *35S::GFP-SECPEP3* looked similar to the root system of plants transformed with an empty vector construct (Fig. 5c). However, eightteen percent of the *35S::SECPEP3* transformed plants (with number 7C, 7F3 and 7F4) and nineteen percent of the *35S::GFP-SECPEP3* transformed plants (with number 9A2, 9B2 and 9C1) show a reduction in lateral root formation and, in case of 7F4 and 9B2, a reduction in primary root elongation (Fig. 5c and Table 3). All plants with altered root growth also showed an altered number of rosette leaves.

Discussion

SECPEP3 is a novel 3.2 kDa effector peptide, which is believed to be secreted into the host cell cytoplasm upon parasitism of the potato cyst nematode *Globodera rostochiensis*. By use of *in-planta* localization studies of SECPEP3 fused to GFP we showed that fusion with SECPEP3 facilitates entrance of GFP into the nucleolus of tobacco epidermal cells. Nucleolar localization has also been observed for other members of the SECPEP family (van Bers, unpublished results), and for several other cyst nematode effector proteins, e.g. the NEMPEP peptide of *G. rostochiensis* (van Bers, *et al.,* chapter 3), the C-terminal domain of the ubiquitin extension protein Hs-Ubi1 from the beet cyst nematode *Heterodera schachtii* ((Tytgat *et al.*, 2004)) and the pioneer parasitism gene 6E07 from the soybean cyst nematode ((Elling *et al.*, 2007)).

Nucleolar localization of SECPEP3 was strongest with GFP fused to the N-terminus of the peptide. We attribute this observation to cleavage occuring upon fusion of GFP to the C-terminus of SECPEP3, resulting in the separation of GFP and SECPEP3. In this case, fluorescence partly corresponds to SECPEP3-GFP and partly to unfused GFP, which does not accumulate in the nucleolus, resulting in a dilution of the fluorescent signal in the nucleolus. This hypothesis is sustained by a fainter band observed for SECPEP3-GFP on western blot, as compared to free GFP and to GFP-SECPEP3 (Fig. 1c).

The balance between the planthormones auxin and cytokinin drives the decision whether tissue explants form roots, shoots or callus. Higher levels of cytokinin induces shoot tissue, higher levels of auxin induces root tissue and media containing a balanced level of auxin and cytokinin result in callus formation (Gordon *et al.*, 2007 and references therein). Potato stem segments transformed with *35S::SECPEP3-GFP* did form callus, but shoot formation was not observed, while control calli did form shoots. A lack in shoot formation from calli was also observed in *Arabidopsis* plants upon downregulation of CYCD3, which was attributed to a reduced cytokinin response (Dewitte *et al.*, 2007). A reduction in cytokinin response would also explain the increase in lateral root formation observed upon expression of SECPEP3-GFP in potato hairy roots, as a reduced cytokinin response together with the increased sensitivity of hairy roots to auxin (Shen *et al.*, 1990, Shen *et al.*, 1988), would shift the auxin/cytokinin balance towards auxin and root formation.

Western blot analysis showed two bands being recognized by anti-GFP in SECPEP3-GFP expressing potato hairy roots. In absence of an alternative in-frame translation initiation signal, we assume that these bands are the result of proteolytic cleavage of SECPEP3-GFP. This cleavage may be important for the observed increase in lateral root formation in potato hairy roots, as lateral root formation was more prominent in a hairy root line in which the majority of SECPEP3-GFP was proteolytically processed. It would be interesting to investigate whether SECPEP3 is further

processed upon cleavage from GFP, and if the peptide remains stable following cleavage. Comparison of C-terminal and N-terminal GFP fusion of SECPEP3 in tobacco leaves indicates that only fusion of GFP to the SECPEP3 C-terminus induces cleavage. Furthermore, cleavage may not take place on the immediate border between SECPEP3 and GFP, but may occur in the SECPEP3 sequence, as the fastest migrating band detected on the western blot may have migrated slightly slower than unfused GFP. Cleavage may result in the removal of a hypothetical C-terminal inhibitory domain from SECPEP3. SECPEP3 contains a putative protein kinase C phosophorylation site near its C-terminus, and cleavage site phosphorylation has been reported to inhibit proteolysis (Tözsér *et al.*, 2003). Fusion to GFP may hamper phosphorylation of SECPEP3, and in that way promote cleavage. Transformation of potato with a *SECPEP3-GFP* overexpressing construct failed to result in transgenic plants, while transformation efficiency doubled upon transformation with a *35S::SECPEP3* or *35S::GFP-SECPEP3* construct, which supports the hypothesis that fusion of GFP to the C-terminus of SECPEP3 alters its activity. Analysis of the phenotype of *Arabidopsis* plants transformed with *35S::SECPEP3-GFP* may shed more light on this. Additionally, it would be valuable to correlate the phenotypes observed in the transformed plants with accumulation of the SECPEP3 peptide or a derivative thereof.

In some of the potato plants transformed with *35S::GFP-SECPEP3* we observed lateral root initiation closer to the root tip than observed in the empty vector transformed control plants. Lateral roots emerge in the differentiation zone of the root, at a fixed distance above the root tip, where pericycle cells progress via the S to G2 phase of the cell cycle (Beeckman *et al.*, 2001, Casimiro *et al.*, 2003, Casimiro *et al.*, 2001). SECPEP3 expression may result in aberrant cell cycle (re)activation resulting in lateral root formation closer to the root tip. However, speeding up cell cycle progression is not sufficient for lateral root initiation (Vanneste *et al.*, 2005). Recently, de Smet and coworkers (De Smet *et al.*, 2007) proposed a model in which local auxin accumulation primes cells of the basal meristem, which is the transition zone between

the root apical meristem and the elongation zone, for cell cycle (re)activation which results in the initiation of lateral roots. This indicates that besides a hypothesized cell cycle reactivation, also the local auxin balance has been altered in some of the roots of potato plants transformed with *35S::GFP-SECPEP3*. It remains to be investigated if this hypothesis can be sustained by correlating the plant lines with altered positioning of lateral roots with a certain expression levels of GFP-SECPEP3 or with processing or modification of the protein.

Cell cycle proliferation of tobacco protoplasts was stimulated by the fraction of small proteins (<3 kDa) present in the secretions of *Globodera rostochiensis* preparasitic juveniles (Goverse *et al.*, 1999). This stimulation required the presence of the plant hormones auxin and cytokinin (Goverse et al., 1999). These plant hormones may prime the protoplasts for the proliferative activity of nematode peptides, in a way similar to proposed for cells of the basal meristem upon lateral root formation, and indeed, local auxin accumulation accompanies feeding cell formation and lateral root formation is promoted when the initial syncytial cell is located in or in the immediate vicinity of the pericycle (reviewed (Goverse *et al.*, 2000). The enhanced sensitivity of hairy root cultures to auxin (Shen et al., 1990, Shen et al., 1988) together with a hypothetical activation of the cell cycle due to overexpression of *SECPEP3* could explain the enhanced lateral root formation observed upon *SECPEP3-GFP* overexpression.

SECPEP3 shares a KEKP sequence motif with the *Arabidopsis* protein MGOUN3 (also known as BRUSHY1 and TONSOKU). Interestingly, some of the *Arabidopsis* plants transformed with a *SECPEP3* overexpression construct show striking phenotypical similarities to plants with a mutation in the MGOUN3 gene. *Mgo3/bru1/tsk* mutants show a.o. perturbation in leaf morphogenesis, the development of more rosette leaves (although this development is delayed as compared to wild-type plants), slower root growth, development of root hairs closer to the root tip and *mgo3/bru1/tonsoku* flowers have a variable number of flower organs (Guyomarc'h *et al.*, 2004, Suzuki *et al.*, 2005, Takeda *et al.*, 2004). Some of the *35S::SECPEP3* or *35S::GFP-SECPEP3*

transformed *Arabidopsis* plants had more rosette leaves and showed a reduction in root growth, and line 9C5 formed flowers with seven petals instead of four observed in wild-type Col-0 flowers. Additionally, two of the potato plants transformed with *35S::GFP-SECPEP3* formed root hairs closer to the root tip. The *mgo3/bru1/tsk* mutant has disorganized meristems (Guyomarc'h et al., 2004, Suzuki *et al.*, 2004), and MGO3 is believed to mediate chromatin modifications involved in expression of the flowering locus C (FLC) gene (Guyomarc'h *et al.*, 2006). Interestingly, chromatin modifications are also important in the control of transcription and cell cycle progression (Shen, 2002), and the *mgo3/bru1/tsk* mutant showed defects in cell cycle progression and has altered ploidy levels (Suzuki et al., 2004, Suzuki et al., 2005). It remains to be investigated whether the *35S::GFP-SECPEP3* transformed plants show a disorganized meristemal organization and/or alterations in ploidy levels.

The KEKP(L) motif is also found in the coding sequence of the cyclin-dependent kinase inhibitor ICK1/KRP1, and *Arabidopsis* plants overexpressing ICK1/KRP1 (Wang *et al.*, 2000, Zhou et al., 2002, Weinl *et al.*, 2005) show phenotypic similarities to a subset of the *SECPEP3* overexpressing potato and *Arabidopsis* plants, e.g. a retardation in root growth and small, serrated leaves (*Arabidopsis* transformant 7F4). ICK1/KRP1 was shown to interact with CDKA;1, CYCD1, CYCD2 and CYCD3 in *Arabidopsis* (Wang et al., 1998, Zhou et al., 2002, Zhou et al., 2003b). CDKA;1 was formerly known as cdc2a, and transcriptional activation of the *cdc2a* promoter was detected inside and around nematode-induced feeding cells (Niebel et al., 1996). Expression of ICK1 resulted in a slight increase in CDKA;1 transcript level (Wang et al., 2000), although CDKA;1 kinase activity was inhibited. ICK1 inhibits the complex formation of CYCD3 with CDKA (Wang et al., 1998). Downregulation of CYCD3, by use of a triple loss-of-function mutant in *Arabidopsis*, resulted in a reduced cytokinin response, which included the inability to initiate shoots from callus tissue (Dewitte et al., 2007). The latter was also observed upon transformation of potato with *35S::SECPEP3-GFP*. A reduced cytokinin response could also explain the enhanced

lateral root formation observed in potato hairy roots (see above), which suggests that SECPEP3 may downregulate CYCD3.

ICK1 encodes a nuclear localized protein consisting of 191 amino-acids, of which the first 109 are suggested to harbour an inhibitory domain (Weinl et al., 2005), which may render the protein unstable (Zhou *et al.*, 2003a). Upon fusion of ICK1 with GFP, some cells showed a punctuate pattern of fluorescence throughout the nucleus (Zhou *et al.*, 2006). Interestingly, nucleolar fluorescence of GFP-SECPEP3 also had a punctuated appearance. The family members of the *Arabidopsis* KRP protein family are highly divergent, and sequence similarity seems to be limited to three motifs at the C-terminus of the protein. Motif 3 is similar to a motif found in the SIAMESE (SIM) protein of *Arabidopsis*, and overexpression of ICK1/KRP1, can correct the phenotype of a *sim* mutant (Weinl et al., 2005). In analogy to ICK1, SIAMESE (SIM) was shown to interact with several D-type cyclins, and with CDKA;1 (Churchman *et al.*, 2006). SIM encodes a 14kDa protein (Churchman et al., 2006), which is positively charged at pH 7, and it contains a C-X-G motif, although in reverse amino-acid primary sequence (G-X-C). SECPEP3 shares a KIP motif with SIM and with the SIAMESE related protein Smr1. Additionally, the (K)PRPP(P) motif found in SIM and Smr1 is also found in SECPEP6 and 7, which are close relatives of SECPEP3. ICK/KRP1 has a C-terminal C-X-G motif, and the five residue motif KEKPL, which is also present in the coding sequence of the mature SECPEP3 peptide, is found at position 177-181 of ICK1/KRP1. Interestingly, deletion of amino-acid 163-191 of ICK1 was shown to severely reduce its ability to bind CDKA and its kinase inhibition activity (Zhou et al., 2003a). The KEKP(L) and (K)PRPP(P) motif may be a motif involved in mediating peptide or protein-protein interactions, as both motifs were also found in two members of the *Arabidopsis thaliana* PROPEP gene family, respectively PROPEP2 and PROPEP6 (Huffaker *et al.*, 2006). The PROPEP gene family encodes for precursors of peptides of about 23 amino-acids (*At*PEP) involved in the amplification of defense pathways induced by pathogens ((Huffaker et al., 2006, Huffaker & Ryan, 2007) and chapter 6 of this thesis).

Both the *mgo3/bru1/tsk* mutant and ICK1 overexpressing *Arabidopsis* plants have cells with altered ploidy levels, due to endoreduplication (Suzuki et al., 2005, Weinl et al., 2005). For ICK1 expressing plants this effect is dependent on the level of ICK1/KRP1, with low levels of ICK1/KRP1 promoting the S phase progression and the occurrence of endocycles, and high levels inhibiting endoreduplication (Weinl et al., 2005, Wang et al., 2000). It would be interesting to investigate if a similar concentration dependent effect can explain the seemingly contrasting observations upon *SECPEP3* expression in hairy roots as compared to "whole plant" transformants, and the absence of an altered phenotype in some of the transformants. Endoreduplication is also one of the cellular changes that accompany syncytium formation, resulting in enlarged cells with prominent nuclei (reviewed by (Gheysen *et al.*, 1997)). It would be interesting to investigate if *SECPEP3* overexpression results in an altered ploidy level, and whether SECPEP3 interacts with cyclins or cyclin dependent kinases, or with a component of the E2F-Rb pathway, which is essential in cell cycle control. Interestingly, Rb proteins are thought to contain a putative peptide binding site (Claudio *et al.*, 2002).

In conclusion, we hypothesize that SECPEP3 interferes with cell cycle progression, which may result in an altered responsiveness to the plant hormone cytokinin, and/or in an alteration in ploidy level of the host cell. SECPEP3 shares sequence motifs with several cyclin-dependent kinase inhibitors and with members of the *Arabidopsis At*PEP family. We are currently investigating if SECPEP3 alters the activation of the CYCD3;1 and CDKA;1 promoter. The observation that overexpression of *SECPEP3* alters development of *Arabidopsis*, creates a wide range of possibilities for complementation studies with some of the transgenic and mutant lines available for this species, e.g. with a CYCD3 overexpressing line. Furthermore, analysis of the interaction partners of SECPEP3 will provide more insight into the functional role of SECPEP3 in feeding cell formation.

References

Beeckman, T., Burssens, S. and Inze, D. (2001) The peri-cell-cycle in Arabidopsis. *J Exp Bot,* **52,** 403-411.

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

Casimiro, I., Beeckman, T., Graham, N., Bhalerao, R., Zhang, H., Casero, P.*, et al.* (2003) Dissecting Arabidopsis lateral root development. *Trends in Plant Science,* **8,** 165-171.

Casimiro, I., Marchant, A., Bhalerao, R. P., Beeckman, T., Dhooge, S., Swarup, R.*, et al.* (2001) Auxin transport promotes Arabidopsis lateral root initiation. *Plant Cell,* **13,** 843-852.

Churchman, M. L., Brown, M. L., Kato, N., Kirik, V., Hu?lskamp, M., Inze, D.*, et al.* (2006) SIAMESE, a plant-specific cell cycle regulator, controls endoreplication onset in Arabidopsis thaliana. *Plant Cell,* **18,** 3145-3157.

Claudio, P. P., Tonini, T. and Giordano, A. (2002) The retinoblastoma family: twins or distant cousins? *Genome Biol,* **3,** reviews3012.

Clough, S. J. and Bent, A. F. (1998) Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant Journal,* **16,** 735-743.

Curtis, M. D. and Grossniklaus, U. (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol,* **133,** 462-469.

de Almeida Engler, J., De Vleesschauwer, V., Burssens, S., Celenza, J. L., Jr., Inze, D., Van Montagu, M.*, et al.* (1999) Molecular markers and cell cycle inhibitors show the importance of cell cycle progression in nematode-induced galls and syncytia. *Plant Cell,* **11,** 793-808.

de Boer, J. M., Smant, G., Goverse, A., Davis, E. L., Overmars, H. A., Pomp, H.*, et al.* (1996) Secretory granule proteins from the subventral esophageal glands of the potato cyst nematode identified by monoclonal antibodies to a protein fraction from second-stage juveniles. *Mol Plant Microbe Interact,* **9,** 39-46.

De Smet, I., Tetsumura, T., De Rybel, B., dit Frey, N. F., Laplaze, L., Casimiro, I., *et al.* (2007) Auxin-dependent regulation of lateral root positioning in the basal meristem of Arabidopsis. *Development,* **134,** 681-690.

De Veylder, L., Beeckman, T., Beemster, G. T., Krols, L., Terras, F., Landrieu, I., *et al.* (2001) Functional analysis of cyclin-dependent kinase inhibitors of Arabidopsis. *Plant Cell,* **13,** 1653-1668.

Dewitte, W., Scofield, S., Alcasabas, A. A., Maughan, S. C., Menges, M., Braun, N., *et al.* (2007) Arabidopsis CYCD3 D-type cyclins link cell proliferation and endocycles and are rate-limiting for cytokinin responses. *Proceedings of the National Academy of Sciences,* **104,** 14537-14542.

Elling, A. A., Davis, E. L., Hussey, R. S. and Baum, T. J. (2007) Active uptake of cyst nematode parasitism proteins into the plant cell nucleus. *International Journal for Parasitology,* **37,** 1269-1279.

Endo, B. Y. (1964) Penetration and development of Heterodera glycines in soybean roots and related anatomical changes. *Phytopathology,* **54,** 79-88.

Evans, K. and Stone, A. R. (1977) A Review of the Distribution and Biology of the Potato Cyst-Nematodes *Globodera rostochiensis* and *G. pallida*. *PANS,* **23,** 178-189.

Gheysen, G., Engler, J. D. A. and Montagu, M. V. (1997) Cell cycle regulation in nematode feeding sites. In: *Cellular and molecular aspects of plant-nematode interactions.* (Fenoll, C., Grundler, F. M. W. and Ohl, S. A., ed.^eds.). Kluwer Academic Publishers, pp. 120-132.

Golinowski, W., Grundler, F. M. W. and Sobczak, M. (1996) Changes in the structure of Arabidopsis thaliana during female development of the plant-parasitic nematode Heterodera schachtii. *Protoplasma,* **194,** 103-116.

Gordon, S. P., Heisler, M. G., Reddy, G. V., Ohno, C., Das, P. and Meyerowitz, E. M. (2007) Pattern formation during de novo assembly of the Arabidopsis shoot meristem. *Development,* **134,** 3539-3548.

Goverse, A., Rouppe van der Voort, J., Roppe van der Voort, C., Kavelaars, A., Smant,

G., Schots, A.*, et al.* (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.

Goverse, A., de Engler, J. A., Verhees, J., van der Krol, S., Helder, J. H. and Gheysen, G. (2000) Cell cycle activation by plant parasitic nematodes. *Plant Mol Biol,* **43,** 747-761.

Grundler, F. M. W., Sobczak, M. and Golinowski, W. (1998) Formation of wall openings in root cells of Arabidopsis thaliana following infection by the plant-parasitic nematode Heterodera schachtii. *European Journal of Plant Pathology,* **104,** 545-551.

Guyomarc'h, S., Benhamed, M., Lemonnier, G., Renou, J. P., Zhou, D. X. and Delarue, M. (2006) MGOUN3: Evidence for chromatin-mediated regulation of FLC expression. *Journal of Experimental Botany,* **57,** 2111-2119.

Guyomarc'h, S., Vernoux, T., Traas, J., Zhou, D. X. and Delarue, M. (2004) MGOUN3, an Arabidopsis gene with TetratricoPeptide-Repeat-related motifs, regulates meristem cellular organization. *J Exp Bot,* **55,** 673-684.

Hoekema, A., Hirsch, P. R., Hooykaas, P. J. J. and Schilperoort, R. A. (1983) A binary plant vector strategy based on separation of vir- and T-region of the Agrobacterium tumefaciens Ti-plasmid. *Nature,* **303,** 179-180.

Horsman, K., Gavrilenko, T., Bergervoet, M., Huigen, D. J., Joe, A. T. W. and Jacobsen, E. (2001) Alteration of the genomic composition of Solanum nigrum (+) potato backcross derivatives by somatic hybridization: Selection of fusion hybrids by DNA measurements and GISH. *Plant Breeding,* **120,** 201-207.

Huffaker, A., Pearce, G. and Ryan, C. A. (2006) An endogenous peptide signal in Arabidopsis activates components of the innate immune response. *Proc Natl Acad Sci U S A,* **103,** 10098-10103.

Huffaker, A. and Ryan, C. A. (2007) Endogenous peptide defense signals in Arabidopsis

differentially amplify signaling for the innate immune response. *Proceedings of the National Academy of Sciences of the United States of America,* **104,** 10732-10736.

Hussey, R. S. and Grundler, F. M. W. (1998) Nematode parasitism of plants. In: *The Physiology and Biochemistry of Free-living and Plant-parasitic Nematodes.* (Perry, R. N. and Wright, D. J., ed.^eds.). CAB International, pp. 213-243.

Inze, D. (2005) Green light for the cell cycle. *Embo J,* **24,** 657-662.

Jans, D. A. and Hu?bner, S. (1996) Regulation of protein transport to the nucleus: Central role of phosphorylation. *Physiological Reviews,* **76,** 651-685.

Karczmarek, A., Overmars, H., Helder, J. and Goverse, A. (2004) Feeding cell development by cyst and root-knot nematodes involves a similar early, local and transient activation of a specific auxin-inducible promoter element. *Molecular Plant Pathology,* **5,** 343-346.

Kudla, U., Milac, A. L., Qin, L., Overmars, H., Roze, E., Holterman, M.*, et al.* (2007) Structural and functional characterization of a novel, host penetration-related pectate lyase from the potato cyst nematode Globodera rostochiensis. *Molecular Plant Pathology,* **8,** 293-305.

Nakai, K. and Horton, P. (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends in Biochemical Sciences,* **24,** 34-35.

Nakazawa, M. and Matsui, M. (2003) Selection of hygromycin-resistant Arabidopsis seedlings. *BioTechniques,* **34,** 28-30.

Niebel, A., de Almeida Engler, J., Hemerly, A., Ferreira, P., Inze, D., Van Montagu, M.*, et al.* (1996) Induction of cdc2a and cyc1At expression in Arabidopsis thaliana during early phases of nematode-induced feeding cell formation. *Plant J,* **10,** 1037-1043.

Shen, W. H. (2002) The plant E2F-Rb pathway and epigenetic control. *Trends Plant Sci,* **7,** 505-511.

Shen, W. H., Davioud, E., David, C., Barbier-Brygoo, H., Tempe, J. and Guern, J. (1990) High sensitivity to auxin is a common feature of hairy root. *Plant Physiology,* **94,** 554-560.

Shen, W. H., Petit, A., Guern, J. and Tempe, J. (1988) Hairy roots are more sensitive to auxin than normal roots. *Proceedings of the National Academy of Sciences of the United States of America,* **85,** 3417-3421.

Shen, W. J. and Forde, B. G. (1989) Efficient transformation of Agrobacterium spp. by high voltage electroporation. *Nucleic Acids Research,* **17,** 8385.

Sigrist, C. J., Cerutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M.*, et al.* (2002) PROSITE: a documented database using patterns and profiles as motif descriptors. *Briefings in bioinformatics,* **3,** 265-274.

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

Sobczak, M., Golinowski, W. and Grundler, F. M. W. (1997) Changes in the structure of *Arabidopsis thaliana* roots induced during development of males of the plant parasitic nematode *Heterodera schachtii*. *European Journal of Plant Pathology,* **103,** 113-124.

Suzuki, T., Inagaki, S., Nakajima, S., Akashi, T., Ohto, M. A., Kobayashi, M.*, et al.* (2004) A novel Arabidopsis gene Tonsoku is required for proper cell arrangement in root and shoot apical meristems. *Plant Journal,* **38,** 673-684.

Suzuki, T., Nakajima, S., Inagaki, S., Hirano-Nakakita, M., Matsuoka, K., Demura, T.*, et al.* (2005) TONSOKU is expressed in S phase of the cell cycle and its defect delays cell cycle progression in Arabidopsis. *Plant and Cell Physiology,* **46,** 736-742.

Takeda, S., Tadele, Z., Hofmann, I., Probst, A. V., Angelis, K. J., Kaya, H.*, et al.* (2004) BRU1, a novel link between responses to DNA damage and epigenetic gene silencing in Arabidopsis. *Genes and Development,* **18,** 782-793.

Tözsér, J., Bagossi, P., Zahuczky, G., Specht, S. I., Majerova, E. and Copeland, T. D. (2003) Effect of caspase cleavage-site phosphorylation on proteolysis. *Biochemical Journal,* **372,** 137-143.

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

van Engelen, F. A., Schouten, A., Molthoff, J. W., Roosien, J., Salinas, J., Dirkse, W. G.*, et al.* (1994) Coordinate expression of antibody subunit genes yields high levels of functional antibodies in roots of transgenic tobacco. *Plant Molecular Biology,* **26,** 1701-1710.

Vanneste, S., Maes, L., De Smet, I., Himanen, K., Naudts, M., Inze, D.*, et al.* (2005) Auxin regulation of cell cycle and its role during lateral root initiation. *Physiologia Plantarum,* **123,** 139-146.

Visser, R. G. F., Jacobsen, E., Witholt, B. and Feenstra, W. J. (1989) Efficient transformation of potato (Solanum tuberosum L.) using a binary vector in Agrobacterium rhizogenes. *Theoretical and Applied Genetics,* **78,** 594-600.

von Mende, N., Gravato-Nobre, M. J. and Perry, R. N. (1998) Host finding, invasion and feeding. In: *The cyst nematodes.* (Sharma, S., ed.^eds.). Dordrecht: Kluwer Academic Publishers.

Wang, H., Qi, Q., Schorr, P., Cutler, A. J., Crosby, W. L. and Fowke, L. C. (1998) ICK1, a cyclin-dependent protein kinase inhibitor from Arabidopsis thaliana interacts with both Cdc2a and CycD3, and its expression is induced by abscisic acid. *Plant Journal,* **15,** 501-510.

Wang, H., Zhou, Y., Gilmer, S., Whitwill, S. and Fowke, L. C. (2000) Expression of the plant cyclin-dependent kinase inhibitor ICK1 affects cell division, plant growth and morphology. *Plant Journal,* **24,** 613-623.

Weinl, C., Marquardt, S., Kuijt, S. J. H., Nowack, M. K., Jakoby, M. J., Hu?lskamp, M.*, et*

al. (2005) Novel functions of plant cyclin-dependent kinase inhibitors, ICK1/KRP1, can act non-cell-autonomously and inhibit entry into mitosis. *Plant Cell,* **17,** 1704-1722.

Yeaman, M. R. and Yount, N. Y. (2007) Unifying themes in host defence effector polypeptides. *Nature Reviews Microbiology,* **5,** 727-740.

Yount, N. Y. and Yeaman, M. R. (2004) Multidimensional signatures in antimicrobial peptides. *Proceedings of the National Academy of Sciences of the United States of America,* **101,** 7363-7368.

Zhou, Y., Fowke, L. and Wang, H. (2002) Plant CDK inhibitors: Studies of interactions with cell cycle regulators in the yeast two-hybrid system and functional comparisons in transgenic Arabidopsis plants. *Plant Cell Reports,* **20,** 967-975.

Zhou, Y., Li, G., Brandizzi, F., Fowke, L. C. and Wang, H. (2003a) The plant cyclin-dependent kinase inhibitor ICK1 has distinct functional domains for in vivo kinase inhibition, protein instability and nuclear localization. *Plant Journal,* **35,** 476-489.

Zhou, Y., Niu, H., Brandizzi, F., Fowke, L. C. and Wang, H. (2006) Molecular control of nuclear and subnuclear targeting of the plant CDK inhibitor ICK1 and ICK1-mediated nuclear transport of CDKA. *Plant Molecular Biology,* **62,** 261-278.

Zhou, Y., Wang, H., Gilmer, S., Whitwill, S. and Fowke, L. C. (2003b) Effects of co-expressing the plant CDK inhibitor ICK1 and D-type cyclin genes on plant growth, cell size and ploidy in Arabidopsis thaliana. *Planta,* **216,** 604-613.

Chapter 5

Novel and highly diverse spliced-leader sequences are *trans*-spliced onto potato cyst nematode *SECPEP1* transcripts

Nikkie van Bers, Ling Qin, Martijn Holterman, Aska Goverse, Jaap Bakker and Geert Smant

To be submitted

Abstract

In Rhabditine nematodes, SL2-like spliced leader sequences are *trans*-spliced onto the 5'-end of transcripts of genes that are located downstream in an operon. SL1 or SL1-like sequences have been found in all nematode species investigated to date. Here we surveyed SL *trans*-splicing of transcripts of the *SECPEP1* gene from the Tylenchine nematode *Globodera rostochiensis*. Besides SL1, we find twenty-six novel SL sequences that can be grouped into four different clusters. The SL RNA sequences of three of these clusters have a secondary structure and genomic organization that may be more similar to *C. elegans* SL2 than to SL1. Furthermore, we find surprising sequence diversity in *SECPEP1* transcripts, which is not present in the *SECPEP1* genomic sequence. Therefore, we believe that RNA editing contributes to the diversification of the SECPEP1 peptide.

Introduction

Cis-splicing and spliced leader (SL) *trans*-splicing are two distinct steps in the maturation of pre-mRNA molecules in some eukaryotes. *Cis*-splicing results in the removal of spliceosomal introns, while in the process of SL *trans*-splicing a leader sequence is added onto the 5'-end of transcripts (Blumenthal & Thomas, 1988)). *Cis*- and SL *trans*-splice sites have the same consensus sequences. The only factor that determines whether SL *trans*-splicing takes place is the absence of a splice donor site upstream of a splice acceptor site (Blumenthal, 1995).

SL *trans*-splicing was first discovered in trypanosomes (phylum Euglenozoa), where all nuclear mRNA are *trans*-spliced (reviewed in (Agabian, 1990). Subsequently SL *trans*-splicing was identified in the phyla Nematoda (Krause & Hirsh, 1987), Platyhelminthes (Rajkovic *et al.*, 1990), Cnidaria (Stover & Steele, 2001), Rotifera (Pouchkina-Stantcheva & Tunnacliffe, 2005) Dinoflagellata (Zhang *et al.*, 2007) and Chordata (Vandenberghe *et al.*, 2001). Seventy percent of the transcribed genes of *Caenorhabditis elegans* is estimated to be SL *trans*-spliced, and fifty percent is *trans*-spliced to the conserved, 22 nucleotide, SL1 sequence (Zorio *et al.*, 1994).

Different functional roles have been suggested for the addition of a spliced leader. In nematodes, SL sequences may function in creating an optimal length of the 5'UTR or in promoting translation initiation or efficiency as a SL is often found just upstream (<10 nucleotides) of the AUG codon (Maroney *et al.*, 1995), (Lall *et al.*, 2004), (Cheng *et al.*, 2007). Additionally, SL sequences function in resolving polycistronic transcripts, which result from the transcription of genes organized in operons. As such, spliced leader sequences may prevent premature transcription termination or protect the downstream mRNA from degradation upon 3'-end formation of the upstream gene (Liu *et al.*, 2001, Evans & Blumenthal, 2000) and reviewed in (Blumenthal, 2004). Operonic organization is estimated for 15-25% of *C. elegans* genes (Zorio et al., 1994, Blumenthal *et al.*, 2002).

Two types of SL sequences have been identified in *C. elegans*, SL1 and SL2 (Huang & Hirsch, 1989, Krause & Hirsh, 1987). SL2 has only been found in nematodes from the suborder Rhabditina, and there is a strong correlation between its presence and the downstream localization of a gene in an operon (Blumenthal et al., 2002, Spieth *et al.*, 1993, Guiliano & Blaxter, 2006). SL sequences are donated by a ~100 nucleotide SL RNA. Other than the SL sequence itself, there is little sequence conservation between SL RNAs of different nematode species, although they potentially adopt similar secondary structures (Bruzik *et al.*, 1988, Nilsen *et al.*, 1989, Redmond & Knox, 2001, Evans & Blumenthal, 2000, Huang & Hirsch, 1989). The secondary structure of SL1 RNA is characterized by three stem loops, with the SL basepaired within the first loop (Bruzik et al., 1988). Furthermore, all nematode SL RNAs contain a highly conserved Sm binding site (RAU_nGR)(Van Doren & Hirsh, 1988).

Recent phylogenetic analysis has shown that the phylum Nematoda can be subdivided into twelve clades (Holterman et al., 2006). The potato cyst *Globodera rostochiensis* is an obligatory plant parasite of a small range of Solanaceous species and belongs to the clade of the Tylenchomorpha in the suborder Tylenchina. The life cycle of *Globodera rostochiensis* consists of four juvenile stages, followed by an adult stage (Turner & Evans, 1998). The infective J2 juvenile transforms a competent host cell into a feeding structure, called syncytium (von Mende *et al.*, 1998). The nematodes' relatively large esophageal glands are packed with secretory granules containing effector molecules, which are believed to be involved in host defense suppression, migration and feeding cell formation (reviewed by (Gheysen & Jones, 2006). *SECPEP1* is specifically expressed in the dorsal esophageal gland of early (pre)parasitic juveniles. It encodes a small (~5 kDa) peptide, which is believed to be secreted into the host cell cytoplasm. SECPEP1 shares its positive charge, hydrophilicity and sequence divergence with nine other *SECPEP* genes, and with venom toxins and antimicrobial peptides (van Bers, chapter 2 of this thesis). In this paper, we use *SECPEP1* transcripts to study SL *trans*-splicing in the tylenchine *Globodera rostochiensis*. We reflect our

finding of a wide diversity of novel spliced-leader sequences onto the process of SL *trans*-splicing in the rhabditine nematode *C. elegans.* To get more insight into *SECPEP1* transcript diversity and diversification, we further investigated the sequence diversity in *SECPEP1* transcripts and compared this to the diversity in *SECPEP1* genomic sequence.

Material and methods

RNA isolation and cloning of full length cDNA

Preparasitic juveniles of *G. rostochiensis* pathotype Ro1-Mierenbos were collected as described previously (De Boer *et al.*, 1992). Total RNA was isolated from preparasitic juveniles using TRIzol reagent (Invitrogen, Breda, the Netherlands). Five µg of total RNA was the starting material for full-length, RNA ligase-mediated rapid amplification of 5' and 3' cDNA ends (RLM-RACE) by use of the GeneRacer Kit (Invitrogen), which was performed according to the manufacturers' protocol. Superscript III reverse transcriptase (Invitrogen) was used for cDNA synthesis, according to the suppliers' protocol. Full-length *SECPEP1* cDNA was amplified by using the gene specific 3'UTR (271-293) RV primer (5'CAAAATGCCTTCAGCAAAATGAC 3') in combination with an adapter specific forward primer provided in the GeneRacer kit. The resulting PCR product was subsequently cloned in the pCR4-TOPO vector (Invitrogen). To confirm the identity of the insert sequence a colony PCR was done with a nested adapter specific primer in combination with the A4 3'RV primer (5' GCGTGGCCTTCTTGTTCCAGTTGG 3'). High fidelity Taq DNA polymerase with proof reading ability (Roche) was used for all PCR reactions.

For the amplification of the 5'-end of *Gr-gpd* the GAPDH(268-291) RV primer (5' CCTCTTTCTCAACGACCAAATTGC 3') was used in combination with an adaptor specific forward primer. The amplification product was cloned into the pCR4-TOPO vector (Invitrogen). Glycerol stocks of the bacterial cultures were used as template for

sequencing at BaseClear (Leiden, the Netherlands). Insert sequences were sequenced from both 5'- and 3'-end with vector specific primers.

Southern blot analysis and amplification of genomic DNA

Genomic DNA was extracted from preparasitic J2 as described by (Curran *et al.*, 1985). About 2.5 µg of genomic DNA was digested by respectively *EcoRI*, *BamHI*, *KpnI* and *BglII* and separated on a 0.8% agarose gel. The southern blot was performed as described by (Sambrook *et al.*, 1989). Hybridization was performed with a 174bp DIG labeled *SECPEP1* cDNA probe amplified from plasmid DNA at 40°C overnight in DIG Easy Hyb solution (Roche Diagnostics). The filter was washed twice in 0.1xSSC/0.1% SDS solution before detection. Alkaline phosphatase activity was detected using DIG Luminiscent Detection Kit (Roche Diagnostics) and the chemiluminiscent signal was exposed on X-ray films.

SECPEP1 gDNA was amplified in a PCR reaction with the (881-904)FW primer (5' TTTATCCAATTCATTGTTCTTCTC 3') in combination with the 3'UTR(271-293)RV primer (5' CAAAATGCCTTCAGCAAAATGAC 3'). For the PCR amplification of the SL RNA genes the library vector pcDNA FW primer (5' GGTGACACTATAGAATACTCAAGCTATGCA 3') was used in combination with a SL specific primer. In all PCR reactions DNA polymerase with proof reading ability was used (High Fidelity Taq DNA polymerase, Invitrogen).

Sequence analysis and statistics

DNA sequences were analysed using the Informax Vector NTI Advance software (Invitrogen). Sequence alignments were performed with the ClustalW 1.83 algorithm, default settings (Thompson *et al.*, 1994). For the mutational analysis, the sequences were aligned and the number and type of nucleotide changes (as compared to the consensus sequence) were manually scored for each clone. Nucleotide changes at position 69, 82 and 87 of *SECPEP1* cDNA and the corresponding positions in *SECPEP1*

gDNA were ignored as they are likely to represent allelic variation. To correct for different lengths of the amplification products, the nucleotide change scores were expressed as nucleotide changes per 1000 basepairs amplified.

Secondary structure predictions of the SL RNA were performed on the Mfold server v 2.3 (Burnet institute)(Zuker, 2003) with t=25°C and the constraint that the Sm binding site remains single stranded, in analogy to (Bruzik et al., 1988). RnaViz2 (De Rijk et al., 2003) was used for the graphical representation of the secondary structure.

For all statistical analyses a student t-test assuming unequal variances was performed in Microsoft Excel.

Phylogenetic analysis

The SL sequences were aligned using the ClustalW algorithm (v1.83, default settings). Treepuzzle 5.2 and Treeview (v. 1.6.6) were used to generate an unrooted phylogenetic tree, in which nodes with a posterior probablilty lower than 0.60 are considered to be unresolved.

Results

26 novel and highly diverse spliced-leader sequences are *trans*-spliced onto the 5'-end of *SECPEP1* transcripts

Spliced Leader 1 (SL1) has been found on the 5'-end of transcripts of all nematode species investigated sofar, with virtually no sequence divergence (Bektesh *et al.*, 1988) and reviewed in (Nilsen, 1997). SL2-like leader sequences, however, have only been found in the suborder Rhabditina (Guiliano & Blaxter, 2006). Here, we study the diversity in SL sequence on the 5'-end of *SECPEP1* transcripts of the potato cyst nematode *Globodera rostochiensis*, a member of the suborder Tylenchina.

Rapid Amplification of cDNA ends (RACE) using an adapter-specific forward primer in combination with a *SECPEP1* reverse primer, resulted in the amplification of the 5'-end of *SECPEP1* transcripts. We sequenced 166 clones, all harbouring the *SECPEP1* sequence. Surprisingly, we found twenty-seven different SL sequences to be *trans*-spliced onto the 5'-end of *SECPEP1* transcripts (Fig. 1). Seventy percent of the clones were *trans*-spliced to a 22 nt sequence identical to SL1, while an additional seven percent of the clones contain a SL that is nearly identical to SL1 (1-2 nucleotides difference)(Fig. 1 and 2a). SL1 variants have been reported before, although only in the tylenchid species *Meloidogyne spp.* and *Aphelenchus avenae* (Koltai *et al.*, 1997, Ray *et al.*, 1994, Lambert *et al.*, 1999, Goyal *et al.*, 2005). However, none of the SL1 variants found at the 5'-end of the *SECPEP1* transcripts have been reported before.

The remaining twenty-three percent of the clones was *trans*-spliced to novel spliced leader sequences of 18-25 nucleotides in length (Fig. 1). Phylogenetic analysis of the spliced leader sequences shows that they cluster in four distinct clusters (Fig. 2b). In this analysis nodes with a posterior probability lower than 0.60 are considered unresolved. This value is lower than the more commonly used 0.95 as the short length of the SL sequences (18-25 nucleotides) gives rise to lower probability values. We named the SL cluster containing SL1 *Gr*-SL1, and the other three SL clusters *Gr*-SL2, *Gr*-SL3 and *Gr*-SL4. No sequence identical or remotely similar to *C. elegans* SL2 has

been found in *G. rostochiensis*, but the SL sequences constituting clusters 2, 3 and 4 all end with the 3-nt motif AAG, which is also found for *C. elegans* SL-2 like genes (Fig. 1 and (Guiliano & Blaxter, 2006). The consensus sequence assigned to each of the novel clusters is shown in Table 1.

```
SL2a              ---GACTTAACCAAAAGTC--CAAAAG
SL2g              ---GACTTAACCAAAAGTC--CAAAAG
SL2h              ---GACTTAACCAAAAGTC--CGAAAG
SL2b              ---GACTTAACCAAAGGTC--CAAAAG
SL2f              ----GACTTACCAAAAGTC--CAAAAG
SL2i              ----GCCTTACCAAAAGTC--CAAAAG
SL2d              ----GACTTTCCATAAATC--CAAAAG
SL2e              ----GACTTTCCATAAGTC--CAAAAG
SL2c              GTTGACTTAACCAAAAGTC--CAAAAG
SL2\C.\elegans    --GGTTTTAACCC-AGTTA--CTCAAG
SLg\C.\elegans    --GGTTTTAACC--AGTTA--ACTAAG
SL5\C.\elegans    --GGTTTTAACCCAAGTTA--ACCAAG
SL1               ---GGTTTAATTACCCAAG--TTTGAG
SL1b              ---GATTTAATTACCCAAG--TTTGAG
SL1e              ---GGTTTAATTACCCAGG--TTTGAG
SL1c              ---GGTTTAATTACCCTAG--TTTGAG
SL1i              ---GGTTTAATTACCCGAG--TTTGAG
SL1d              ---GGCTTAATTACCCAAG--TTTGAG
SL1a              ---GGTTTAATTACCCAAG--TTTGGG
SL1h              ---GGTTTAACTACCCAAG--TTTGAG
SL1f              ---GGTTTAATCACCCAAG--TTTGAG
SL1g              ---GGTTTAATTACCCAAA--TATGAG
SL3a              -----GCTATTTATAGTCCATAGTAAG
SL3b              ----GGCTATTTATAGTCCATAGTAAG
SL3c              ----GGCTATTTATAGTCC-TAGTAAG
SL3f              ----GGCTATTTATAGTCC-TGGTAAG
SL3d              ----GGCTATTGATAGTCC-TAGTAAG
SL3e              ----GGTTATTTATAGTCC-TAGTAAG
SL4a              -----ACTAGTTTTAGTCC-AAATAAG
SL4f              ----GACTAGTTTTATTCC-AAATAAG
SL4b              ----GTCTAGTTTTAGTCC-AAATAAG
SL4d              ---------AGTTTTAGTCC-AAATAAG
SL4e              ----GACTGGTTTTAGTCC-AAATAAG
SL4c              -------GAGTTTTAGTCC-AAATAAG
```

Figure 1: Alignment of the primary structures of the spliced leaders (Gr-SL) found on the 5'-end of *SECPEP1* transcripts. As a comparison, the Ce-SL2, Ce-SLg and Ce-SL5 found on *C.elegans* transcripts are also included.

None of the SL-like sequences are found in the genomic DNA sequence 800 nucleotides upstream of the *SECPEP1* translation initiation codon, indicating that they are indeed *trans*-spliced onto *SECPEP1* transcripts. Additionally, a splice acceptor sequence (TTTTAG) consistent to a *G. rostochiensis* splice acceptor site (Qin *et al.*, 1998), and nearly identical to the consensus *C.elegans* splice acceptor sequence (TTTCAG)(Ross et al., 1995) is found at position -11 to -6 relative from the start codon.

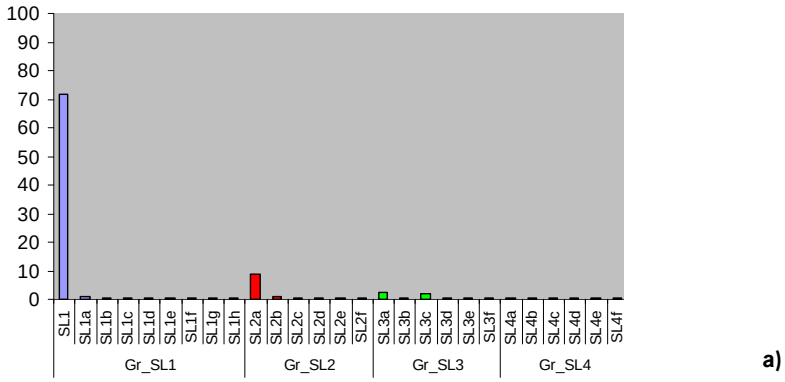
Table 1: Consensus sequences of *Gr-SL2*, *Gr-SL3* and *Gr-SL4*, derived from alignment of their primary structures

SL cluster	Consensus
*Gr*_SL2	$CTTX_{1-2}CCAXAX_2TCCXAAG$
*Gr*_SL3	$AXXGTCCX_{2-5}AAG$
*Gr*_SL4	$GT_4AXTCCAAATAAG$

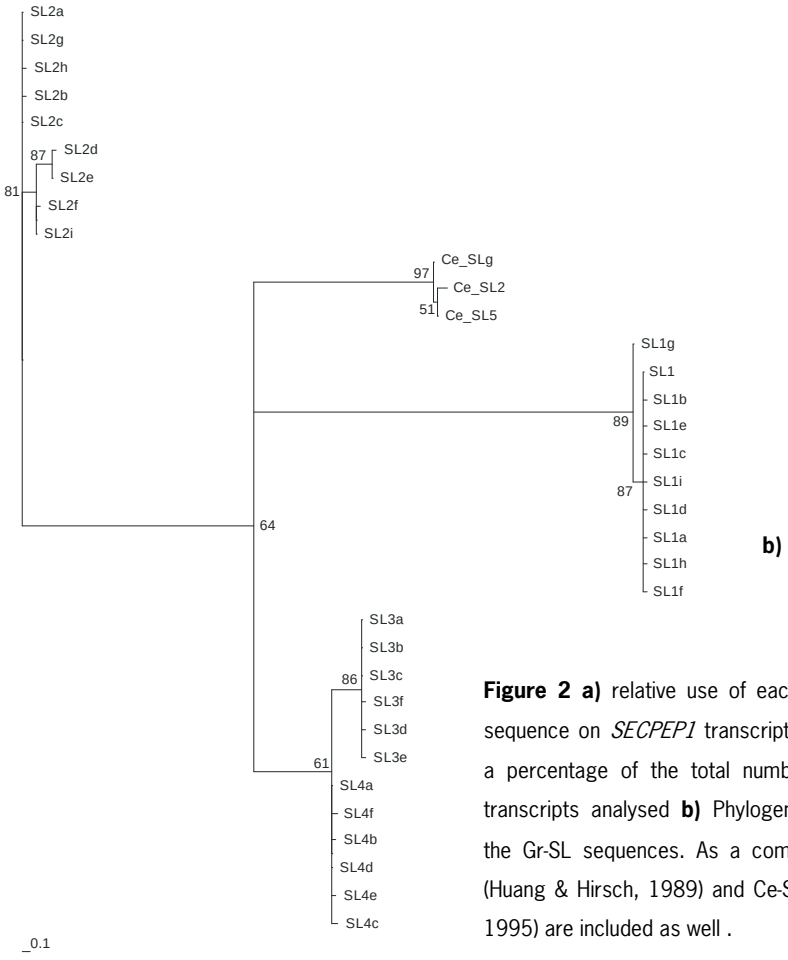
Variant SLs are also found on *Gr-gpd* transcripts

SL2 was first discovered at the ultimate 5'-end of *C. elegans* glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNAs (Huang & Hirsch, 1989). *C. elegans* has four GAPDH genes, *gpd-1* to *gpd-4*. *Gpd-2* and *gpd-3* are downstream genes in an operon, and ten percent of the *gpd-2* transcripts and nearly all the transcripts of *gdp-3* are *trans*-splied to a SL2 sequence (Spieth et al., 1993, Huang & Hirsch, 1989).

For *Globodera rostochiensis* four, nearly identical, ESTs corresponding to a *gpd*-like gene have been found, which was named *Gr-gpd*. Gr-gpd shows 79-80% amino-acid sequence similarity to the four *C. elegans gpd* gene products, and its transcripts have been shown to receive a SL1 leader sequence by *trans*-splicing (Qin *et al.*, 1998).



a)



b)

Figure 2 a) relative use of each spliced leader sequence on *SECPEP1* transcripts, expressed as a percentage of the total number of *SECPEP1* transcripts analysed **b)** Phylogenetic analysis of the Gr-SL sequences. As a comparison, Ce-SL2 (Huang & Hirsch, 1989) and Ce-SL5 (Ross et al., 1995) are included as well .

We wondered whether *Gr-gpd* transcripts are exclusively *trans*-spliced to SL1, or if also alternative SL sequences are used in their *trans*-splicing. Therefore, we sequenced 50 full-length *Gr-gpd* RACE clones from both ends. All the amplified transcripts have a SL sequence at their 5'-end. Eighty-eight percent of the clones start with the conserved SL1 sequence. Interestingly, twelve percent of the *Gr-gpd* clones have a 5'-end consisting of a SL other than SL1 (Fig. 3). These SLs are identical to SLs found on *SECPEP1* transcripts, of which two group into cluster 2, while the third one groups into cluster 3.

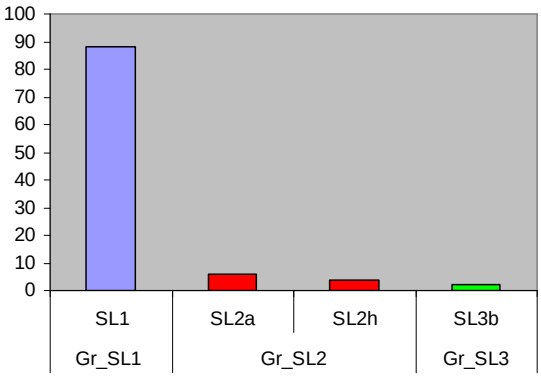


Figure 3: Relative use of each spliced leader sequence on the 5'-end of fifty full-length *Gr-gpd transcripts*, expressed as a percentage of the total number of *Gr-gpd* transcripts analysed.

*SECPEP*1 may or may not be a single copy gene

In order to find out if the 27 SL sequences are all *trans*-spliced onto the transcripts of one, single copy, *SECPEP1* gene we further analysed the sequence downstream of the SL. Analysis of 172 *SECPEP1* coding sequences shows that seventy-five percent of the insert sequences of the clones are not identical to the consensus *SECPEP1* sequence. The sequences of these 129 clones include either one or more point mutations, but also deletions or insertions are present. Nucleotide changes are found at 22 different positions of the 247 bp fragment amplified. A nucleotide change at position 69 (counted relative to the first nucleotide of SL1) and a correlated change at the positions 82 and 87 is found in thirty-one per cent of the clones. As these changes may represent allelic variation rather than occasional point mutations, they are ignored in

further calculations. On average, we observed 0.19 synonymous nucleotide changes per 1000 nucleotides amplified (Fig. 4a).

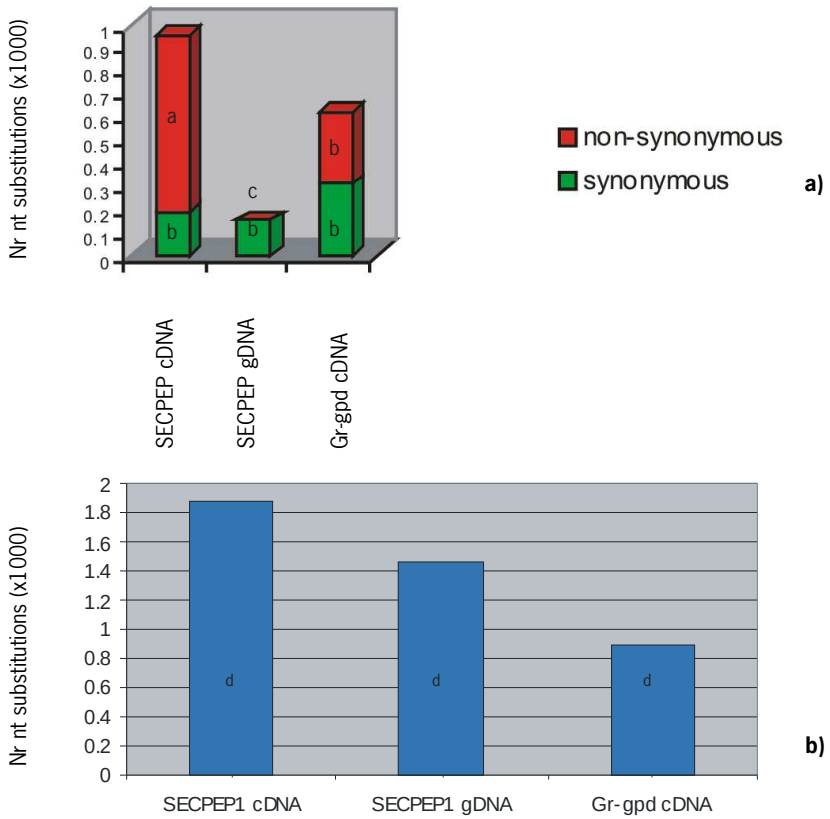


Figure 4: a) The average number of nucleotide changes per 1000 amplified nucleotides of the coding region as observed for each clone of *SECPEP*1 cDNA, *SECPEP*1 gDNA and *Gr-gpd1* cDNA. The bars were subsequently subdivided into synonymous and non-synonymous changes. Insertions and deletions are classified as non-synonymous changes. Different lettercodes are given to significantly different areas of the graph **b)** The average number of nucleotide changes per 1000 amplified nucleotides of the non-coding region observed for each clone containing *SECPEP*1 cDNA, *SECPEP*1 gDNA and *Gr-gpd1* cDNA. Different lettercodes are given to significantly different areas of the graph.

However, significantly more ($P<0.005$) nucleotide changes are non-synonymous (0.76 per 1000 nucleotides amplified). Some of these non-synonymous mutations result in radical changes such as a nucleotide change in the translation initiation codon or Arg→Pro, Arg→Gly, Gly→Asp and Ser→Ala changes. One substitution is found in the 3'UTR region resulting in an additional in-frame stop codon. No correlation could be established between a mutation at a certain position and the use of a certain kind of SL.

The different transcripts may be derived from multiple *SECPEP1* gene copies in the *G. rostochiensis* genome. To get more information about the *SECPEP1* gene copy number, we performed a southern blot analysis. Genomic DNA was digested with four different enzymes and hybridized with a 174 bp probe designed on the *SECPEP1* cDNA sequence. In all cases a prominent, single band was observed (Fig. 5), indicating that *SECPEP1* is likely to be a single copy gene. An EcoRI restriction site is present in the *SECPEP1* coding sequence, and hybridization of EcoRI digested genomic DNA showed a single band corresponding to a size of about 1600 bp. The discrepancy between the expected gene copy number and the number of nucleotide changes observed in the *SECPEP1* transcripts suggests that the nucleotide changes were induced during or post-transcription. We sequenced 25 clones containing a 369 bp fragment of the *SECPEP1* genomic sequence, which was amplified with the same reverse primer as used for the amplification of the *SECPEP1* transcripts. Eighty-three percent of the clones were not identical to the consensus *SECPEP1* sequence, which can be mainly attributed to nucleotide changes at position 69 and position 82/87 (positions are numbered as in cDNA), which is found in seventy-six percent of the clones. In total, mutations are found at 6 different positions. Except for the mutations at position 69 and 87 none of the observed mutations is non-synonymous, which is significantly less than observed for *SECPEP1* cDNA (Fig. 4a). No nucleotide insertions or deletions have been found in the clones containing the *SECPEP1* genomic fragment. The average number of synonymous nucleotide changes and the number of changes in non-coding

regions of *SECPEP1* gDNA are not significantly different from *SECPEP1* cDNA (Fig. 4a and b).

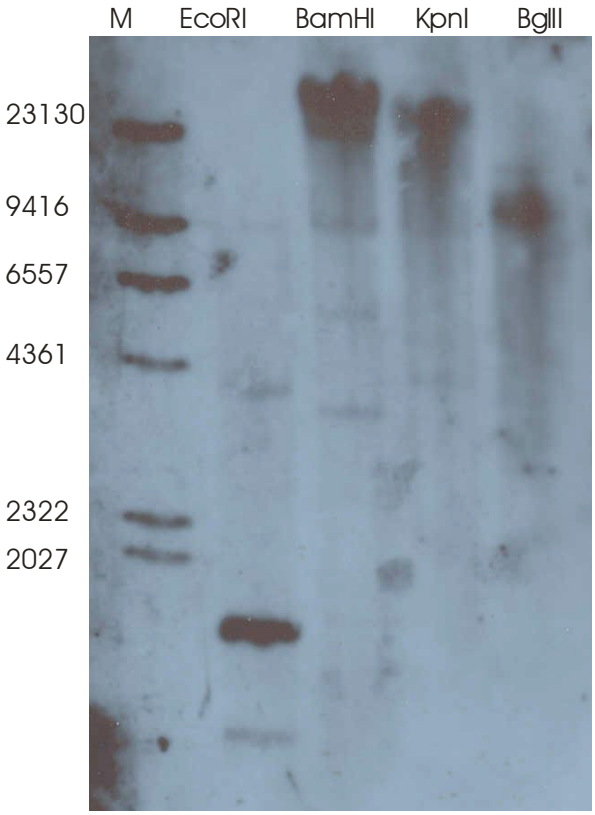


Figure 5: Southern blot analysis with a 174 bp*SECPEP*1 probe. 2.5 µg nematode genomic DNA was digested with respectively *EcoRI*, *BamHI*, *KpnI* and *BglII* . M corresponds to the size marker.

Some of the nucleotide changes observed in the *SECPEP1* transcripts may be artifacts induced during reverse transcription or PCR. To further investigate this, we amplified transcripts of *Gr-gpd*, and sequenced 50 of the corresponding clones. Also in the *Gr-gpd* transcripts point mutations at seven different positions and in one case a deletion were observed. Eightteen percent of the *Gr-gpd* harbouring clones contain a nucleotide change in the *Gr-gpd* sequence. An average number of 0.31 non-synonymous nucleotide changes (per 1000 bp amplified) is found for *Gr-gpd* cDNA, which is significantly less than observed for *SECPEP1* cDNA ($p<0.05$)(Fig. 4a).

Additionally, all non-synonymous changes result in relatively conserved changes (Lys→Arg and twice Phe→Leu). The number of synonymous nucleotide changes, and the number of changes found in the UTR region are not significantly different from the ones observed for *SECPEP1* cDNA (Fig. 4a and 4b).

Gr-SL RNA genes have a conserved Sm binding site and secondary structure

In *C. elegans* SL1 and SL2 are donated by small nuclear RNAs, called SL RNAs, which are 95 and 110 nucleotides in length, respectively (Huang & Hirsch, 1989). As a further confirmation that the sequences at the 5'-end of *SECPEP1* and *Gr-gpd* are indeed novel spliced leaders, we screened a genomic library for corresponding SL RNA genes. An oligonucleotide corresponding to the SL sequence was used as a primer in combination with a library-vector specific primer. Using the SL2a sequence as a primer resulted in amplification of a 353 bp product, while amplification with either the SL3a or the SL4b primer resulted in a product of 414 and 464 bp respectively (Fig. 6). The SL2a and SL4b sequences were immediately followed by the typical GT dinucleotide splice donor site, which is also found in *C. elegans* SL RNA genes, whereas the SL3a sequence is followed by GC.

All nematode SL RNAs identified up to date have a Sm antigen binding site (5'(A/G)AU_{4-6}G(A/G) 3') (Thomas *et al.*, 1988, Liu *et al.*, 2003). This site is also found in the three novel SL RNA gene sequences (Fig. 6). SL RNA genes of different organisms show little primary sequence conservation (Bruzik et al., 1988, Vandenberghe et al., 2001), although in several cases the secondary structure seems to be conserved to some extent (Agabian, 1990, Nilsen et al., 1989, Bruzik et al., 1988). *C. elegans* SL RNAs fold into a three or four stem-loop structure in which the SL is base-paired in the first stem-loop, while the Sm antigen-binding site is located between two stem-loop structures (Fig. 7). The first 100 nucleotides of the novel *G. rostochiensis* SL RNA genes show no sequence similarity to *C. elegans* SL1 RNA genes (12-15%), to the SL5 RNA gene (19-27%) or to the *Globodera rostochiensis* SL1 RNA gene (15-23%)(Fig. 7).

On the other hand, the primary sequences of the SL3a and SL4b RNA gene are 73% identical, while the SL2a RNA gene have 31% and 56% identity with respectively the SL3a and SL4b RNA gene. Although the primary sequences of the novel SL RNA genes are not conserved with *C. elegans* SL RNA sequences, their predicted secondary structures are strikingly similar (Fig. 7). The leader exon is base paired in the first or first two stem-loop structures, while the Sm binding site resides between two stem-loop structures. Especially the predicted Gr-SL3a RNA and Gr-SL4b RNA structure are remarkably similar. Furthermore, the secondary structures of the novel *G. rostochiensis* SL RNA genes seem to show more similarity to the secondary structure of Ce-SL2a RNA than to the Ce-SL1 RNA structure.

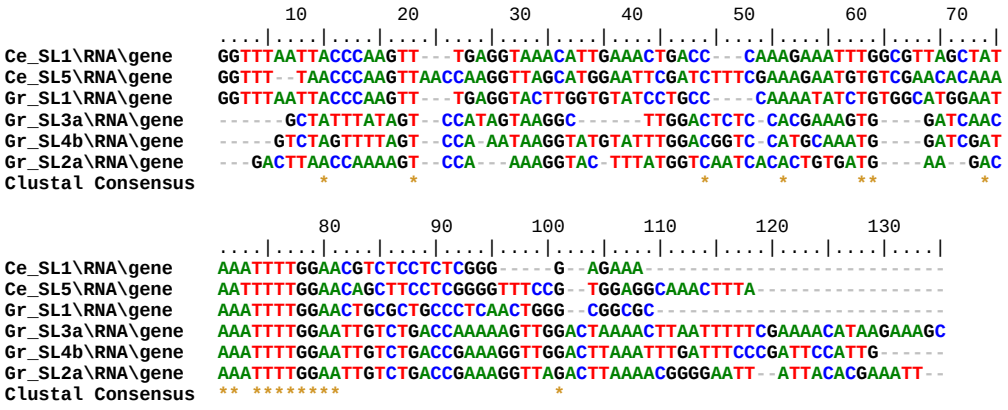


Figure 6: Alignment of *C. elegans* and *G. rostochiensis* SL RNA genes. The spliced leader exons and Sm antigen-binding site are shaded. Ce-SL1 RNA is as described by (Krause & Hirsh, 1987), Ce-SL5 RNA is as described by (Ross *et al.*, 1995) and Gr-SL1 RNA is as described by (Stratford & Shields, 1994).

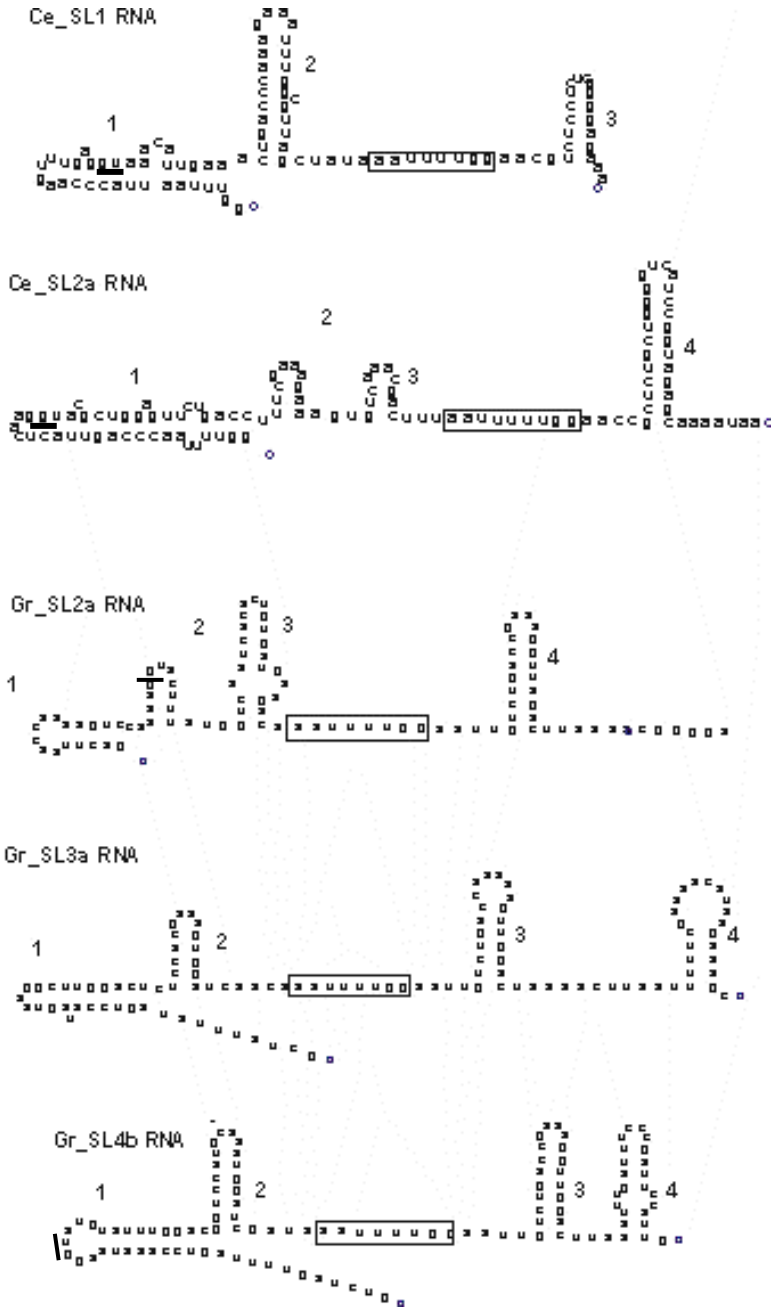


Figure 7: The secondary structure of the SL RNAs as predicted by mfold v 2.3 (Zuker, 2003), with an unpaired Sm binding site as a constraint. RnaViz2 (De Rijk *et al.*, 2003) is used for the graphical representation of the structure. The Sm binding site is boxed and the splice donor site is underlined.

Gr-SL RNA genes are organized in repeats

The ~110 copies of the *C. elegans* SL1 RNA gene are organized on a 1-kb repeat that also specifies the 5S rRNA genes (Krause & Hirsh, 1987). Also *G. rostochiensis* SL1 RNA was found in the 5S rRNA spacer (Stratford & Shields, 1994). The 19 copies of *C. elegans* SL2-like RNA genes are dispersed across the genome, although about half of them are organized in a few small clusters (reviewed in (Stein *et al.*, 2003) and (Guiliano & Blaxter, 2006). A similarity search does not indicate the presence of a 5S rRNA sequence flanking the novel *G. rostochiensis* RNA sequences, however the amplified region is too short to exclude clustering with 5S rRNA genes. Interestingly, we found that all three novel SL RNA genes are organized in a cluster with at least one other SL RNA gene, as another SL sequence is observed 300-400 nt downstream of the first SL exon (Fig. 8). In all cases, this downstream SL sequence belongs to the same phylogenetic cluster as the upstream SL exon. In case of the SL4b RNA gene, the downstream SL sequence is identical to SL4b, and also the following 27 nucleotides of the SL RNA intron are identical (Fig. 8). The PCR-amplified nucleotides of the SL RNA introns of the gene downstream of the SL2a and SL3a RNA gene show about 80% similarity to the intronic region of the upstream gene.

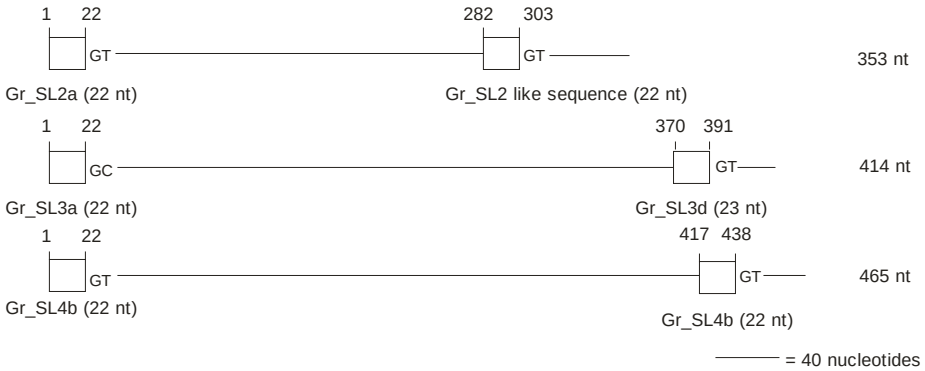
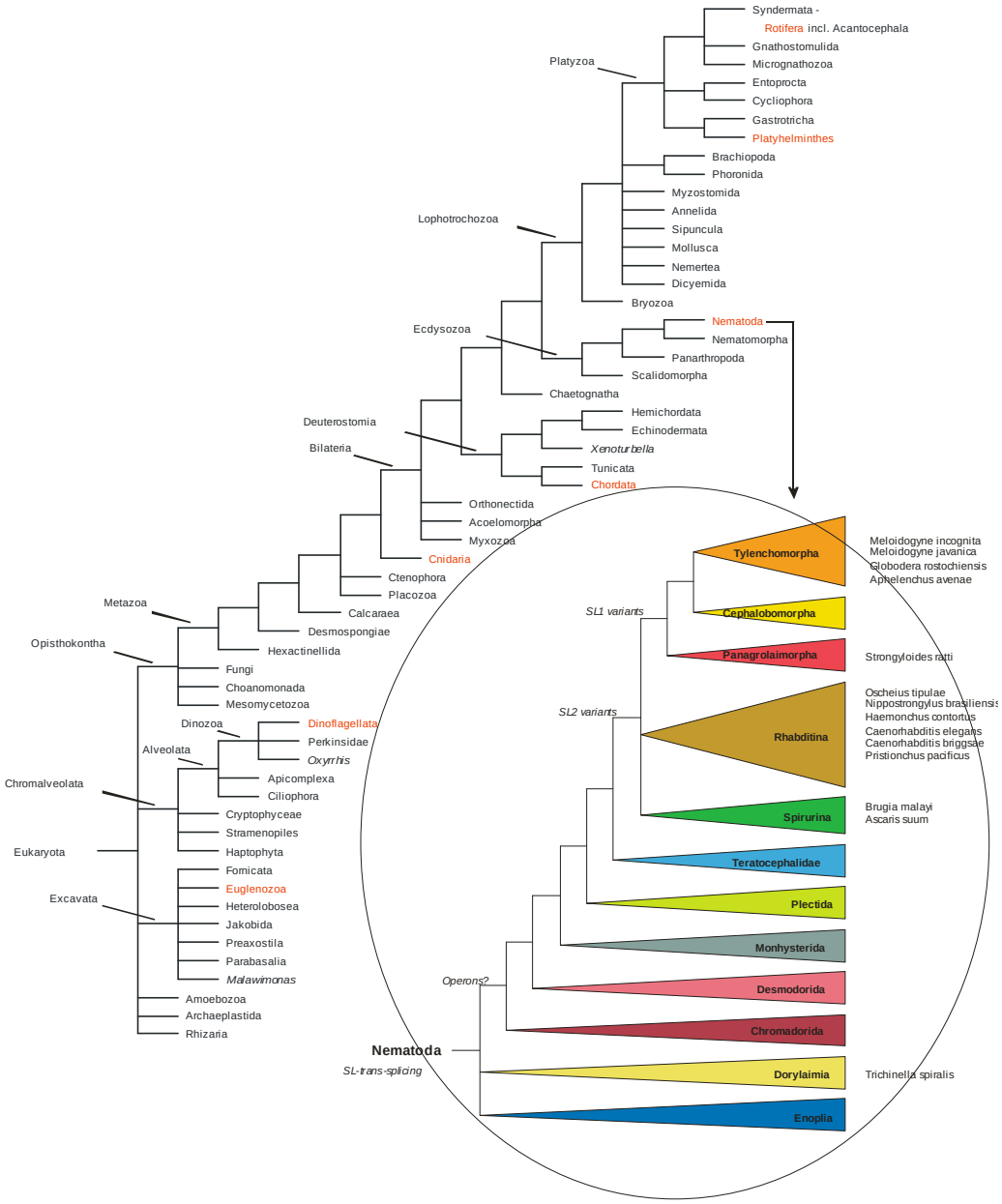


Figure 8: Schematic representation of the genomic organization of novel *G. rostochiensis* SL RNA genes. The SL exons are shown as rectangles, followed by a dinucleotide splice donor site. The intronic and flanking regions of the genes are shown as a line.

***Globodera rostochiensis* is the first nematode species outside suborder Rhabditina to have SL1, SL1-like and other spliced leader variants**

Guiliano and Blaxter (2006) showed that SL2 variants are restricted to the suborder Rhabditina, while SL1 variants (1 nucleotide difference) have previously been found in the suborder Tylenchina to which the potato cyst nematode belongs. However, our finding of SL1, SL1 variants and three novel SL classes results in an adapted mapping of SL usage within the phylum Nematoda (Fig. 9).

Figure 9: The SL usage mapped onto a phylogeny of the phylum Nematoda (adapted from (Holterman *et al.*, 2006); Guiliano, 2006 #785}). SL1 *trans*-splicing may be conserved throughout the phylum, while the use of SL1-like SLs seems to be restricted to the Tylenchina. SL2-like SLs have been observed on transcripts derived from Rhabditina as well as Tylenchina genes.


Syndermata -
Rotifera incl. Acantocephala
Gnathostomulida
Micrognathozoa
Entoprocta
Cycliophora
Gastrotricha
Platyhelminthes
Platyzoa
Brachiopoda
Phoronida
Myzostomida
Annelida
Sipuncula
Mollusca
Nemertea
Dicyemida
Lophotrochozoa
Bryozoa
Nematoda
Nematomorpha
Panarthropoda
Scalidomorpha
Ecdysozoa
Chaetognatha
Hemichordata
Echinodermata
Xenoturbella
Tunicata
Chordata
Deuterostomia
Bilateria
Orthonectida
Acoelomorpha
Myxozoa
Cnidaria
Ctenophora
Placozoa
Calcaraea
Desmospongiae
Hexactinellida
Metazoa
Fungi
Choanomonada
Mesomycetozoa
Opisthokontha
Dinoflagellata
Perkinsidae
Oxyrrhis
Dinozoa
Apicomplexa
Ciliophora
Alveolata
Cryptophyceae
Stramenopiles
Haptophyta
Chromalveolata
Fornicata
Euglenozoa
Heterolobosea
Jakobida
Preaxostila
Parabasalia
Malawimonas
Eukaryota
Amoebozoa
Archaeplastida
Rhizaria
Excavata
Nematoda
SL-trans-splicing
Operons?
SL1 variants
SL2 variants
Tylenchomorpha
Cephalobomorpha
Panagrolaimorpha
Rhabditina
Spirurina
Teratocephalidae
Plectida
Monhysterida
Desmodorida
Chromadorida
Dorylaimia
Enoplia
Meloidogyne incognita
Meloidogyne javanica
Globodera rostochiensis
Aphelenchus avenae
Strongyloides ratti
Oscheius tipulae
Nippostrongylus brasiliensis
Haemonchus contortus
Caenorhabditis elegans
Caenorhabditis briggsae
Pristionchus pacificus
Brugia malayi
Ascaris suum
Trichinella spiralis

Discussion

In this paper we showed that the spliced leader sequences used for *trans*-splicing by the tylenchid nematode *Globodera rostochiensis* are surprisingly diverse, both in sequence and in number. Some of the novel SLs seem to share more similarities with SL2-like sequences than with SL1 or SL1-like sequences. Up to now, SL2-like *trans*-splicing has only been observed for the Rhabditina (Guiliano & Blaxter, 2006). We hypothesize that SL2-like *trans*-splicing is actually more widespread within the phylum Nematoda than anticipated before.

SL1 is found in all nematode species investigated, with virtually no sequence divergence (reviewed by (Blaxter & Liu, 1996). Only in the tylenchid species *Meloidogyne javanica*, *M. incognita* and *Aphelenchus avenae* SL1-like sequences are found that differ in some positions from the canonical SL1 (Koltai et al., 1997, Stratford & Shields, 1994, Ray et al., 1994, Goyal et al., 2005). In 1989, SL2 was discovered in *C. elegans* (Huang & Hirsch, 1989). Genome sequencing indicates that eighteen SL2-like sequences are present in the *C. elegans* and *C.briggsae* genomes (Stein et al., 2003). Although less conserved than SL1, the SL2 sequences share 63-100% identity. At the 5'-end of transcripts of the *Globodera rostochiensis* gene *SECPEP1,* we observed a surprising number of twenty-seven variant SL sequences. These SLs are considerably more diverse than the ones described for *Caenorhabditis spp.*, having only 18-100% identity. This abundancy in SL variants found in *Globodera rostochiensis* was recently also found for *Trichinells spiralis* (Pettitt *et al.*, 2008), but is appently lacking in most other *trans*-splicing organisms. Sequencing of the 5'-end of 1,500 random full-length cDNA clones of *Ascaris suum* showed the presence of only one spliced leader sequence; SL1 (Lall et al., 2004). In *Hydra vulgaris*, a member of the phylum Cnidaria, two different SLs were found upon analysis of fifty 5'UTR sequences (Stover & Steele, 2001). In flatworms, phylum Platyhelminthes, more variation has been found in SL sequence and length as compared to the phylum Nematoda (Davis, 1997, Zayas *et al.*, 2005). However, this variance is mainly found in an interspecies comparison, as

variation within a single species seems to be limited. Sequence analysis of the 5'-end of 300 *trans*-spliced ESTs from the platyhelminth *Schmidtea mediterranea* indicated the presence of two SL sequences (Zayas et al., 2005). Analysis of 2078 randomly sampled 5'-full-length ESTs from the chordate *Ciona intestinalis* showed that a single major SL is *trans*-spliced onto ~50% of the transcripts in this species (Satou *et al.*, 2006). Also in another member of this phylum, *Oikopleura dioica*, only one type of SL has been found, which is *trans*-spliced to at least 25% of *O. dioica* mRNAs (Ganot *et al.*, 2004).

The novel SL sequences observed on the transcripts of *SECPEP1* and *Gr-gpd* are clearly different from the SL1 variants observed on transcripts of other nematodes of the suborder Tylenchina (Koltai et al., 1997, Stratford & Shields, 1994, Ray et al., 1994, Goyal et al., 2005). Sequence similarity search with the novel SLs showed that, except for the SL1-like sequences, no exact or near exact matches were found with the 5'-ends of cDNA sequences isolated from nematodes other than *Globodera rostochiensis.* Sequences identical to SL2f and SL4e are found on the 5'-ends of two library clones (respectively GRAA-aaa55c06 and GRAA-aaa68c10) derived from mRNA isolated from parasitic J2 from *Globodera rostochiensis*.

Sequence-based techniques, like nothern blotting or sequence homology search, are the most common way for investigating the phylogenetic dispersion of SL *trans*-splicing. The high level of sequence variance in the novel SL sequences may explain why they have not been observed before. However, an investigation of the current sequence database shows that out of 11,851 *Globodera rostochiensis* ESTs only 96 contain a SL-1 like sequence (van Bers, unpublished results). None of these sequences contain all the 22 nucleotides of the SL1 sequence; in all of them the ultimate 5'-end seems to be absent. This is striking as for the generation of 5906 *Globodera rostochiensis* ESTs the SMART technology was used, which should enrich the library for full-length sequences (Zhu *et al.*, 2001). However, a bias against *trans*-spliced ESTs upon use of the SMART technology has been suggested before when only

1 out of 2662 *Heterodera schachtii* ESTs was shown to have a SL sequence (Vanholme *et al.*, 2006). Out of 20,334 *Meloidogyne incognita* ESTs 1634 show sequence similarity to SL1 (E-value <0.01; van Bers, unpublished results), of which 1087 contain a complete SL1 sequence (22 nucleotide). However, 914 of these are derived from a library generated with a SL1 primer. Strikingly, none of the remaining 173 ESTs containing a full length SL1 sequence is derived from mRNA isolated from (pre)parasitic juveniles. In *C. elegans,* the 5'-end of *trans*-spliced mRNA consists of a tri-methylguanosine (TMG) cap, instead of the "usual" mono-methylguanosine cap (Thomas et al., 1988, Van Doren & Hirsh, 1988). It remains to be investigated whether this alternative cap hampers the oligonucleotide extension which is used in the SMART technology, or if the 5'-cap of *trans*-spliced mRNA of (pre)parasitic juveniles of plant parasitic nematodes inhibits cDNA cloning.

In *C. elegans*, the use of SL2 is restricted to mRNA of genes that are located downstream in an operon, although operons have been observed in nematode species that seem to lack SL2 *trans*-splicing (Blumenthal et al., 2002) and reviewed by (Guiliano & Blaxter, 2006). SL2 may function in protecting the downstream mRNA from degradation upon 3'-end formation of the upstream gene (reviewed in (Blumenthal, 2004). *Trans*-splicing has allowed the evolution of polycistronic transcription units and in that way creates a possibility for reducing genome size (Zorio et al., 1994). Interestingly, polycistronic transcription has also been observed for the chordate *Oikopleura dioica*, which has a small genome (<70 Mb)(Ganot et al., 2004). *Globodera rostochiensis* also has a small genome (estimated size 80 Mb; Qin, unpublished data), so it is tempting to speculate on the existence of polycistronic transcription units in *Globodera*. Operons have been found for another member of the suborder Tylenchina; the animal parasitic nematode *Strongyloides ratti* (Guiliano & Blaxter, 2006).

The novel SL sequences are more similar to SL2-like sequences and SL2 RNA than to SL1 and SL1 RNA, which is indicated by their sequence variance, the secondary structure of the corresponding SL RNA genes, and the clustered genomic organization

of the novel SL RNA genes. Twelve percent of the *Gr-gpd* transcripts are *trans*-spliced to a SL different from SL1 or SL-1 like sequences. Interestingly, in *C. elegans*, *gpd-2* is located downstream in an operon, and ten percent of the *gdp-2* transcripts are *trans*-spliced to SL2 (Spieth et al., 1993). The use of a SL sequence completely different from SL1 is not specific for *Globodera* within the Tylenchina, but is also found for *Heterodera glycines* (Baum, personal communication). The use of highly polymorphic, non-canonical SL-sequences on mRNAs of single gene was recently also found for *Trichinella spiralis* (Pettitt et al., 2008). This may imply that the conserved SL1 sequence found in the Rhabditina evolved after divergence of the major nematode clades (Pettitt et al., 2008).

About twenty-five percent of the *SECPEP1* transcripts are *trans*-spliced to a SL from the Gr-SL2, Gr-SL3 or Gr-SL4 cluster, which leads us to speculate that *SECPEP1* is localized downstream in an operon. The single band on the southern blot suggests that *SECPEP1* is a single copy gene. However, we hypothesize that *SECPEP1* is in fact a multiple copy gene that is localized in a tandem repeat. This would explain the sequence variation observed for *SECPEP1* genomic DNA, while the organization in a tandem repeat would still account for the single band observed on the southern blot. For the analysis of the sequence variation in *SECPEP1* gDNA we used DNA isolated from a population of nematodes. Therefore, the observed variation could be attributed to both genetic variation between individual nematodes and to allelic variation. However, amplification of *SECPEP1* gDNA from a single nematode results in more than two *SECPEP1* sequences (van Bers, unpublished data), while *G. rostochiensis* is diploid. These results support the hypothesis that *SECPEP1* is a multiple copy gene. The band observed on the southern blot of *EcoRI* digested gDNA was about 1650 bp. One *EcoRI* restriction site is present in the *SECPEP1* coding sequence. Interestingly, we could amplify a band of about 1650 bp when using two *SECPEP1* primers that were pointing outwards from eachother (van Bers, unpublished data), which suggests that *SECPEP1* is indeed organized in a tandem repeat.

SECPEP1 is a member of a highly divergent family of genes encoding positively charged peptides. Sequence similarity between the nine members of this family (*SECPEP1-SECPEP9)* is limited to the signal peptide and 3'UTR region. The mature peptides, which are delivered into the host tissue, are highly divergent between the family members. This extreme sequence variation was partly explained by positive selection acting on some residues of the mature peptides (van Bers *et al., in prep.*). The transcriptional analysis presented here showed that the number of non-synonymous nucleotide changes observed in *SECPEP1* cDNA is significantly higher than observed for *SECPEP1* gDNA or *Gr-gpd* cDNA. This indicates that besides positive selection an additional mechanism, e.g. RNA editing, may contribute to SECPEP1 diversity. RNA editing results in insertion or deletion of nucleotides, or in the conversion of one base to another (reviewed by (Gott & Emeson, 2000). Many of the nucleotide conversions are thought to arise from deamination (A-to-I and C-to-U) and amination (U-to-C) reactions. The translational machinery and reverse transcriptases read inosine as if it were guanosine and uracil as thymine, resulting in discrepancy between the genomic and cDNA sequence of a gene (Nishikura, 2006). The majority of the non-synonymous changes observed for *SECPEP1* result in charge or polarity changes of the mature peptide, while the few non-synonymous changes observed for *Gr-gpd* are all conservative changes. This indicates a selective pressure towards rapid change of the SECPEP1 peptide which may point towards a role for SECPEP1 in host pathogen interactions.

The results presented in this chapter indicate that *SECPEP1* is *trans*-spliced to SL2-like spliced leader sequences, which is novel for the suborder Tylenchina. This shows that SL-2-like *trans*-splicing within the phylum Nematoda is more widespread than previously anticipated (Guiliano & Blaxter, 2006). Currently, we are performing a BAC-library screening to further investigate the *SECPEP1* genomic organization.

References

Agabian, N. (1990) Trans splicing of nuclear pre-mRNAs. *Cell,* **61,** 1157-1160.

Bektesh, S., Van Doren, K. and Hirsh, D. (1988) Presence of the Caenorhabditis elegans spliced leader on different mRNAs and in different genera of nematodes. *Genes & development,* **2,** 1277-1283.

Blaxter, M. and Liu, L. (1996) Nematode spliced leaders - Ubiquity, evolution and utility. *International Journal for Parasitology,* **26,** 1025-1033.

Blumenthal, T. (1995) Trans-splicing and polycistronic transcription in Caenorhabditis elegans. *Trends Genet,* **11,** 132-136.

Blumenthal, T. (2004) Operons in eukaryotes. *Brief Funct Genomic Proteomic,* **3,** 199-211.

Blumenthal, T., Evans, D., Link, C. D., Guffanti, A., Lawson, D., Thierry-Mieg, J.*, et al.* (2002) A global analysis of Caenorhabditis elegans operons. *Nature,* **417,** 851-854.

Blumenthal, T. and Thomas, J. (1988) Cis and trans mRNA splicing in C. elegans. *Trends in Genetics,* **4,** 305-308.

Bruzik, J. P., Van Doren, K., Hirsh, D. and Steitz, J. A. (1988) Trans splicing involves a novel form of small nuclear ribonucleoprotein particles. *Nature,* **335,** 559-562.

Cheng, G., Cohen, L., Mikhli, C., Jankowska-Anyszka, M., Stepinski, J., Darzynkiewicz, E.*, et al.* (2007) In vivo translation and stability of trans-spliced mRNAs in nematode embryos. *Mol Biochem Parasitol*.

Curran, J., Baillie, D. L. and Webster, J. M. (1985) Use of genomic DNA restriction fragment length differences to identify nematode species. *Parasitology,* **90,** 137-144.

Davis, R. E. (1997) Surprising diversity and distribution of spliced leader RNAs in flatworms. *Mol Biochem Parasitol,* **87,** 29-48.

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.

De Rijk, P., Wuyts, J. and De Wachter, R. (2003) Rna Viz 2: An improved representation of RNA secondary structure. *Bioinformatics,* **19,** 299-300.

Evans, D. and Blumenthal, T. (2000) trans splicing of polycistronic Caenorhabditis elegans pre-mRNAs: analysis of the SL2 RNA. *Mol Cell Biol,* **20,** 6659-6667.

Ganot, P., Kallesoe, T., Reinhardt, R., Chourrout, D. and Thompson, E. M. (2004) Spliced-leader RNA trans splicing in a chordate, Oikopleura dioica, with a compact genome. *Mol Cell Biol,* **24,** 7795-7805.

Gheysen, G. and Jones, J. T. (2006) Molecular Aspects of Plant-Nematode Interactions. In: *Plant Nematology.* (Perry, R. N. and Moens, M., ed.^eds.). Wallingford: CABI, pp. 234-254.

Gott, J. M. and Emeson, R. B. (2000) Functions and mechanisms of RNA editing. *Annu Rev Genet,* **34,** 499-531.

Goyal, K., Browne, J. A., Burnell, A. M. and Tunnacliffe, A. (2005) Dehydration-induced tps gene transcripts from an anhydrobiotic nematode contain novel spliced leaders and encode atypical GT-20 family proteins. *Biochimie,* **87,** 565-574.

Guiliano, D. B. and Blaxter, M. L. (2006) Operon conservation and the evolution of trans-splicing in the phylum nematoda. *PLoS Genetics,* **2,** 1871-1882.

Holterman, M., Van Der Wurff, A., Van Den Elsen, S., Van Megen, H., Bongers, T., Holovachov, O.*, et al.* (2006) Phylum-wide analysis of SSU rDNA reveals deep phylogenetic relationships among nematodes and accelerated evolution toward crown clades. *Molecular Biology and Evolution,* **23,** 1792-1800.

Huang, X.-Y. and Hirsch, D. (1989) A second trans-spliced RNA leader sequence in the nematode Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America,* **86,** 8640-8644.

Koltai, H., Spiegel, Y. and Blaxter, M. L. (1997) Regulated use of an alternative spliced leader exon in the plant parasitic nematode Meloidogyne javanica. *Molecular and Biochemical Parasitology,* **86,** 107-110.

Krause, M. and Hirsh, D. (1987) A trans-spliced leader sequence on actin mRNA in C. elegans. *Cell,* **49,** 753-761.

Lall, S., Friedman, C. C., Jankowska-Anyszka, M., Stepinski, J., Darzynkiewicz, E. and Davis, R. E. (2004) Contribution of trans-splicing, 5' -leader length, cap-poly(A) synergism, and initiation factors to nematode translation in an Ascaris suum embryo cell-free system. *J Biol Chem,* **279,** 45573-45585.

Lambert, K. N., Allen, K. D. and Sussex, I. M. (1999) Cloning and characterization of an esophageal-gland-specific chorismate mutase from the phytoparasitic nematode Meloidogyne javanica. *Mol Plant Microbe Interact,* **12,** 328-336.

Liu, Y., Huang, T., MacMorris, M. and Blumenthal, T. (2001) Interplay between AAUAAA and the trans-splice site in processing of a Caenorhabditis elegans operon pre-mRNA. *RNA,* **7,** 176-181.

Liu, Y., Kuersten, S., Huang, T., Larsen, A., MacMorris, M. and Blumenthal, T. (2003) An uncapped RNA suggests a model for Caenorhabditis elegans polycistronic pre-mRNA processing. *Rna,* **9,** 677-687.

Maroney, P. A., Denker, J. A., Darzynkiewicz, E., Laneve, R. and Nilsen, T. W. (1995) Most mRNAs in the nematode Ascaris lumbricoides are trans-spliced: A role for spliced leader addition in translational efficiency. *RNA,* **1,** 714-723.

Nilsen, T. W. (1997) *Trans*-splicing. In: *Eukaryotic mRNA processing.* (Krainer, A., ed.^eds.). Oxford: IRL Press, pp. 310-334.

Nilsen, T. W., Shambaugh, J., Denker, J., Chubb, G., Faser, C., Putnam, L.*, et al.* (1989) Characterization and expression of a spliced leader RNA in the parasitic nematode Ascaris lumbricoides var. suum. *Molecular and Cellular Biology,* **9,** 3543-3547.

Nishikura, K. (2006) Editor meets silencer: Crosstalk between RNA editing and RNA interference. *Nature Reviews Molecular Cell Biology,* **7,** 919-931.

Pettitt, J., Muller, B., Stansfield, I. and Connolly, B. (2008) Spliced leader trans-splicing in the nematode Trichinella spiralis uses highly polymorphic, noncanonical spliced leaders. *RNA***,** rna.948008.

Pouchkina-Stantcheva, N. N. and Tunnacliffe, A. (2005) Spliced leader RNA-mediated trans-splicing in phylum Rotifera. *Mol Biol Evol,* **22,** 1482-1489.

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.

Rajkovic, A., Davis, R. E., Simonsen, J. N. and Rottman, F. M. (1990) A spliced leader is present on a subset of mRNAs from the human parasite Schistosoma mansoni. *Proceedings of the National Academy of Sciences of the United States of America,* **87,** 8879-8883.

Ray, C., Abbott, A. G. and Hussey, R. S. (1994) Trans-splicing of a Meloidogyne incognita mRNA encoding a putative esophageal gland protein. *Molecular and Biochemical Parasitology,* **68,** 93-101.

Redmond, D. L. and Knox, D. P. (2001) Haemonchus contortus SL2 trans-spliced RNA leader sequence. *Mol Biochem Parasitol,* **117,** 107-110.

Ross, L. H., Freedman, J. H. and Rubin, C. S. (1995) Structure and expression of novel spliced leader RNA genes in Caenorhabditis elegans. *J Biol Chem,* **270,** 22066-22075.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*. New York.

Satou, Y., Hamaguchi, M., Takeuchi, K., Hastings, K. E. and Satoh, N. (2006) Genomic overview of mRNA 5'-leader trans-splicing in the ascidian Ciona intestinalis. *Nucleic Acids Res,* **34,** 3378-3388.

Spieth, J., Brooke, G., Kuersten, S., Lea, K. and Blumenthal, T. (1993) Operons in C. elegans: polycistronic mRNA precursors are processed by trans-splicing of SL2 to downstream coding regions. *Cell,* **73,** 521-532.

Stein, L. D., Bao, Z., Blasiar, D., Blumenthal, T., Brent, M. R., Chen, N.*, et al.* (2003) The genome sequence of Caenorhabditis briggsae: A platform for comparative genomics. *PLoS Biology,* **1**.

Stover, N. A. and Steele, R. E. (2001) Trans-spliced leader addition to mRNAs in a cnidarian. *Proc Natl Acad Sci U S A,* **98,** 5693-5698.

Stratford, R. and Shields, R. (1994) A trans-spliced leader RNA sequence in plant parasitic nematodes. *Mol Biochem Parasitol,* **67,** 147-155.

Thomas, J. D., Conrad, R. C. and Blumenthal, T. (1988) The C. elegans trans-spliced leader RNA is bound to Sm and has a trimethylguanosine cap. *Cell,* **54,** 533-539.

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

Turner, S. J. and Evans, K. (1998) The origins, global distribution and biology of potato cyst nematodes (*Globodera rostochiensis* (Woll.) and *Globodera pallida* Stone). In: *Potato cyst nematodes: biology, distribution and control.* (Marks, R. J. and Brodie, B. B., ed.^eds.). Wallingford: CAB International, pp. 7-26.

Van Doren, K. and Hirsh, D. (1988) Trans-spliced leader RNA exists as small nuclear ribonucleoprotein particles in Caenorhabditis elegans. *Nature,* **335,** 556-559.

Vandenberghe, A. E., Meedel, T. H. and Hastings, K. E. (2001) mRNA 5'-leader trans-splicing in the chordates. *Genes Dev,* **15,** 294-303.

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

von Mende, N., Gravato-Nobre, M. J. and Perry, R. N. (1998) Host finding, invasion and feeding. In: *The cyst nematodes.* (Sharma, S., ed.^eds.). Dordrecht: Kluwer Academic Publishers.

Zayas, R. M., Bold, T. D. and Newmark, P. A. (2005) Spliced-leader trans-splicing in freshwater planarians. *Molecular Biology and Evolution,* **22,** 2048-2054.

Zhang, H., Hou, Y., Miranda, L., Campbell, D. A., Sturm, N. R., Gaasterland, T.*, et al.* (2007) Spliced leader RNA trans-splicing in dinoflagellates. *Proc Natl Acad Sci U S A,* **104,** 4618-4623.

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

Zorio, D. A., Cheng, N. N., Blumenthal, T. and Spieth, J. (1994) Operons as a common form of chromosomal organization in C. elegans. *Nature,* **372,** 270-272.

Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research,* **31,** 3406-3415.

Chapter 6

General discussion: Peptides in plant-pathogen interactions

Nikkie van Bers, Aska Goverse, Jaap Bakker and Geert Smant

To be submitted

Abstract

A wide range of plant-parasites interfere with existing plant signaling pathways, i.e. to circumvent recognition by the host defense system and for the formation of sophisticated feeding structures. Interestingly, a considerable number of the effector molecules delivered by parasites into host plant cells is a peptide. In this thesis, we described two novel classes of peptide effectors believed to be secreted by the potato cyst nematode, the highly divergent SECPEP family, with at least nine members in the potato cyst nematode, and the NEMPEP peptide. In this conclusive chapter, we discuss similarities observed among a wide range of peptides important in parasitism on plants, and we summarize insights into the mechanisms driving the molecular evolution of genes coding for effectors. We conclude that a majority of the peptides involved in plant-microbe interactions is positively charged and hypervariable. These are characteristics shared with the “host defense polypeptides”, a class of peptides widely distributed among eukaryotes.

Introduction

Effectors and the plant immune system

Generally speaking, effectors are molecules produced by a pathogen, that are intended to contribute to its virulence, but that do not have a "housekeeping" function in pathogen growth and development (Bent & Mackey, 2007). Although exceptions exist and the difference is becoming less clearly defined, the dispensability of effectors discriminates them from MAMPs (microbe associated molecular patterns) which are defense elicitors that form a core component of the micro-organism and are therefore often evolutionary stable (Bent & Mackey, 2007). Recognition of pathogen effectors by R-gene products can be described by two models, concerning respectively a direct and an indirect interaction between a R-gene and its cognate effector (reviewed by (Bent & Mackey, 2007). Direct recognition is similar to ligand(effector)-receptor(R-gene) interaction, resulting in the manifestation of a host defense response (Dodds *et al.*, 2006). The indirect interaction model is also known as the guard hypothesis, in which presence of the pathogen is sensed by R-gene products through a modified state of other (guarded) host proteins (Van der Hoorn *et al.*, 2002)). These host proteins may be the virulence target of the effector protein. Natural selection drives pathogens to avoid recognition by change or even complete loss of effector genes.

Peptides and host defense peptides

Peptide signaling plays an important role in several signaling processes in plants, e.g. in mediating cell-to-cell communication, in self/non-self recognition, in determining cell identity and in host defense (reviewed in (Farrokhi *et al.*, 2008, Matsubayashi & Sakagami, 2006)). Several definitions on the concept "peptide" have been formulated. In this review we refer to peptides as molecules consisting of fewer than 150 amino-acids. For proteins that are synthesized as a preprotein, we only considered the mature peptide. The term host defense peptides encompasses a wide range of peptides that generally share a small size, a high level of sequence diversity and a positive net

charge at pH 7. Examples of host defense peptides are antimicrobial peptides (AMPs) and neurotoxic peptides. AMPs are found in virtually all organisms, and function in, among others, defense against pathogens, in modulating immune responses and in promoting woundhealing (Brown & Hancock, 2006, Peschel & Sahl, 2006, Silverstein *et al.*, 2007). Neurotoxic peptides are found in the venom of Conus snails, scorpion and spiders and play a role in defense and in predation (reviewed by (Olivera & Cruz, 2000))(Zeng *et al.*, 2005, Escoubas, 2006, Terlau & Olivera, 2004). The positive charge of host defense peptides is believed to be important for the interaction with negatively charged groups, i.e. on the microbial cell surface (Yeaman & Yount, 2003).

Recently, Yeaman and Yount discovered that a majority of cysteine containing host defense peptides share a tri-residue C-X-G γ-core motif (Yeaman & Yount, 2007, Yount & Yeaman, 2004). The C-X-G γ-core motif can occur in either forward (dextromeric) or in reverse (levomeric) orientation. It is a structural motif, predicted to result in the formation of a loop structure, similar in shape to the Greek letter γ. The exact importance of the motif is unknown, but a role in mediating membrane or receptor activation has been suggested (Yeaman & Yount, 2007). Host defense peptides often show a high level of sequence divergence, which is limited to or predominant in the mature peptide region. For several of the peptide genes this hypervariability in the mature peptide region was attributed to positive selection pressure (a.o. (Tennessen, 2005, Win *et al.*, 2007, Silverstein *et al.*, 2005)).

Yount and coworkers searched the database for peptides having a C-x-G γ-core motif, and observed the motif also in peptides that were not previously anticipated to function in host defense. Strikingly, all of the peptides tested were found to exert direct antimicrobial activity (Yount & Yeaman, 2004). Boman (Boman, 2003) suggested to use the level hydrophilicity of peptides to discriminate between relatively hydrophobic "classical" antimicrobial peptides, that only exert antimicrobial activity, and hydrophilic peptide hormones which act as signaling molecules by interacting with receptors (protein-protein interaction). The GRAVY score (Kyte & Doolittle, 1982) is a method to

numerically express the hydrophilicity of a protein, with a mean value of -0.4 kcal/mol reported for soluble proteins and values exceeding -0.4 for more hydrophobic, membrane-spanning, proteins. In analogy to suggested by Boman, the GRAVY score may give an indication about the mode of action of a peptide, e.g. membrane or protein-protein interaction.

In this chapter, we investigate whether peptides involved in plant-pathogen interactions show characteristics typical for host defense peptides, like a positive net charge, hypervariability, and a C-x-G γ-core motif. Thereto, we generated an overview with these characteristics and the GRAVY score of a wide range of peptides involved in plant-pathogen interactions. We investigate whether the GRAVY score reflects the mode of action of the peptides, e.g. protein-protein or protein-membrane interaction, which could help in predicting the mode of action of pioneer effectors.

Plant peptides

Systemin

The 18 amino-acid peptide systemin (charge +2) was first isolated from tomato leaves (Pearce *et al.*, 1991). Systemin derives by proteolytic processing from the C-terminus of a 200 amino-acids precursor protein (McGurl *et al.*, 1992). At low concentrations, systemin induces the synthesis of two wounding inducible proteinase inhibitor proteins that may be involved in plant defense (Pearce et al., 1991). The ability to elicit effects at very low concentrations classes systemin as a peptide hormone (Pearce *et al.*, 2001).

Systemin peptides have been found in several other species of the family Solanaceae, e.g. potato, pepper and black nightshade (Constabel *et al.*, 1998, Pearce et al., 2001). However, the systemin isolated from black nightshade (*Solanum nigrum)* has a tenfold lower capacity of inducing proteinase inhibitors than tomato systemin, and does not mediate direct defense responses (Constabel et al., 1998, Schmidt &

Baldwin, 2006), suggesting a different function for this systemin. The sequences from the potato, pepper and black nightshade systemin precursor gene, prosystemin, share 73-80% sequence identity with tomato prosystemin. This is in contrast to the high level of sequence conservation in peptide hormones in animals, which led Boller (Boller, 2005) to suggest that systemin-type signaling may be subject to diversifying selection.
In tobacco, systemin-like peptides are found, which are named TobHypSysI and II. TobHypSysI and II show no sequence similarity to either tomato systemin nor to its precursor, prosystemin, but are, like all systemin-like peptides, rich in P, O, S and T residues {reviewed by (Ryan & Pearce, 2003). The acronym TobHypSys derives from tobacco hydroxyproline rich systemins, and TobHypSys I and II are processed from a single, 165 amino-acids precursor protein (Pearce et al., 2001). Three tomato homologous of TobHypSys have been found, TomHypSys I, II and III, which encode for 15-20 amino-acids peptides. TomHypSys I, II and III are all processed from a single, 148 amino-acids, precursor protein (Ryan & Pearce, 2003). The tomato and tobacco HypSys peptides induce the synthesis of defensive proteinase-inhibitor proteins in a manner similar to that of systemin.

The tomato prosystemin gene has a repetitive structure, and consists of five homologous pairs of exons, in which the first exons of the pairs are homologous to each other and the same is true for the second exons of the pairs. The prosystemin gene also contains one non-homologous exon, which encodes systemin (McGurl & Ryan, 1992). This organization suggests that prosystemin has evolved from a much smaller ancestral gene through duplication and elongation events. The C-terminal region, encoding the systemin peptide, is the most conserved region of prosystemin (Constabel et al., 1998).

Systemin signaling is believed to be mediated through protein-protein interaction between the peptide and Leucine Rich Repeat (LRR)-receptor-like kinase, which causes a rapid activation of a MAP kinase (Kandoth *et al.*, 2007, Stratmann *et al.*, 2000). This mode of action by receptor-ligand interaction is reflected in the

negative GRAVY score of systemin (-1.72 kcal/mol, Table 1), which classes systemin as a hydrophilic protein.

AtPEP1

Recently, a novel defense related peptide was discovered in *Arabidopsis*, *At*PEP1 (Huffaker *et al.*, 2006). *At*PEP1 is a 23 amino-acid peptide that induces the synthesis of H_2O_2, and activates the expression of the defensin *PDF1.2* and of its own precursor gene through the jasmonate/ethylene signaling pathway (Huffaker et al., 2006, Huffaker & Ryan, 2007). *At*PEP1 is a positively charged, hydrophilic protein (GRAVY score of -1.69 kcal/mol, Table1) and serves as a ligand for the cell-surface receptor PEPR1, which is a LRR-kinase (Huffaker & Ryan, 2007). The peptide derives from the C-terminus of a 92 amino-acid precursor protein, PROPEP1. PROPEP1 does not contain a signal peptide for secretion, indicating that *At*PEP1 functions intracellularly (Huffaker et al., 2006).

Six homologs of the *PROPEP1* gene have been identified in *Arabidopsis thaliana*, and in addition to *PROPEP*1 at least also *PROPEP*2 and *PROPEP*3 appear to function in a feedback loop that amplifies defense signaling pathways initiated by pathogens (Huffaker & Ryan, 2007). Constitutive overexpression of *PROPEP1* and *PROPEP2* enhances the resistance of *Arabidopsis* plants against the pathogen *Pythium irregulare* (Huffaker et al., 2006). Sequence comparison shows that the *PROPEP* genes share a low overall level of amino-acid similarity (12%-47%), with highest levels (35%-65%) found in the C-terminal region, where *At*PEP1 resides in PROPEP1 (Huffaker et al., 2006).

Antimicrobial peptides in plants

Recently, Silverstein and coworkers (Silverstein et al., 2007) identified ~13,000 plant genes within the Uniprot dataset and the *Arabidopsis thaliana* and *Oryza* sativa genomes, encoding divergent, charged or polar peptides with conserved cysteine

residues. This number included several classes of peptides that exert antimicrobial activity, and here we will briefly describe the four most abundant classes (in number of ESTs) in their dataset, i.e. thionins, defensins, lipid transfer proteins and heveins.

Thionins are small cysteine-rich and often positively charged peptides, implicated in plant defense against pathogenic invaders (reviewed by (Stec, 2006)). Plant thionins can be divided into two distinct groups, α/ß-thionins and γ-thionins. The latter should more appropriately be called plant defensins (see below). More than 100 thionins were identified in 15 plant species, and for an overview and discussion of the mechanism of toxicity we refer to the review by (Stec, 2006). Here we limit us to describing one example of the α/ß-thionins, α-hordothionin (Table 1). α-Hordothionin was isolated from Barley seed (Ponz *et al.*, 1986), and is believed to inhibit fungal growth as a result of direct protein-membrane interactions (Thevissen *et al.*, 1996). The α-hordothionin precursor protein consists of a N-terminal signal peptide for secretion, a C-terminal acidic region which is removed post-translationally, and a mature peptide of 45 amino-acids with a high positive net charge of +10 at pH 7 (Ponz et al., 1986, Rodríguez-Palenzuela *et al.*, 1988)(Table 1). Divergence between different types of thionins in wheat has occurred through a process of accelerated evolution, also known as diversifying selection, which has affected the amino-acid sequence of the mature thionin, but not of the precursor domains (Castagnaro *et al.*, 1992). Plant thionins have highly conserved C-X-G γ-core signatures ((Yeaman & Yount, 2007)).

Plant defensins were originally referred to as γ-thionins because they have a similar size and the same number of disulfide bridges as α/ß-thionins. However, later NMR studies showed that γ-thionins share structural properties with mammalian and insect defensins and should therefore be regarded as plant defensins, which are unrelated to α/ß-thionins (Broekaert *et al.*, 1995, Stec, 2006). Plant defensins have mainly been identified from seeds and leaves, and typically contain several disulfide-linked cysteines. Rs-AFP2 and Dm-AMP1 (Table 1), isolated from respectively radish and dahlia seed, are two plant defensins. Thevissen and coworkers (Thevissen et al., 1996)

showed that these defensins do not exert antifungal action by direct protein-membrane interactions, but may act via a receptor mediated mechanism. This is reflected in the GRAVY score of DM-AMP1, which is -0.69 kcal/mol, indicating it is a hydrophilic protein (Table 1). Hydrophilicity facilitates protein-protein interaction, while it hampers interaction with (hydrophobic) membranes.

Plant genome initiatives give insight into the abundance and diversity of peptide genes in the plant kingdom. Silverstein and coworkers revealed the existence of 317 defensin-like genes (DEFLs) in the *Arabidopsis thaliana* genome (Silverstein et al., 2005). Nearly all DEFLs are composed of two exons, of which the first one encodes a signal peptide for secretion and the second exon codes for the mature peptide of 60-70 amino-acids. For defensins, the mature peptide results from co-translational cleavage of a signal peptide for secretion. Many of the DEFLs occur in the Arabidopsis genome as clusters, which have evolved by duplication events, followed by selection. Within a cluster, the N-terminal signal peptides targeting the DEFLs for secretion into the apoplast are highly conserved, even as intron position and intron size. However, depending on the gene cluster, the mature peptides are showing evidence of either purifying or diversifying selection (Silverstein et al., 2005).

Plant non-specific lipid transfer proteins (nsLTPs) are positively charged peptides that inhibit the growth of both bacterial and fungal pathogens. The mode of action of nsLTPs is unknown, but they may insert themselves in fungal membranes and form a pore that results in ion leakage and cell death, which is a common mode of action for antimicrobial peptides (reviewed by (De Lucca *et al.*, 2005)). The hypothesized membrane interaction potential of nsLTPs is sustained by a GRAVY score of +0.24 kcal/mol for IWF1(Table 1), which is a nsLTPs isolated from sugar beet leaves (Nielsen *et al.*, 1996).

Hevein is an antifungal peptide isolated from the rubber tree *Hevea brasiliensis* (Van Parijs *et al.*, 1991, Rozynek *et al.*, 1998). Hevein, and hevein-like peptides exert antifungal activity by binding to chitin, which is a constituent of the fungal cell wall

(Broekaert *et al.*, 1992, Van Parijs et al., 1991). Ac-AMP1 is a hevein-like peptide, consisting of 29 amino-acids, isolated from *Amaranthus caudatus* (Table 1). Ac-AMP1 contains six cysteine residues, and is highly homologous to a cysteine/glycine rich domain occurring in many chitin-binding proteins (Broekaert et al., 1992). Ac-AMP1 was first identified in *A. caudatus* seeds by direct amino-terminal sequencing. Therefore, the size and character of this precursor protein remain to be investigated. Processing may be similar as observed for the hevein peptide, which is processed from a 204 amino-acid precursor protein, prohevein (Rozynek et al., 1998). The 17 N-terminal residues of prohevein encode a signal peptide for secretion, which is followed by the 43 amino-acids of the hevein domain. Ac-AMP1 is highly homologous to AC-AMP2, another hevein-like molecule from *A. caudatus*, which is suggested to be produced from the same precursor protein (Broekaert et al., 1992).

Bacterial peptides

Phytopathogenic bacteria, such as *Pseudomonas syringae*, use type III effector proteins to inhibit host defenses, to grow in plants and to produce disease lesions (Abramovitch & Martin, 2004, Alfano & Collmer, 2004). These effectors (e.g. *AvrPto*, *AvrB*, *AvrRPM1* and *HopPtoD2*) generally consist of >250 amino-acids, which is in striking contrast to the effector proteins secreted by eukaryotic pathogens of plants. However, a bacterial type I effector, AvrXa21, is hypothesized to be a GG leader peptide involved in plant-microbe interactions. GG leader peptides are secreted by the bacterial type I secretion system, and exhibit anti-bacterial activity. The peptides share little sequence similarity, but share a GG-leader sequence, which is required for secretion (Da Silva *et al.*, 2004). AvrXa21 activity was detected in *Xanthomonas oryzae* and *X. campestris*, and elicits a defense response upon recognition by the plant receptor kinase Xa21 (Song *et al.*, 1995, Da Silva et al., 2004). The role of the peptide in promoting virulence of the pathogen is likely to be as an antimicrobial toxin (bacteriocidin) restraining the growth of related bacteria occupying the same niche (Da Silva et al., 2004). Also a role in

quorum-sensing was suggested for AvrXa21 (Lee *et al.*, 2006). The exact nature and size of AvrXa21 is still unknown, and it needs to be confirmed whether AvrXa21 is actually a protein, as a combination of heat and proteinase treatment only slightly reduced its activity (Lee et al., 2006). AvrXa21 activity depends on factors of the type I secretion system associated with secretion of GG-leader peptides, which led to the belief that AvrXa21 is a GG-leader peptide (Da Silva et al., 2004). AvrXa21 activity is likely to be conserved between different *Xanthomonas oryzae* strains, and was detected in at least two *Xanthomonas* species. This conserved nature is a characteristic of a MAMP, which makes AvrXa21 a molecule that shares characteristics of a MAMP and of an effector molecule (Lee et al., 2006).

A positively charged peptide from *Xanthomonas axopodis* shows significant structural and sequence similarity to the *Arabidopsis thaliana* PNP-A protein (Nembaware *et al.*, 2004). PNP stands for plant natriuretic peptides, and PNPs regulate several homeostatic processes in plants, e.g. stomatal opening, modulation of ion fluxes and H_2O uptake in proplasts (Nembaware et al., 2004). The PNP-like peptide from *X. axopodis* (Table 1) does not show sequence similarity to other bacterial proteins, and its presence in *X. axopodis* was contributed to lateral gene transfer (Nembaware et al., 2004). The role of the PNP-like peptide in promoting bacterial virulence is unknown, but is likely correlated to plant hydration and the formation of wet lesions (Nembaware et al., 2004).

Fungal peptides

AVR from Cladosporium fulvum

Cladosporium fulvum is a non-obligate biotrophic fungus that causes leaf mold on tomato (Thomma *et al.*, 2005). Conidia of the fungus can infect tomato plants if they settle and germinate on the abaxial side of the leaf. The resulting hyphae enter the plant through open stomata and continue their growth preferentially towards the vascular tissue (Thomma et al., 2005). To date, eight effector genes have been cloned and

characterized from *C. fulvum*, and all code for secreted peptides with an even number of cysteine residues. None of the *C. fulvum* effectors have significant sequence homology with eachother or with any gene or protein in the database (Luderer *et al.*, 2002a).

Avr9 was the first fungal effector gene resulting in avirulence cloned. The corresponding effector peptide, avr9, mediates resistance on tomato plants harboring the *Cf*-9 resistance gene (Van Den Ackerveken *et al.*, 1992). A gene-for-gene basis may underlie the *avr9-Cf-9* interaction (Van Den Ackerveken et al., 1992). Although suggested, a direct interaction between AVR9 and Cf-9 has never been detected (Luderer & Joosten, 2001). The *Avr9* gene encodes a 63 amino-acids precursor protein that is processed in two steps into a 28 amino-acid mature peptide (AVR9)(Table 1). The first processing step involves the removal of a 23 amino-acid signal peptide, and the second step involves *in-planta* processing by plant proteases (Van Den Ackerveken *et al.*, 1993b). AVR9 is structurally homologous to a carboxy peptidase inhibitor found in potato, and contains a typical cysteine knot motif, which is found in several small proteins (e.g. proteinase inhibitors, ion channel blockers and growth factors)(Vervoort *et al.*, 1997).

C. fulvum strains carrying the *Avr2* gene induce a defense response on tomato plants carrying the *Cf-2* resistance gene. This response is mediated via binding of AVR2 to, and inhibition of, the extracellular tomato cysteine protease Rcr3 (Rooney *et al.*, 2005). *Avr2* encodes a 78 amino-acid precursor protein (Luderer et al., 2002a), including a signal peptide for secretion of 19 amino-acids at its amino- terminus (Table 1).

AVR4 is recognized by the tomato *Cf-4* resistance gene, and binds very efficiently to chitin. Fungal cell walls contain chitin, and chitin binding may protect the fungal cell walls from the action of plant chitinases (Van Den Burg *et al.*, 2006, Westerink *et al.*, 2002). AVR4 is encoded by a 135 amino-acids precursor protein, which after removal of the signal peptide and N-terminal processing by plant or fungal

proteases results in a peptide of 105 amino-acids. Further C-terminal processing results in a mature peptide of 86 amino-acids (Table 1), although various intermediates have been found as well (Joosten *et al.*, 1997). The *Cf4* locus consists of five tandem duplicated genes. One of the *Cf4* homologs on this locus (*Cf-4E*) confers reistance to *C. fulvum* strains carrying the *Avr4E* gene (Takken *et al.*, 1999). AVR4E is a negatively charged, secreted, cystein-rich peptide (Westerink *et al.*, 2004), lacking a C-x-G γ-core motif (Table 1). This is in contrast to the other three *C. fulvum* AVR peptides described here, which are positively charged and contain a CxG γ-core motif (Table 1).

In the *C. fulvum Avr* genes, point mutations, insertions and deletions, sometimes even of a complete gene, are associated with the transition from avirulence to virulence, indicating adaptive evolution of the pathogen to cognate resistance genes (Stergiopoulos *et al.*, 2007, Luderer *et al.*, 2002b). Comparison of *Avr* genes of virulent and avirulent *C. fulvum* strains showed that in all cases most variation was found in the protein coding part of the sequences (Stergiopoulos et al., 2007). However, the type of change differed for each of the *Avr* genes, suggesting different modes of adaptation (Stergiopoulos et al., 2007). *Avr2* genes showed mainly indels leading to truncated, non-functional AVR2 proteins, but also the insertion of a transposable element into the *Avr2* open reading frame was observed to result in loss of avirulence activity (Stergiopoulos et al., 2007, Luderer et al., 2002b). Loss of a complete gene is the main strategy to overcome *Avr9* and *Avr4E* mediated recognition, although a *AVR4E* variant with two amino-acid changes, resulting in loss of recognition, was found (Westerink et al., 2004, Stergiopoulos et al., 2007). *Avr4* genes of virulent strains showed single nucleotide polymorphisms (SNPs), resulting in non-synonymous changes. Sequence diversification as a way to overcome host resistance is more effective for effectors that directly interact with a R-gene than for effectors that are recognized indirectly (Bent & Mackey, 2007, Dodds et al., 2006). For effectors that are recognized indirectly, point mutations do not change recognition as long as interaction with the guard protein keeps on existing. Therefore, for effector genes encoding indirectly

recognized effectors, gene loss is a more common strategy to overcome recognition (Bent & Mackey, 2007). This theory is sustained by the strategy observed for evading recognition of AVR2, which is indirectly recognized by *Cf-2* (see above). A direct interaction between AVR4 and Cf4 has been hypothesized to drive the mutations in the *Avr4* sequences (Stergiopoulos et al., 2007). The hypothesized peptide-receptor interaction is reflected in the negative GRAVY score (-0.78 kcal/mol) of AVR4. In contrast, AVR2 and AVR9 have higher GRAVY score (respectively –0.48 and –0.25 kcal/mol) which may reflect the suggested indirect recognition by the corresponding resistance genes.

ECP

C. fulvum strains made deficient in the production of the secreted ECP1 and/or ECP2 proteins were only poorly able to colonize leaf tissue of six weeks old tomato plants (Laugé *et al.*, 1997), indicating that these proteins are effectors. ECP stands for extracellular protein, and to date 4 ECPs have been cloned (ECP1, 2, 4 and 5). ECPs are recognized by tomato genotypes carrying the corresponding *Cf-ECP* resistance gene (Laugé *et al.*, 1998). ECP1 is a peptide of 65 residues long, which is processed from a precursor of 96 amino-acids. Processing is (probably) a two step process, one removing the signal peptide and the second one takes place *in planta* and is mediated by fungal or plant proteases (Van den Ackerveken *et al.*, 1993a)(Table 1). ECP2 derives from a 165 amino-acids precursor and removal of the signal peptide results in an intermediate of 143 amino-acids, which is processed into a 142 amino-acid mature protein in the plant (Van den Ackerveken et al., 1993a)(Table 1). The role of the ECPs in promoting virulence of the pathogen is unknown, but ECP1 and 2 may be involved in suppressing host defense responses. Deficiency in ECP2 resulted in lower levels of *in planta* secreted AVR4 and AVR9 proteins (Laugé et al., 1997). The *Ecp* genes of virulent and avirulent *C. fulvum* strains showed less sequence variation than the *Avr* genes of the fungus, and most mutations resulted in synonymous changes

(Stergiopoulos et al., 2007). The apparent sequence conservation in the *Ecp* genes suggests that they are essential for virulence of *C. fulvum*. Alternatively, the lower amount of variation was attributed to the absence of *Cf-Ecp* genes in commercial tomato cultivars (Stergiopoulos et al., 2007).

SIX

SIX1 (secreted in xylem 1) was first isolated as a protein from fungal origin in xylem sap of tomato plants infected with *Fusarium oxysporum* f. sp. *Lycopersici*, the causal agent of tomato wilt disease (Rep *et al.*, 2002). SIX1 confers avirulence to *F. oxysporum* strains infecting tomato plants harboring the *I3* resistance gene (Rep *et al.*, 2004). *SIX1* codes for a precursor protein of 248 amino-acids, containing a 21 amino-acid signal peptide for secretion. The SIX1 peptide isolated from tomato xylem sap migrates on SDS-PAGE as a 12 kDa protein, which is 18 kDa smaller than the predicted molecular mass of the precursor protein after cleavage of the signal peptide (30 kDa). This indicates that the SIX1 precursor is further processed (Rep et al., 2004). Peptide sequencing of SIX1 fragments from xylem sap shows that SIX1 is located in the central part of the precursor protein.

The *SIX1* gene is located only at only 8 kb distance from *SIX2* in the genome of *F. oxysporum*. *SIX2* also codes for a protein that is found in the xylem sap of *F. oxysporum* infected tomato plants. Several other genes encoding secreted proteins are positioned to this same chromosomal region, including the pseudogene *SIX1*-H, which is a truncated version of *SIX1* resulting from the insertion of a transposon (reviewed by (Van Der Does & Rep, 2007). Both *SIX1* and *SIX1-H* have undergone diversifying selection (Van Der Does & Rep, 2007). No homolog of *SIX1*, nor of most of the other genes directly flanking the *SIX1* locus, were found in other formae speciales (f. sp.) of *F. oxysporum*.

PWL

PWL (pathogenicity towards weeping lovegrass) proteins are secreted, hydrophilic, glycine-rich peptides, isolated from the fungus *Magnaporthe grisea* (Kang *et al.*, 1995). The host range of *M. grisea* is extensive, and includes several agronomically important cereals, e.g. rice, barley and wheat. However, individual isolates of the fungus are limited to infecting only a number of grass species. The *PWL* gene family consists of four members, *PWL1-4* (Kang et al., 1995). The role of the PWL peptides in promoting virulence of the pathogen remains to be elucidated, but transformants harboring *PWL2* were unable to infect weeping lovegrass (*Eragrostis curvula*), which is a host grass of *M. grisea* (Sweigard *et al.*, 1995). This indicates that presence of PWL2 (Table 1) limits the host range of *M. grisea* (Sweigard et al., 1995, Kang et al., 1995).

The PWL gene family is a rapidly evolving gene family, of which some of the members have undergone faster sequence changes than the rest of the genome (Kang et al., 1995). The sequence similarity between members of the family is consistently lower at amino-acid level than at nucleotide level, indicating that the genes are under diversifying selection (Kang et al., 1995). However, ClustalW analysis shows that the PWL peptides have lower sequence similarity in the region coding for the signal peptide (33%-52%) than in the region coding for the mature peptide (50%-78%). This is in contrast to the more conserved signal peptide region observed in several other classes of peptides under positive selection pressure (Win et al., 2007, Silverstein et al., 2005)(van Bers *et al,* chapter 2, this thesis). The PWL locus is bordered by repetitive DNA sequences, which may have played a role in the evolution of the *PWL* genes by recombination (Kang et al., 1995).

NIP1

NIP1 is a effector of 60 amino-acids of *Rhynchosporium secalis*, the causal agent of barley scald on barley (Table 1). Pathogenicity of *R. secalis* is associated with necrotic lesions on barley leaves (Araz & Maden, 2006). Injection of NIP1 into leaves of barley

causes the formation of scald-like lesions, and disruption of the NIP1 encoding gene slightly reduced fungal virulence on susceptible plants. These findings suggest a role for NIP1 in virulence of *R. secalis* (Wevelsiep *et al.*, 1991). At subnecrotic concentrations, NIP1 indirectly stimulates plasma membrane localized H^+-ATP-ase (Wevelsiep *et al.*, 1993, Van't Slot *et al.*, 2007), which may be involved in stomatal opening and impairment of water balance observed upon infection with *R. secalis*.

In addition to a role in virulence, NIP1 elicits the synthesis of defense related proteins in plants carrying the *Rrs1* resistance gene (Hahn *et al.*, 1993, Rohe *et al.*, 1995). NIP1 is encoded by the *NIP1* (*AvrRrs1*) gene, which influences cultivar specificity of *R. secalis*. A resistance response in barley carrying the *Rrs1* gene follows the gene-for-gene model, and correlates with the presence of *NIP1* in *R. secalis* (Laugé & De Wit, 1998). *NIP1* codes for a 82 amino-acid precursor protein, which is processed into the mature NIP1 peptide (Table 1) by removal of a 22 amino-acid signal peptide for secretion. Two other necrosis-inducing peptides have been found in *R.secalis*, NIP2 and NIP3, but their sequences have not been published (Wevelsiep et al., 1991, Schürch *et al.*, 2004). Schürch and coworkers (Schürch et al.) analyzed the *NIP1* sequence of 614 *R. secalis* isolates, and observed a surprisingly high deletion frequency of the *NIP1* gene (45% of the isolates). They sequenced *NIP1* of 196 isolates and found that diversifying selection is acting on *NIP1*. Sequence diversification was four times higher in the region coding for the mature peptide than in non-coding regions or in the region coding for the signal peptide. Assessment of a relationship between the *NIP1* sequence and pathogenicity showed that loss of avirulence was often achieved through deletion of *NIP1*, and rarely through point mutations.

AvrL567 and *HESPs*

Rust fungi form specialized feeding structures, called haustoria, which penetrate the host cell wall but remain separated from the host cytoplasm. Some of the alleles of the AVRL567 locus in the flax rust fungus *Melampsora lini* induce a hypersensitive response

on plants carrying the L5, L6 or L7 resistance gene (Dodds *et al.*, 2004). Twelve alleles, named L1 to L12, of the resistance gene *L* have been identified in flax (*Linum usitatissimum*) so far. These alleles are highly polymorphic, with, for example, 33 amino-acid polymorphisms between L6 and L11. 32 of these polymorphisms reside in the LRR region, a region associated with recognition of effector molecules. AvrL567 alleles are also highly polymorphic, with 35 polymorphic amino-acid sites found within the 127 amino-acids of the mature peptide of six rust strains (Dodds et al., 2006). The polymorphisms observed in the *AvrL567* alleles are the result of diversifying selection (Dodds et al., 2006). The limited host range of the *M. lini* may inflict a co-evolutionary arms race between *M. lini* and its host, because overcoming recognition is essential for survival of *M. lini* (Dodds et al., 2006). This arms race results in accelerated sequence diversification of both the *AvrL567* and the *L* gene (Dodds et al., 2006). AvrL567 interacts directly with L5 and L6, and sequence diversification is often observed as a mechanism to overcome recognition in effectors directly interacting with the R-gene product. (Dodds et al., 2006).

Recently, a novel set of 20 putative effectors, named HESPs after haustorically expressed secreted proteins, was discovered in *M. lini* (Catanzariti *et al.*, 2006). Interestingly, the majority of the proteins (12 out of 20) are peptides. For the alleles of two of the HESP genes, *AvrP4* and *AvrM*, significantly more nucleotide substitutions were found in the coding regions as compared to the flanking regions of the genes. This indicates that these genes are under diversifying selection pressure (Catanzariti et al., 2006). The *AvrM* product may interact directly with the M resistance protein (Dodds *et al.*, 2007). The *AvrP123* product is recognized by the resistance genes P1, P2 and P3, and this recognition is likely to be through a direct interaction (Catanzariti et al., 2006).

Oomycete peptides

Similar to rust and mildew fungi, oomycetes form haustoria that are essential for the extraction of nutrients from plant tissue. Oomycetes are believed to secrete hundreds of effector proteins, which are functional in either in the extracellular space or in the host cell cytoplasm (Morgan & Kamoun, 2007). A RXLR-motif, flanked by a high frequency of acidic residues (D/E), in the N-terminal region of oomycete effector molecules is essential for delivery into the host cell cytoplasm (Whisson *et al.*, 2007). The RXLR motif has not been detected in fungi and may be specific for oomycetes (Ellis *et al.*, 2006). However, the motif is similar in sequence to the host cell targeting sequence found in proteins from malaria parasites (*Plasmodium* species)(Morgan & Kamoun, 2007). Analysis of the *Phytophthora infestans* genome sequence further revealed 425 putative genes coding for secreted proteins containing a RXLR motif. A comparison of the genomes of *Phytophthora sojae* and *Phythophthora ramorum* revealed as expected an overall high level of similarity (Tyler *et al.*, 2006). However, RxLR proteins (e.g. *Avr1b-1*, see below) have a strong tendency to reside in species-specific clusters (indel-blocks) (Jiang *et al.*, 2006a). Below we will discuss five of the best described RXLR effectors from *Hyaloperonospora parasitica* and *Phytophthora spp.*.

Atr13

ATR13 (also known as *Ppat17*) is a highly polymorphic gene from *Hyaloperonospora parasitica* that is up-regulated during infection of Arabidopsis (Bittner-Eddy *et al.*, 2003 1915). This expression pattern suggests that the ATR13 peptide is involved in promoting pathogen virulence, however its mode of action remains to be elucidated. The *ATR13* gene in *H. parasitica* is sufficient to trigger *RPP13*-dependent resistance in *Arabidopsis thaliana*. Moreover, bacteria and viruses modified such that they express the *ATR13* gene are also controlled by the RPP13-dependent resistance in *Arabidopsis thaliana* (Rentel *et al.*, 2008).

Depending on the isolate of *H. parasitica*, the ATR13 precursor protein consists either of 154 or 187 amino-acids, of which the N-terminal 19 residues are a signal peptide for secretion (Allen *et al.*, 2004). The precursor of 187 amino acids is processed into an ATR13 protein with four imperfect repeats of 11 amino acids, which are missing in the variant deriving from the 154 amino acid precursor. Sequence analysis of five ATR13 alleles showed the presence of two synonymous, two indels, and 26 nonsynonymous polymorphisms (Allen et al., 2004). The nonsynonymous polymorhisms were mainly located in the region corresponding to the C-terminal part of the protein, in the mature peptide. Interestingly, also the matching *RPP13* resistance gene shows an extreme variability in sequence, although a direct interaction between ATR13 and RPP13 could not be confirmed (Allen et al., 2004).

Avr1b-1 and *Avr3a*

Avr1b-1 is a positively charged (Table 1), species-specific effector peptide from *Phythophthora sojae* (Jiang et al., 2006a), which is strongly expressed 1-2 days after infection of soybean, indicating a role in promoting virulence (Shan *et al.*, 2004). Avr1b-1 shows no significant sequence simililarity to a functionally annotated gene, and its role in virulence remains to be elucidated. The avr1b-1 peptide is required for avirulence on soybean plants carrying the *Rps1b* resistance gene (Shan et al., 2004). Interestingly, *AVR1b-1* alleles of *P. sojae* isolates virulent on *Rps1b* plants showed numerous nonsynonymous mutations, which is indicative of strong diversifying selection. In some other virulent isolates, *AVR1b-1* mRNA did not accumulate at all (Shan et al., 2004), which is another strategy to avoid recognition by the host.

The avr3a peptide from *P. infestans* shares a low level of sequence similarity (29%) with avr1b-1 from *P. sojae*. Armstrong and coworkers ((Armstrong *et al.*, 2005) analyzed the *Avr3a* sequence from 55 *P. infestans* strains, and found only three single nucleotide polymorphisms. These polymorphisms resulted in amino-acid changes at position 19 (S/C), 80 (E/K) and 103 (M/I), of which position 80 and 103 and located in

the mature peptide. Interestingly, the presence of AVR3a^{SEM} correlated with virulence, while the presence of Avr3a^{CKI} was associated with avirulence on plants containing the *R3a* resistance from *Solanum demissum*. Activation of *R3a* by Avr3a^{KI} was shown to be dependent on ubiquitin-ligase-associated protein SGT1 and on heat-shock protein HSP90 (Bos *et al.*, 2006). Avr3a^{KI} is able to suppress cell death induced by the elicitin INF1 of *P. infestans*, which suggests a virulence role in host defense suppression for Avr3a^{KI} (Bos et al., 2006). The C-terminal 75 amino-acids of Avr3a, which exclude the RxLR region, were shown to be sufficient for both avirulence and suppression activities (Bos et al., 2006).

INF1

Elicitin (ELI) and elicitin-like (ELL) peptides are secreted by the oomycetes *Phytophthora* and *Pythium* spp. (Kamoun *et al.*, 1994, Huet *et al.*, 1995). The *inf* gene family codes for highly conserved 10 kDa elicitin peptides (Table 1). Elicitins induce defense responses on tomato cultivars carrying the *Cf9* resistance gene, and on all cultivars of tobacco. As such, they limit the host range of *Phytophthora* and *Pythium* spp. (Kamoun *et al.*, 1999, Kamoun *et al.*, 1998b).

All members of the elicitin family encode peptides or proteins that share the highly conserved 98 amino-acids elicitin domain corresponding to the mature, negatively charged INF1 peptide (Table 1). The function of elicitins in promoting virulence of the pathogen is related to sterol-binding and/or processing. This is essential for *Phytophthora*, which cannot synthesize sterols itself and must retrieve them from external sources (Mikes *et al.*, 1998). Sterols are hydrophobic molecules, and the hydrophobicity of INF1 (GRAVY score +0.26 kcal/mol) reflects its sterol-binding properties.

Elicitin-like proteins possess either shorter or longer elicitin domains that are more diverse at the sequence level (Jiang *et al.*, 2006b). Phylogenetic analysis of 156 elicitin domains derived from elicitins and elicitin-like sequences from different

Phytophthora and *Pythium* species resulted in 17 clades, which have all undergone purifying selection pressure (Jiang et al., 2006b). Some *inf* genes from *P. infestans* are clustered in the genome. These clusters are dispersed over a large region that harbors repeats and transposable elements (Jiang *et al.*, 2005), which may play a role in recombination events that drive diversity within the *inf* gene family.

Kamoun and coworkers (Kamoun *et al.*, 1998a) analyzed the presence of INF1 in culture filtrates of 102 isolates of *P. infestans*. Five isolates did not secrete the INF1 peptide, although the corresponding gene was present. These isolates were derived from a restricted geographical area, and all belonged to the same clonal lineage that dominated the world population until the 1980s. Additionally, silencing of the *inf 1* gene in *P. infestans* resulted in strains that were able to colonize the non-host *N. bethamiana* (Kamoun et al., 1998b). Furthermore, lack of elicitin production correlated with increased virulence of *P. parasitica* on susceptible plants (Kamoun et al., 1994).

Scr74 and Pcf

SCR74 is a secreted cysteine rich protein consisting of 74 amino-acids (Liu *et al.*, 2005) found in *P. infestans*. SCR74 shows sequence similarity to PcF, a necrosis inducing peptide (52 amino-acids) from *P. cactorum* (Liu et al., 2005). *Scr74* is up-regulated during infection of tomato and potato (Liu et al., 2005), which suggests that SCR74 could have a role as effector. Sequence analysis of the *scf74* locus in different *P. infestans* isolates revealed highly polymorphic alleles (Liu et al., 2005). Liu and coworkers sequenced a 304 bp fragment from 45 *P. infestans* samples, representing six different isolates. Twenty-one polymorphic amino-acid sites were found, nineteen of which were located in the mature peptide-encoding region. The polymorphisms resulted from diversifying selection, which was mostly affecting the mature peptide encoding region (Liu et al., 2005). On top of this, three *scr74* gene copies were found in a 300 kb region of the *P. infestans* genome indicating that gene duplication and selection may also have contributed to the diversity in *scr74* genes (Liu et al., 2005).

Nematode peptides

Plant parasitic nematodes are obligate parasites, obtaining nutrition only from living plant cells. Sedentary endoparasitic nematodes (root knot and cyst nematodes) invade the plant root and induce highly sophisticated feeding structures, often near vascular tissue, from which they feed for several weeks (Williamson & Hussey, 1996). During parasitism, the nematode releases effectors into the extracellular matrix and injects effectors into the cytoplasm of host cells through an oral stylet. The nematode effectors are produced in three specialized secretory cells, which are connected to the lumen of the stylet. Nematode effectors are believed to play a role in host invasion, feeding site formation and feeding, and in suppression of host defense responses. Below we discuss four classes of peptides whose functions are implicated in parasitism.

CLE-like peptides

SYV-46 is a peptide of 139 amino acids specifically produced in parasitic stages of the soybean cyst nematode *Heterodera glycines*. Although the peptide is unique to *H. glycines*, it does include a short C-terminal motif with similarity to the CLAVATA3/ESR-related (CLE) protein family in *Arabidopsis* (Olsen & Skriver, 2003, Wang *et al.*, 2001). CLV3 is the founding member of the CLE protein family, and it is a negative regulator of *WUSCHEL* (*WUS*). *WUS* is a transcription factor, restricting the size of the stem cell population in the shoot and floral meristems (Cock & McCormick, 2001, Schoof *et al.*, 2000). Expression of *Hg-SYV46* rescued the mutant phenotype of the clv3-1 mutant in *A. thaliana*, suggesting a similar functionality for both peptides (Wang *et al.*, 2005).

At present, 179 non-redundant CLE-like sequences have been found in various plant species (Strabala *et al.*, 2006). The proteins encoded by these sequences have a conserved stretch of 14 amino-acids, designated as the CLE domain, close to or at the C-terminus in common. The mature peptide of CLV3 consists of 12 amino-acids (Kondo

et al., 2006), and processing of the CLV3 proprotein into this mature peptide is crucial for CLV3 functionality. Besides a role in meristem maintenance, subsets of the CLE peptides are believed to have a role in vascular development and root growth (recently reviewed by (Jun *et al.*, 2007)and (Mitchum *et al.*)). The *Arabidopsis* family of *CLE*-like genes consist of 32 members, most of which are scattered across the genome (Jun et al., 2007). The diversity of the CLE-family is believed to result from whole genome duplications, reshuffling and local duplications. The latter is indicated by the presence of a tight cluster of CLE4, CLE5 and CLE6 on chromosome 2 (Jun et al., 2007).

Syv-46 is unique to *H. glycines,* but in the root knot nematode *Meloidogyne incognita* the 16D10 peptide was discovered, which is also a CLE-like peptide. The 16D10 gene is conserved between root knot nematode species. The 16D10 gene is specifically expressed in the subventral esophageal gland cells of preparasitic and parasitic juveniles of *Meloidogyne incognita* (Huang *et al.*, 2006b) The gene codes for a precursor protein of 43 amino-acids (Table 1), of which the first 30 amino-acids are a signal peptide for secretion. The 13 amino-acids of the mature peptide are similar to the conserved C-terminal motif of the plant CLE protein family. However, expression of 16D10 could not restore the wild-type phenotype of the *clv3* mutant in *Arabidopsis* (Huang et al., 2006b).

Overexpression of 16D10 in *Arabidopsis* resulted in an increase in lateral root formation, a phenotype that could be complemented by crossing the 16D10 overexpressing plants with plants producing double stranded RNA, resulting in silencing of 16D10 (Huang et al., 2006b, Huang *et al.*, 2006a). The 16D10 peptide is very hydrophilic with a GRAVY score of -1.96 kcal/mol, and it interacts with SCARECROW like transcription factors (Huang et al., 2006b). So far, root-knot and cyst nematode CLEs are the only CLE-like peptides outside the plant kingdom, suggesting that nematodes may have acquired the CLE genes by lateral transfer from plants (Oelkers *et al.*, 2008).

Hs-Ubi1

The ubiquitin extension protein is a peptide of about 122 amino acids, consisting of a N-terminal signal peptide for secretion, a central highly conserved mono-ubiquitin domain, and a variable positively charged C-terminal domain (Tytgat *et al.*, 2004). Fusion of the ubiquitin extension protein of *Heterodera schachtii*, Hs-ubi1, to GFP shows that the C-terminal domain is targeted to the nucleolus, but the exact function of the peptide in parasitism remains to be elucidated.

NEMPEP

Nucleolar accumulation was also found for NEMPEP peptide from the potato cyst nematode *Globodera rostochiensis* (chapter 3, this thesis). *NEMPEP* shows no significant sequence similarity to other sequences in the database, and has only been found in the potato cyst nematode *Globodera rostochiensis* (chapter 3, this thesis).

The cyst nematode life cycle consists of four juvenile stages, followed by an adult stage (Lee, 2002, Zuncke & Eisenback, 1998). The parasitic second stage juvenile (J2) migrates through the host root, and transforms a suitable cell near the vasculature into a sophisticated feeding cell. *NEMPEP* is specifically expressed in the dorsal esophageal gland of preparasitic juveniles. The timing of the expression of *NEMPEP* in preparasitic juvenile nematodes, together with the presence of a cleavable signal peptide for secretion, makes NEMPEP an effector protein likely to be secreted during parasitism. Besides a high level of *NEMPEP* expression in preparasitic juveniles, *NEMPEP* is expressed in the parasitic juvenile stage after induction of a feeding cell (J3). However, expression of *NEMPEP* is strikingly absent from invading juveniles (chapter 3, this thesis). Host defence responses against nematodes occur early in the interaction, mainly during the invasion phase (Cabrera Poch *et al.*, 2006). The absence of *NEMPEP* expression in migratory juveniles may be an effective way to avoid recognition by the host.

NEMPEP is a positively charged peptide, with a C-x-G γ-core motif (Table 1).

Currently we are investigating if NEMPEP has antimicrobial activity. On the other hand, the hydrophilic nature of NEMPEP (Table 1) may point to a role as a peptide hormone. The feeding cell induced by the potato cyst nematode, the so-called syncytium, results from the (partial) fusion of neighbouring cells. Syncytium formation is accompanied by cell-cycle reactivation, and the nuclei of a syncytium are enlarged due to repeated rounds of DNA replication (endoreduplication)(Gheysen *et al.*, 1997).

Transformation of potato and *Arabidopsis* with a *35S::NEMPEP* overexpression construct resulted in a lowered transformation efficiency, as compared to an empty vector control (Chapter 3, this thesis). The two potato transformants obtained had a smaller root system, a shorter shoot and shortened internodes between their leaves. This phenotype was also observed for potato plants overexpressing the cytokinin biosynthetic isopentenyltransferase (*IPT*) gene from *Agrobacterium tumefaciens* (Ivana *et al.*, 1997). Currently, we are investigating if overexpression of NEMPEP results in interference with cytokinin signalling.

SECPEPs

The *SECPEP* gene family is a highly divergent family of genes code for nematode effectors from the potato cyst nematode *Globodera rostochiensis* (van Bers *et al.*, chapter 2, this thesis, and Table 1). The family consists of at least nine members, of which two members are believed to be pseudogenes. The *SECPEPs* are specifically expressed in the dorsal gland of preparasitic juveniles of the potato cyst nematode *Globodera rostochiensis*, and antibodies recognizing SECPEP1 labelled granular structures, which probably correspond to secretory vesicles, in the lobe and the extension of this gland (van Bers *et al.*, chapter 2, this thesis). The mature peptide encoding regions of the *SECPEP*s are highly divergent in sequence, which is in contrast to the signal peptide region and the non-coding intron- and 3'UTR regions of the genes. Diversifying selection acts on several sites in the first exon of the *SECPEPs*. The first exon of the *SECPEPs* codes for the signal peptide for secretion and for the first 8 to 10

residues of the mature peptide. Interestingly, all the sites in the first exon that are undergoing diversifying selection correspond to residues of the mature peptide. The second exon of the *SECPEPs* codes for the C-terminal part of the mature peptide, however the sequence of this exon is too diverse to be analysed by the maximum likelihood method (PAML)(van Bers *et al.*, chapter 2, this thesis).

In addition to diversification on genomic level, additional sequence variation is introduced into the *SECPEPs* during or after transcription (van Bers *et al.,* chapter 2 and 5, this thesis). For one of the members, *SECPEP5*, we have found alternatively spliced transcripts in which one of the exons is completely lacking. SECPEP5 has (at least) 8 exons, and exon 3, 4 and 5 are highly similar in sequence, which suggests relatively recent exon duplication. For *SECPEP1*, we observed significantly higher rate of non-synonymous nucleotide changes, due to point mutations and insertions/deletions of single nucleotides, in *SECPEP1* cDNA than observed in *SECPEP1* gDNA or in cDNA of transcripts of a GAPDH gene (*Gr-gpd*), which conserved across many phyla (van Bers *et al.*, Chapter 5, this thesis). Because the changes are absent in the genomic *SECPEP1* sequence, we hypothesize that they are introduced into the transcripts by RNA-editing (Gott & Emeson, 2000, Nishikura, 2006{Tonkin, 2002 #1926)}. The surprising diversity of the *SECPEPs* may in part be attributed to exon shuffling. This was suggested because all the introns interrupting the coding region of the *SECPEP* genes are in phase 1 (van Bers *et al.*, chapter 2, this thesis). Exons flanked by two phase 1 introns have been shown to have a higher chance to have resulted from an exon shuffling event (Kaessmann *et al.*, 2002). Preliminary data resulting from the sequencing of a BAC-library clone indicates that at least two copies of *SECPEP1* and *SECPEP2*, *SECPEP5*, *SECPEP6*, *SECPEP9* and a putative novel homolog are located within a stretch of ~97 kb (van Bers *et al.*, unpublished data).

The *SECPEP*s show neither significant sequence similarity to anything with a known function in the sequence database, nor to a sequence outside *Globodera rostochiensis*. All SECPEP peptides are positively charged, hydrophilic peptides, and

the C-x-G γ-core motif is encoded in all the *SECPEP* genes (Table 1). These characteristics group the SECPEPs together with host defense polypeptides (Yeaman & Yount, 2007). Currently, we are investigating if the SECPEPs show antimicrobial activity, which could protect the host plant against secondary infections by opportunistic microbes. Such a protective role would be beneficial for the potato cyst nematode, because it is for its feeding and reproduction over weeks time dependent on a viable host. As a competing hypothesis, the high hydrophilicity of the SECPEPs may point to a role as peptide hormones (Boman, 2003). Preliminary data show that expression of *SECPEP3* under control of the constitutive 35S promoter results in altered root growth and leaf morphology of potato and *Arabidopsis* plants, for which processing of the peptide may be important (van Bers *et* al., chapter 4, this thesis). The SECPEP3 mature peptide includes a four residue KEKP(L) motif. This motif, of unknown function, is also found in the cyclin-dependent kinase ICK1 of *Arabidopsis*, and several copies of a similar motif (KEK) was found in the systemin precursor prosystemin (Constabel et al., 1998), where it was suggested to be involved in protein-protein interactions. SECPEP6 and SECPEP7 share an amino-acid motif ((K)PRPP(P)) with the mammalian cyclin-dependent kinase inhibitor $p27^{kip}$ and with two *Arabidopsis* peptides that are believed to function as cyclin-dependent kinase inhibitors, SIM1 and Smr1 (Weinl *et al.*, 2005, Churchman *et al.*, 2006). The (K)PRPP(P) motif is not found in SECPEP3, but SECPEP3 shares an additional three residue motif (KIP) with SIM and Smr1 (Chapter 4, this thesis). Interestingly, both the KEKP(L) and the (K)PRPP(P) motif are also present in the C-terminal 23 amino-acids of respectively *PROPEP2* and *PROPEP6* of *Arabidopsis*. This C-terminal region is believed to code for the mature peptide molecules involved in host defense ((Huffaker et al., 2006) see above). These motifs may be involved in protein-protein interactions, and in their presence sustains the hypothesis that SECPEP3 functions as a peptide hormone.

As a peptide hormone, SECPEP3 may alter cell cycle progression in nematode feeding sites. The nuclei of cells constituting the feeding cell induced by the potato cyst

nematode have undergone repeated endoreduplication cycles (reviewed by (Gheysen et al., 1997). SECPEP3 shares sequence motifs with cyclin-dependent kinase inhibitors, which play a prominent role in controlling endoreduplication (Weinl et al., 2005, Schnittger *et al.*, 2003, Verkest *et al.*, 2005, Churchman et al., 2006). Previously we showed that the fraction of small proteins (<3 kDa) co-stimulates cell proliferation of tobacco protoplasts (Goverse *et al.*, 1999). Currently, we are investigating if SECPEP3, and/or other members of the *SECPEP* family alter cell cycle progression.

Concluding remarks

In this review, we discussed a wide range of peptides involved in plant-pathogen interactions. Remarkably, while there are many examples of peptide effectors from eukaryotic pathogens of plants, bacterial effectors seem to be generally larger in size (>250 amino-acids). For several phytopathogenic bacteria, the complete inventory of type III effectors is known (Chang *et al.*, 2005, Alfano & Collmer, 2004, Genin & Boucher, 2004), which makes it unlikely that peptide effectors from bacteria will be uncovered by future research. Possibly, the type III secretion system in bacteria is not compatible with the translocation of peptides into the host cell cytoplasm.

A common denominator of the majority of the peptides (discussed here) is their net positive charge (Table 1), which seems to be driven by three mechanisms. First, the mode of action of many peptides involves the interaction with negatively charged membranes. In case of antimicrobial peptides this type of interactions may lead to pore formation and detrimental ion leakage of targeted cells. For example, the antifungal peptide α-hordothionin with an extremely high net charge is believed to interact with membranes of invading fungi. Pathogens may adapt to avoid recognition by antimicrobial peptides, however, their freedom to change is functionally constraint by the membranes.

Secondly, the net positive charge of the peptide effectors may be required for their binding to specific receptor proteins. The hydrophilicity of amino-acids side chains

in a peptide determines its ability to be involved in intermolecular protein-protein interactions (Boman, 2003). The positively charged residues arginine and lysine have the most hydrophilic side chains, and tend to position themselves in a protein such that they are solvent exposed. In addition to a higher protein-protein interaction potential, the prevalence of positively charged residues may be coherent to the small size of the peptides, which may be essential for receptor binding. Simplistically phrased, there is little room in the three dimensional structure of small peptides to bury hydrophobic and more negatively charged residues. For instance, Hg-Cle (charge +6, 116 amino-acids, 18% positively charged residues) is a functional homolog of CLV3, which activates CLV1/2 receptor signalling in *A. thaliana* (Rojo *et al.*, 2002). Similarly, the 23 residues of AtPep (charge +6, 37% positively charged residues) interact with the PEPR1 receptor protein.

A third possible mechanism explaining the bias towards positively charged peptides may depend on the accelerated diversification of the corresponding genes. Of the four classes of negatively charged peptides (Table 1), two are under purifying selection pressure (i.e. INF1 and the ECPs). Also PWL2 has a net negative charge, and its signal peptide seems less conserved than the mature peptide (see above). Avr3a of *P. infestans* has no positive net charge (0), and the alleles of Avr3a show a low level of sequence diversity. Avr3a and Avr1b-1 (charge +4) show some sequence similarity (29%), however, alleles of *avr*1b-1 are highly divergent. The positively charged residue arginine is encoded by six different codons, five of which contain CG or GG dinucleotides. Methylation of CG dinucleotides plays an important role in protection against deamination by activation-induced cytidine deaminase (AID)(Larijani *et al.*, 2005), which is an essential enzyme in hypermutation. A protection against further hypermutation could result in accumulation of CG dinucleotides, which will often result in positive charged residues, in hypermutated genes.

The majority of effectors peptides described here give evidence of diversifying selection. The peptides mostly accumulate changes in the mature peptide, whereas

diversifying selection seems absent in the signal peptide sequence and in the non-coding regions. Depending on the role of the peptide, these changes may evolve to adapt to an evasive target (i.e. antimicrobial peptide), or they may occur to avoid recognition by the immune system of a host (i.e. loss of avirulence activity). In the latter case, effectors which are directly interacting with immune receptors in plants are more likely to be hypervariable than effectors that are noted by the changes they induce in other host proteins (Bent & Mackey, 2007, Dodds et al., 2006).

Yeaman and Yount (Yeaman & Yount, 2007) suggested a role for the C-x-G γ-core motif in membrane and receptor interactions. We showed that the C-x-G γ-core motif is also present in a large subset of the peptides discussed in this paper. It is tempting to speculate on a direct interaction of the peptides with receptor-like products encoded by resistance genes. The C-x-G γ-core motif is absent from all oomycete effectors and their precursor proteins discussed here. The motif is also absent from systemin and from the CLE peptides secreted by nematodes, which are examples of peptides likely to interact with receptors. The C-x-G γ-core motif may be the signature of a shared evolutionary origin of several classes of peptides (Yeaman & Yount, 2007), and this origin may not be shared with the oomycete effector peptides and with systemin and the nematode CLE peptides.

The C-x-G γ-core motif may provide a basis for a novel search strategy to identify novel peptide effectors in fungi and nematodes. To date, peptide effectors have been overlooked in nematode sequence projects. The SECPEPs and NEMPEP are not processed from larger proteins, and have transcripts smaller than the usual cut-off limits for cDNA construction. The absence of genes coding for other peptide effectors in sequence databses is therefore likely to be an artifact (van Bers *et al.,* chapter 2, this thesis). *NEMPEP* and the *SECPEPs* were originally identified by cDNA-AFLP, a technique for which no discrimination against small size is expected. Given the fact that the *SECPEPs* are hypervariable, they have no match in the sequence database. For instance, the mature sequences of two neighboring loci coding for SECPEPs do not

significantly match in a BLAST algorithm (unpublished data). Only the intron sequences in these homologs point to a common ancestry of the loci. This stresses the potential of the C-x-G γ-core motif as the sole indicator for hypervariable effector peptides in nematodes.

The biological role in promoting virulence of most of the effector peptides described in this review is still unknown. This is largely due to the accelerated diversification of genes coding for these effectors, which restricts the similarity to sequences with a known function in the database. On top of this hypervariability in genomic sequence, additional variation may be introduced into the transcripts of the genes by alternative splicing and RNA-editing, as was observed for the *SECPEPs*. Despite this hypervariability in primary sequence, the majority of the effector peptides share three characteristics: a positive net charge positive, a C-x-G γ-core motif and a relative hydrophilicity. These characteristics provide a novel strategy in identifying effector peptides, and may aid in future studies on elucidating the role of pioneer effector molecules in promoting pathogen virulence. The similarities of the SECPEPs and NEMPEP to host defense peptides, in combination with the phenotypes we observed upon expression in plants, led us to propose on two models, which are not mutually exclusive, regarding the functional role of these peptides in plant-nematode interaction. The first one postulates a role as antimicrobial peptide in protecting the host plant against opportunistic microbes. The second model, which is based on the plant phenotypes, on shared sequence motifs and on the hydrophilicity of the peptides, postulates a role as peptide hormones in interference with hormonal regulation and cell cycle signaling.

Table 1: Characteristics of peptides involved in plant-microbe interactions

	Peptide	Organism	Size[a] (aa)	Charge	Function	C-x-G	GRAVY	Reference[b]	Acc. number
Plant peptides	**Systemin**	*Lycopersicon esculentum*	18 (200)	+2	Activates defense-related genes	-	-1.72	(Ryan & Pearce, 2003)	gil228051
	***At*PEP1 (PROPEP1)**	*Arabidopsis thaliana*	23 (92)	+6		+[g]	-1.69	(Huffaker et al., 2006)	NP_569001
	PROPEP2		36? (109)	+4		+[g]	-1.77		NM_125887
	PROPEP3		23? (96)	+6	Unknown	+[g]	-1.87		NM_125889
	PROPEP4		27? (81)	+7		-	-1.39		NM_121035
	PROPEP5		27? (86)	+4		-	-0.92		NM_121036
	PROPEP6		23? (104)	+5		+[g]	-1.35		NM_127769
	PROPEP7		23? (75)	+4		?	-1.1		Unannotated
	α-Hordothionin	*Hordeum vulgare*	45 (127)	+10	Antifungal peptides	+	-0.32	(Ponz et al., 1986)	CAA29330
	DM-AMP1 (defensin)	*Dahlia merckii*	50 (y[e])	+1		+[c]	-0.69	(Osborn *et al.*, 1995)	AAB34972
	IWF1 (nsLTP)	*Beta vulgaris*	91 (117)	+7		+	+0.24	(Nielsen et al., 1996)	CAA63407
	Ac-AMP1 (Hevein-like)	*Amaranthus caudatus*	29	+3		+[c]	-0.20	(Broekaert et al., 1992)	AAB22103
Bacterial peptides	**PNP-like protein**	*Xanthomonas axonopodis*	143	+8	Osmoregulation	+[C]	+0.12	(da Silva Ac Fau - Ferro *et al.*)	NP_642965

Fungal Peptides	**Avr2**	*Cladosporium fulvum*	59 (78)	+7	Inhibits Rcr3 protease	+	-0.48	(Luderer et al., 2002b)	CAD16675
	Avr4		86[f] (135)	+3	Protects against plant chitinases	+	-0.78	(Joosten et al., 1997)	CAA69643
	Avr4E		101 (121)	-3	Unkn own	-	-0.21	(Westerink et al., 2004)	AAT28197
	Avr9		28[f] (63)	+1	Unknown, similarity to peptidase inhibitor	+	-0.25	(Van Den Ackerveken et al., 1993b)	CAA01455
	ECP1		73 (96)	-1	Suppressing host defense?	+[c]	-0.59	(Van den Ackerveken et al., 1993a)	CAA78400
	ECP2		142 (165)	-6	Suppressing host defense?	+[c]	-0.55		CAA78401
	Six1	*Fusarium oxysporum f. sp. lycopersici*	~120 (248)	~+4[d]	Unknown; Matches the I-3 reststance gene	+[c]	~ -0.94[d]	(Rep et al., 2004)	AJ608702
	PWL2	*Magnaporthe grisea*	124 (145)	-6	Unknown; determines host range	-	-1.51	(Kang et al., 1995)	AAA91019
	NIP1	*Rhynchosporium secalis*	60 (82)	0	Stimulates H^+-ATPase; matches *Rrs1* gene	+	-0.28	(Van't Slot et al., 2007)	ABR92629
	AvrL567-A	*Melampsora lini*	127 (150)	+2	Unknown; binds to L567 resistance gene	-	-0.50	(Dodds et al., 2004)	AAS66948
Oomycete peptides	**Atr13**	*Hyaloperonospora parasitica*	135 (154)	0	Unknown; matches *RPP13* gene	-	-0.44	(Allen et al., 2004)	AAW63768
	Avr1b-1	*Phytophthora sojae*	117 (138)	+4	Unknown; matches *Rps1b* resistance gene	-	-0.95	(Shan et al., 2004)	AAR05402
	Avr3a	*Phytophthora infestans*	126 (147)	0	Unknown; Triggers *R3a*-dependent HR	-	-0.71	(Armstrong et al., 2005)	CAI72345
	Scr74		53 (74)	0	Unknown	-	-0.26	(Liu et al., 2005)	PC015G09

Oomycete peptides	**INF1**	*Phytophthora infestans*	98 (118)	-2	Sterol carrier; elicits defense responses	-	+0.26	(Jiang et al., 2005)	AAV92913
Nematode peptides	**Hg-Cle**	*Heterodera glycines*	116 (139)	+6	Peptide hormone	-	-0.70	(Mitchum et al.)	AAG21331
	16D10	*Meloidogyne incognita*	13 (43)	+2	Peptide hormone	-	-1.96	(Huang et al., 2006b)	AAZ77751
	Hs-Ubi1	*Heterodera schachtii*	98 (122)	+7	Unknown	+[g]	-0.92	(Tytgat et al., 2004)	AY286305
	N EMPEP	*Globodera rostochiensis*	40 (62)	+7	Unknown; Peptide hormone?	+	-1.3	Van Bers *et al.,in prep.*	CAD60975
	SECPEP1		54 (72)	+6	Unknown; Peptide hormone or antimicrobial peptide?	+	-1.10		CAC21847
	SECPEP2		44 (68)	+1	Unknown; Peptide hormone or antimicrobial peptide?	+[g]	-1.83		BM345411
	SECPEP3		31 (51)	+4		+	-1.35		-
	SECPEP5		108 (134)	+6		+[g]	-1.51		-
	SECPEP6		32 (51)	+5		+	-1.80		-
	SECPEP7		31 (49)	+2		+	-1.58		-
	SECPEP8		31 (49)	+3		+	-0.85		-

[a]Size of precursor protein in brackets

[b]Space limitations restrict us to one selected reference.Consult this reference for further reading.

[C] A CxG motif is found in reverse orientation: G-X-C

[d]In absence of conclusive evidence on the sequence of the SIX1 mature peptide, we used the central 100 amino-acids (100-200) of SIX1 for sequence analysis

[e] Peptide was identified by direct peptide sequencing, and size of the proprotein is as yet unpublished

[f]Size of the mature protein was determined by mass spectrometry

[g]CxG in precursor protein

References

Abramovitch, R. B. and Martin, G. B. (2004) Strategies used by bacterial pathogens to suppress plant defenses. *Current Opinion in Plant Biology,* **7,** 356-364.

Alfano, J. R. and Collmer, A. (2004) Type III secretion system effector proteins: Double agents in bacterial disease and plant defense. *Annual Review of Phytopathology,* **42,** 385-414.

Allen, R. L., Bittner-Eddy, P. D., Grenville-Briggs, L. J., Meitz, J. C., Rehmany, A. P., Rose, L. E.*, et al.* (2004) Host-parasite coevolutionary conflict between Arabidopsis and downy mildew. *Science,* **306,** 1957-1960.

Araz, A. and Maden, S. (2006) Pathogenic variation among isolates of Rhynchosporium secalis from cultivated barley growing in Central Anatolia, Turkey. *Plant Pathology Journal,* **5,** 244-247.

Armstrong, M. R., Whisson, S. C., Pritchard, L., Bos, J. I. B., Venter, E., Avrova, A. O.*, et al.* (2005) An ancestral oomycete locus contains late blight avirulence gene Avr3a, encoding a protein that is recognized in the host cytoplasm. *Proceedings of the National Academy of Sciences,* **102,** 7766-7771.

Bent, A. F. and Mackey, D. (2007) Elicitors, Effectors, and R Genes: The New Paradigm and a Lifetime Supply of Questions. *Annual Review of Phytopathology,* **45,** 399-436.

Bittner-Eddy, P. D., Allen, R. L., Rehmany, A. P., Birch, P. and Beynon, J. L. (2003) Use of suppression subtractive hybridization to identify downy mildew genes expressed during infection of Arabidopsis thaliana. *Molecular Plant Pathology,* **4,** 501-507.

Boller, T. (2005) Peptide signalling in plant development and self/non-self perception. *Current Opinion in Cell Biology,* **17,** 116-122.

Boman, H. G. (2003) Antibacterial peptides: Basic facts and emerging concepts. *Journal of Internal Medicine,* **254,** 197-215.

Bos, J. I. B., Kanneganti, T. D., Young, C., Cakir, C., Huitema, E., Win, J.*, et al.* (2006)

The C-terminal half of Phytophthora infestans RXLR effector AVR3a is sufficient to trigger R3a-mediated hypersensitivity and suppress INF1-induced cell death in Nicotiana benthamiana. *Plant Journal,* **48,** 165-176.

Broekaert, W. F., Marie?n, W., Terras, F. R. G., De Bolle, M. F. C., Proost, P., Van Damme, J.*, et al.* (1992) Antimicrobial peptides from Amaranthus caudatus seeds with sequence homology to the cysteine/glycine-rich domain of chitin-binding proteins. *Biochemistry,* **31,** 4308-4314.

Broekaert, W. F., Terras, F. R. G., Cammue, B. P. A. and Osborn, R. W. (1995) Plant Defensins: Novel Antimicrobial Peptides as Components of the Host Defense System. *Plant Physiol.,* **108,** 1353-1358.

Brown, K. L. and Hancock, R. E. W. (2006) Cationic host defense (antimicrobial) peptides. *Current Opinion in Immunology,* **18,** 24-30.

Cabrera Poch, H. L., Manzanilla Lo?pez, R. H. and Kanyuka, K. (2006) Functionality of resistance gene Hero, which controls plant root-infecting potato cyst nematodes, in leaves of tomato. *Plant, Cell and Environment,* **29,** 1372-1378.

Castagnaro, A., Marana, C., Carbonero, P. and Garcia-Olmedo, F. (1992) Extreme divergence of a novel wheat thionin generated by a mutational burst specifically affecting the mature protein domain of the precursor. *Journal of Molecular Biology,* **224,** 1003-1009.

Catanzariti, A. M., Dodds, P. N., Lawrence, G. J., Ayliffe, M. A. and Ellis, J. G. (2006) Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell,* **18,** 243-256.

Chang, J. H., Urbach, J. M., Law, T. F., Arnold, L. W., Hu, A., Gombar, S.*, et al.* (2005) A high-throughput, near-saturating screen for type III effector genes from Pseudomonas syringae. *Proceedings of the National Academy of Sciences of the United States of America,* **102,** 2549-2554.

Churchman, M. L., Brown, M. L., Kato, N., Kirik, V., Hu?lskamp, M., Inze, D.*, et al.* (2006) SIAMESE, a plant-specific cell cycle regulator, controls endoreplication

onset in Arabidopsis thaliana. *Plant Cell,* **18,** 3145-3157.

Cock, J. M. and McCormick, S. (2001) A large family of genes that share homology with CLAVATA3. *Plant Physiol,* **126,** 939-942.

Constabel, C. P., Yip, L. and Ryan, C. A. (1998) Prosystemin from potato, black nightshade, and bell pepper: Primary structure and biological activity of predicted systemin polypeptides. *Plant Molecular Biology,* **36,** 55-62.

da Silva Ac Fau - Ferro, J. A., Ferro Ja Fau - Reinach, F. C., Reinach Fc Fau - Farah, C. S., Farah Cs Fau - Furlan, L. R., Furlan Lr Fau - Quaggio, R. B., Quaggio Rb Fau - Monteiro-Vitorello, C. B.*, et al.* Comparison of the genomes of two Xanthomonas pathogens with differing host specificities.

Da Silva, F. G., Shen, Y., Dardick, C., Burdman, S., Yadav, R. C., De Leon, A. L.*, et al.* (2004) Bacterial genes involved in type I secretion and sulfation are required to elicit the rice Xa21-mediated innate immune response. *Molecular Plant-Microbe Interactions,* **17,** 593-601.

De Lucca, A. J., Cleveland, T. E. and Wedge, D. E. (2005) Plant-derived antifungal proteins and peptides. *Canadian Journal of Microbiology,* **51,** 1001-1014.

Dodds, P. N., Catanzariti, A. M., Lawrence, G. J. and Ellis, J. G. (2007) Avirulence proteins of rust fungi: Penetrating the host-haustorium barrier. *Australian Journal of Agricultural Research,* **58,** 512-517.

Dodds, P. N., Lawrence, G. J., Catanzariti, A. M., Ayliffe, M. A. and Ellis, J. G. (2004) The Melampsora lini AvrL567 avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *Plant Cell,* **16,** 755-768.

Dodds, P. N., Lawrence, G. J., Catanzariti, A. M., Teh, T., Wang, C. I. A., Ayliffe, M. A.*, et al.* (2006) Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences of the United States of America,* **103,** 8888-8893.

Ellis, J., Catanzariti, A. M. and Dodds, P. (2006) The problem of how fungal and

oomycete avirulence proteins enter plant cells. *Trends in Plant Science,* **11,** 61-63.

Escoubas, P. (2006) Molecular diversification in spider venoms: A web of combinatorial peptide libraries. *Mol Divers,* **10,** 545-554.

Farrokhi, N., Whitelegge, J. P. and Brusslan, J. A. (2008) Plant peptides and peptidomics. *Plant Biotechnology Journal,* **6,** 105-134.

Genin, S. and Boucher, C. (2004) Lessons learned from the genome analysis of Ralstonia solanacearum. *Annual Review of Phytopathology,* **42,** 107-134.

Gheysen, G., Engler, J. D. A. and Montagu, M. V. (1997) Cell cycle regulation in nematode feeding sites. In: *Cellular and molecular aspects of plant-nematode interactions.* (Fenoll, C., Grundler, F. M. W. and Ohl, S. A., ed.^eds.). Kluwer Academic Publishers, pp. 120-132.

Gott, J. M. and Emeson, R. B. (2000) Functions and mechanisms of RNA editing. *Annu Rev Genet,* **34,** 499-531.

Goverse, A., Rouppe van der Voort, J., Roppe van der Voort, C., Kavelaars, A., Smant, G., Schots, A.*, et al.* (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.

Hahn, M., Jungling, S. and Knogge, W. (1993) Cultivar-specific elicitation of barley defense reactions by the phytotoxic peptide NIP1 from Rhynchosporium secalis. *Molecular Plant-Microbe Interactions,* **6,** 745-754.

Huang, G., Allen, R., Davis, E. L., Baum, T. J. and Hussey, R. S. (2006a) Engineering broad root-knot resistance in transgenic plants by RNAi silencing of a conserved and essential root-knot nematode parasitism gene. *Proc Natl Acad Sci U S A,* **103,** 14302-14306.

Huang, G., Dong, R., Allen, R., Davis, E. L., Baum, T. J. and Hussey, R. S. (2006b) A root-knot nematode secretory peptide functions as a ligand for a plant

transcription factor. *Molecular Plant-Microbe Interactions,* **19,** 463-470.

Huet, J. C., Le Caer, J. P., Nespoulous, C. and Pernollet, J. C. (1995) The relationships between the toxicity and the primary and secondary structures of elicitinlike protein elicitors secreted by the phytopathogenic fungus Pythium vexans. *Molecular Plant-Microbe Interactions,* **8,** 302-310.

Huffaker, A., Pearce, G. and Ryan, C. A. (2006) An endogenous peptide signal in Arabidopsis activates components of the innate immune response. *Proc Natl Acad Sci U S A,* **103,** 10098-10103.

Huffaker, A. and Ryan, C. A. (2007) Endogenous peptide defense signals in Arabidopsis differentially amplify signaling for the innate immune response. *Proceedings of the National Academy of Sciences of the United States of America,* **104,** 10732-10736.

Ivana, M., Lidiya, S., Milos, O., Oksana, Z., Tatyana, K., Josef, E.*, et al.* (1997) Growth pattern, tuber formation and hormonal balance in in vitro potato plants carrying ipt gene. *Plant Growth Regulation,* **21,** 27-36.

Jiang, R. H. Y., Dawe, A. L., Weide, R., Van Staveren, M., Peters, S., Nuss, D. L.*, et al.* (2005) Elicitin genes in Phytophthora infestans are clustered and interspersed with various transposon-like elements. *Molecular Genetics and Genomics,* **273,** 20-32.

Jiang, R. H. Y., Tyler, B. M. and Govers, F. (2006a) Comparative analysis of Phytophthora genes encoding secreted proteins reveals conserved synteny and lineage-specific gene duplications and deletions. *Molecular Plant-Microbe Interactions,* **19,** 1311-1321.

Jiang, R. H. Y., Tyler, B. M., Whisson, S. C., Hardham, A. R. and Govers, F. (2006b) Ancient origin of elicitin gene clusters in Phytophthora genomes. *Molecular Biology and Evolution,* **23,** 338-351.

Joosten, M. H. A. J., Vogelsang, R., Cozijnsen, T. J., Verberne, M. C. and De Wit, P. J. G. M. (1997) The biotrophic fungus Cladosporium fulvum circumvents Cf-4-

mediated resistance by producing unstable AVR4 elicitors. *Plant Cell,* **9,** 367-379.

Jun, J. H., Fiume, E. and Fletcher, J. C. (2007) The CLE family of plant polypeptide signaling molecules. *Cellular and Molecular Life Sciences*.

Kaessmann, H., Zo?llner, S., Nekrutenko, A. and Li, W. H. (2002) Signatures of domain shuffling in the human genome. *Genome Research,* **12,** 1642-1650.

Kamoun, S., Hone?e, G., Weide, R., Lauge, R., Kooman-Gersmann, M., De Groot, K.*, et al.* (1999) The fungal gene Avr9 and the oomycete gene inf1 confer avirulence to potato virus X on tobacco. *Molecular Plant-Microbe Interactions,* **12,** 459-462.

Kamoun, S., Van Der Lee, T., Van Den Berg-Velthuis, G., De Groot, K. E. and Govers, F. (1998a) Loss of production of the elicitor protein INF1 in the clonal lineage US-1 of Phytophthora infestans. *Phytopathology,* **88,** 1315-1323.

Kamoun, S., Van West, P., Vleeshouwers, V. G. A. A., De Groot, K. E. and Govers, F. (1998b) Resistance of Nicotiana benthamiana to Phytophthora infestans is mediated by the recognition of the elicitor protein INF1. *Plant Cell,* **10,** 1413-1425.

Kamoun, S., Young, M., Forster, H., Coffey, M. D. and Tyler, B. M. (1994) Potential role of elicitins in the interaction between Phytophthora species and tobacco. *Applied and Environmental Microbiology,* **60,** 1593-1598.

Kandoth, P. K., Ranf, S., Pancholi, S. S., Jayanty, S., Walla, M. D., Miller, W.*, et al.* (2007) Tomato MAPKs LeMPK1, LeMPK2, and LeMPK3 function in the systemin-mediated defense response against herbivorous insects. *Proceedings of the National Academy of Sciences of the United States of America,* **104,** 12205-12210.

Kang, S., Sweigard, J. A. and Valent, B. (1995) The PWL host specificity gene family in the blast fungus Magnaporthe grisea. *Molecular Plant-Microbe Interactions,* **8,** 939-948.

Kondo, T., Sawa, S., Kinoshita, A., Mizuno, S., Kakimoto, T., Fukuda, H., *et al.* (2006) A plant peptide encoded by CLV3 identified by in situ MALDI-TOF MS analysis. *Science,* **313,** 845-848.

Kyte, J. and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology,* **157,** 105-132.

Larijani, M., Frieder, D., Sonbuchner, T. M., Bransteitter, R., Goodman, M. F., Bouhassira, E. E., *et al.* (2005) Methylation protects cytidines from AID-mediated deamination. *Molecular Immunology,* **42,** 599-604.

Laugé, R. and De Wit, P. J. G. M. (1998) Fungal Avirulence Genes: Structure and Possible Functions. *Fungal Genetics and Biology,* **24,** 285-297.

Laugé, R., Joosten, M. H. A. J., Haanstra, J. P. W., Goodwin, P. H., Lindhout, P. and De Wit, P. J. G. M. (1998) Successful search for a resistance gene in tomato targeted against a virulence factor of a fungal pathogen. *Proceedings of the National Academy of Sciences of the United States of America,* **95,** 9014-9018.

Laugé, R., Joosten, M. H. A. J., Van Den Ackerveken, G. F. J. M., Van Den Broek, H. W. J. and De Wit, P. J. G. M. (1997) The in planta-produced extracellular proteins ECP1 and ECP2 of Cladosporium fulvum are virulence factors. *Molecular Plant-Microbe Interactions,* **10,** 725-734.

Lee, D. L. (2002) Life cycles. In: *The Biology of Nematodes.* (Lee, D. L., ed.^eds.). London: Taylor and Francis, pp. 61-72.

Lee, S. W., Han, S. W., Bartley, L. E. and Ronald, P. C. (2006) Unique characteristics of Xanthomonas oryzae pv. oryzae AvrXa21 and implications for plant innate immunity. *Proceedings of the National Academy of Sciences of the United States of America,* **103,** 18395-18400.

Liu, Z., Bos, J. I. B., Armstrong, M., Whisson, S. C., Da Cunha, L., Torto-Alalibo, T., *et al.* (2005) Patterns of diversifying selection in the phytotoxin-like scr74 gene family of Phytophthora infestans. *Molecular Biology and Evolution,* **22,** 659-

672.

Luderer, R., De Kock, M. J. D., Dees, R. H. L., De Wit, P. J. G. M. and Joosten, M. H. A. J. (2002a) Functional analysis of cysteine residues of ECP elicitor proteins of the fungal tomato pathogen Cladosporium fulvum. *Molecular Plant Pathology,* **3,** 91-95.

Luderer, R. and Joosten, M. H. A. J. (2001) Avirulence proteins of plant pathogens: Determinants of victory and defeat. *Molecular Plant Pathology,* **2,** 355-364.

Luderer, R., Takken, F. L. W., De Wit, P. J. G. M. and Joosten, M. H. A. J. (2002b) Cladosporium fulvum overcomes Cf-2-mediated resistance by producing truncated AVR2 elicitor proteins. *Molecular Microbiology,* **45,** 875-884.

Matsubayashi, Y. and Sakagami, Y. (2006) Peptide Hormones in Plants. *Annu Rev Plant Biol*.

McGurl, B., Pearce, G., Orozco-Cardenas, M. and Ryan, C. A. (1992) Structure, expression, and antisense inhibition of the systemin precursor gene. *Science,* **255,** 1570-1573.

McGurl, B. and Ryan, C. A. (1992) The organization of the prosystemin gene. *Plant Molecular Biology,* **20,** 405-409.

Mikes, V., Milat, M. L., Ponchet, M., Panabie?res, F., Ricci, P. and Blein, J. P. (1998) Elicitins, proteinaceous elicitors of plant defense, are a new class of sterol carrier proteins. *Biochemical and Biophysical Research Communications,* **245,** 133-139.

Mitchum, M. G., Wang, X. and Davis, E. L. Diverse and conserved roles of CLE peptides. *Current Opinion in Plant Biology*.

Morgan, W. and Kamoun, S. (2007) RXLR effectors of plant pathogenic oomycetes. *Current Opinion in Microbiology,* **10,** 332-338.

Nembaware, V., Seoighe, C., Sayed, M. and Gehring, C. (2004) A plant natriuretic peptide-like gene in the bacterial pathogen Xanthomonas axonopodis may induce hyper-hydration in the plant host: a hypothesis of molecular mimicry.

BMC Evolutionary Biology, **4,** 10.

Nielsen, K. K., Nielsen, J. E., Madrid, S. M. and Mikkelsen, J. D. (1996) New antifungal proteins from sugar beet (Beta vulgaris L.) showing homology to non-specific lipid transfer proteins. *Plant Molecular Biology,* **31,** 539-552.

Nishikura, K. (2006) Editor meets silencer: Crosstalk between RNA editing and RNA interference. *Nature Reviews Molecular Cell Biology,* **7,** 919-931.

Oelkers, K., Goffard, N., Weiller, G., Gresshoff, P., Mathesius, U. and Frickey, T. (2008) Bioinformatic analysis of the CLE signaling peptide family. *BMC Plant Biology,* **8**.

Olivera, B. M. and Cruz, L. J. (2000) Conotoxins, in retrospect. *Toxicon,* **39,** 7-14.

Olsen, A. N. and Skriver, K. (2003) Ligand mimicry? Plant-parasitic nematode polypeptide with similarity to CLAVATA3. *Trends in Plant Science,* **8,** 55-57.

Osborn, R. W., De Samblanx, G. W., Thevissen, K., Goderis, I., Torrekens, S., Van Leuven, F.*, et al.* (1995) Isolation and characterisation of plant defensins from seeds of Asteraceae, Fabaceae, Hippocastanaceae and Saxifragaceae. *FEBS Letters,* **368,** 257-262.

Pearce, G., Moura, D. S., Stratmann, J. and Ryan, C. A. (2001) Production of multiple plant hormones from a single polyprotein precursor. *Nature,* **411,** 817-820.

Pearce, G., Strydom, D., Johnson, S. and Ryan, C. A. (1991) A polypeptide from tomato leaves induces wound-inducible proteinase inhibitor proteins. *Science,* **253,** 895-898.

Peschel, A. and Sahl, H. G. (2006) The co-evolution of host cationic antimicrobial peptides and microbial resistance. *Nat Rev Microbiol,* **4,** 529-536.

Ponz, F., Paz-Ares, J., Hernandez-Lucas, C., Garcia-Olmedo, F. and Carbonero, P. (1986) Cloning and nucleotide sequence of a cDNA encoding the precursor of the barley toxin α-hordothionin. *European Journal of Biochemistry,* **156,** 131-135.

Rentel, M. C., Leonelli, L., Dahlbeck, D., Zhao, B. and Staskawicz, B. J. (2008)

Recognition of the Hyaloperonospora parasitica effector ATR13 triggers resistance against oomycete, bacterial, and viral pathogens. *Proceedings of the National Academy of Sciences of the United States of America,* **105,** 1091-1096.

Rep, M., Dekker, H. L., Vossen, J. H., De Boer, A. D., Houterman, P. M., Speijer, D.*, et al.* (2002) Mass spectrometric identification of isoforms of PR proteins in xylem sap of fungus-infected tomato. *Plant Physiology,* **130,** 904-917.

Rep, M., Van Der Does, H. C., Meijer, M., Van Wijk, R., Houterman, P. M., Dekker, H. L.*, et al.* (2004) A small, cysteine-rich protein secreted by Fusarium oxysporum during colonization of xylem vessels is required for I-3-mediated resistance in tomato. *Molecular Microbiology,* **53,** 1373-1383.

Rodríguez-Palenzuela, P., Pintor-Toro, J. A., Carbonero, P. and Garci?a-Olmedo, F. (1988) Nucleotide sequence and endosperm-specific expression of the structural gene for the toxin alpha-hordothionin in barley (Hordeum vulgare L.). *Gene,* **70,** 271-281.

Rohe, M., Gierlich, A., Hermann, H., Hahn, M., Schmidt, B., Rosahl, S.*, et al.* (1995) The race-specific elicitor, NIP1, from the barley pathogen, Rhynchosporium secalis, determines avirulence on host plants of the Rrs1 resistance genotype. *EMBO Journal,* **14,** 4168-4177.

Rojo, E., Sharma, V. K., Kovaleva, V., Raikhel, N. V. and Fletcher, J. C. (2002) CLV3 is localized to the extracellular space, where it activates the Arabidopsis CLAVATA stem cell signaling pathway. *Plant Cell,* **14,** 969-977.

Rooney, H. C. E., Van't Klooster, J. W., Van Der Hoorn, R. A. L., Joosten, M. H. A. J., Jones, J. D. G. and De Wit, P. J. G. M. (2005) Plant science: Cladosporium Avr2 inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance. *Science,* **308,** 1783-1786.

Rozynek, P., Posch, A. and Baur, X. (1998) Cloning, expression and characterization of the major latex allergen prohevein. *Clinical and Experimental Allergy,* **28,**

1418-1426.

Ryan, C. A. and Pearce, G. (2003) Systemins: a functionally defined family of peptide signals that regulate defensive genes in Solanaceae species. *Proc Natl Acad Sci U S A,* **100 Suppl 2,** 14577-14580.

Schmidt, S. and Baldwin, I. T. (2006) Systemin in Solanum nigrum. The tomato-homologous polypeptide does not mediate direct defense responses. *Plant Physiology,* **142,** 1751-1758.

Schnittger, A., Weinl, C., Bouyer, D., Scho?binger, U. and Hu?lskamp, M. (2003) Misexpression of the cyclin-dependent kinase inhibitor ICK1/KRP1 in single-celled arabidopsis trichomes reduces endoreduplication and cell size and induces cell death. *Plant Cell,* **15,** 303-315.

Schoof, H., Lenhard, M., Haecker, A., Mayer, K. F. X., Ju?rgens, G. and Laux, T. (2000) The stem cell population of Arabidopsis shoot meristems is maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell,* **100,** 635-644.

Schürch, S., Linde, C. C., Knogge, W., Jackson, L. F. and McDonald, B. A. (2004) Molecular Population Genetic Analysis Differentiates Two Virulence Mechanisms of the Fungal Avirulence Gene NIP1. *Molecular Plant-Microbe Interactions,* **17,** 1114-1125.

Shan, W., Cao, M., Leung, D. and Tyler, B. M. (2004) The Avr1b Locus of Phytophthora sojae Encodes an Elicitor and a Regulator Required for Avirulence on Soybean Plants Carrying Resistance Gene Rps1b. *Molecular Plant-Microbe Interactions,* **17,** 394-403.

Silverstein, K. A. T., Graham, M. A., Paape, T. D. and Vandenbosch, K. A. (2005) Genome organization of more than 300 defensin-like genes in Arabidopsis. *Plant Physiology,* **138,** 600-610.

Silverstein, K. A. T., Moskal Jr, W. A., Wu, H. C., Underwood, B. A., Graham, M. A., Town, C. D.*, et al.* (2007) Small cysteine-rich peptides resembling antimicrobial

peptides have been under-predicted in plants. *Plant Journal,* **51,** 262-280.

Song, W. Y., Wang, G. L., Chen, L. L., Kim, H. S., Pi, L. Y., Holsten, T.*, et al.* (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. *Science,* **270,** 1804-1806.

Stec, B. (2006) Plant thionins - The structural perspective. *Cellular and Molecular Life Sciences,* **63,** 1370-1385.

Stergiopoulos, I., De Kock, M. J. D., Lindhout, P. and De Wit, P. J. G. M. (2007) Allelic variation in the effector genes of the tomato pathogen Cladosporium fulvum reveals different modes of adaptive evolution. *Molecular Plant-Microbe Interactions,* **20,** 1271-1283.

Strabala, T. J., O'Donnell P, J., Smit, A. M., Ampomah-Dwamena, C., Martin, E. J., Netzler, N.*, et al.* (2006) Gain-of-function phenotypes of many CLAVATA3/ESR genes, including four new family members, correlate with tandem variations in the conserved CLAVATA3/ESR domain. *Plant Physiol,* **140,** 1331-1344.

Stratmann, J., Scheer, J. and Ryan, C. A. (2000) Suramin inhibits initiation of defense signaling by systemin, chitosan, and a ?-glucan elicitor in suspension-cultured Lycopersicon peruvianum cells. *Proceedings of the National Academy of Sciences of the United States of America,* **97,** 8862-8867.

Sweigard, J. A., Carroll, A. M., Kang, S., Farrall, L., Chumley, F. G. and Valent, B. (1995) Identification, Cloning, and Characterization of PWL2, a Gene for Host Species Specificity in the Rice Blast Fungus. *Plant Cell,* **7,** 1221-1233.

Takken, F. L. W., Thomas, C. M., Joosten, M. H. A. J., Golstein, C., Westerink, N., Hille, J.*, et al.* (1999) A second gene at the tomato Cf-4 locus confers resistance to Cladosporium fulvum through recognition of a novel avirulence determinant. *Plant Journal,* **20,** 279-288.

Tennessen, J. A. (2005) Molecular evolution of animal antimicrobial peptides: Widespread moderate positive selection. *Journal of Evolutionary Biology,* **18,** 1387-1394.

Terlau, H. and Olivera, B. M. (2004) Conus Venoms: A Rich Source of Novel Ion Channel-Targeted Peptides. *Physiological Reviews,* **84,** 41-68.

Thevissen, K., Ghazi, A., De Samblanx, G. W., Brownlee, C., Osborn, R. W. and Broekaert, W. F. (1996) Fungal Membrane Responses Induced by Plant Defensins and Thionins. *J. Biol. Chem.,* **271,** 15018-15025.

Thomma, B. P. H. J., Van Esse, H. P., Crous, P. W. and De Wit, P. J. G. M. (2005) Cladosporium fulvum (syn. Passalora fulva), a highly specialized plant pathogen as a model for functional studies on plant pathogenic Mycosphaerellaceae. *Molecular Plant Pathology,* **6,** 379-393.

Tyler, B. M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R. H. Y., Aerts, A.*, et al.* (2006) Phytophthora genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science,* **313,** 1261-1266.

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

Van't Slot, K. A. E., Gierlich, A. and Knogge, W. (2007) A single binding site mediates resistance- and disease-associated activities of the effector protein NIP1 from the barley pathogen Rhynchosporium secalis. *Plant Physiology,* **144,** 1654-1666.

Van den Ackerveken, G. F., Van Kan, J. A., Joosten, M. H., Muisers, J. M., Verbakel, H. M. and De Wit, P. J. (1993a) Characterization of two putative pathogenicity genes of the fungal tomato pathogen Cladosporium fulvum. *Molecular plant-microbe interactions : MPMI,* **6,** 210-215.

Van Den Ackerveken, G. F. J. M., Van Kan, J. A. L. and De Wit, P. J. G. M. (1992) Molecular analysis of the avirulence gene avr9 of the fungal tomato pathogen Cladosporium fulvum fully supports the gene-for-gene hypothesis. *Plant Journal,* **2,** 359-366.

Van Den Ackerveken, G. F. J. M., Vossen, P. and De Wit, P. J. G. M. (1993b) The AVR9 race-specific elicitor of Cladosporium fulvum is processed by endogenous and plant proteases. *Plant Physiology,* **103,** 91-96.

Van Den Burg, H. A., Harrison, S. J., Joosten, M. H. A. J., Vervoort, J. and De Wit, P. J. G. M. (2006) Cladosporium fulvum Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection. *Molecular Plant-Microbe Interactions,* **19,** 1420-1430.

Van Der Does, H. C. and Rep, M. (2007) Virulence genes and the evolution of host specificity in plant-pathogenic fungi. *Molecular Plant-Microbe Interactions,* **20,** 1175-1182.

Van der Hoorn, R. A., De Wit, P. J. and Joosten, M. H. (2002) Balancing selection favors guarding resistance proteins. *Trends Plant Sci,* **7,** 67-71.

Van Parijs, J., Broekaert, W. F., Goldstein, I. J. and Peumans, W. J. (1991) Hevein: an antifungal protein from rubber-tree (Hevea brasiliensis) latex. *Planta,* **183,** 258-264.

Verkest, A., Manes, C. L. D. O., Vercruysse, S., Maes, S., Van Der Schueren, E., Beeckman, T.*, et al.* (2005) The cyclin-dependent kinase inhibitor KRP2 controls the onset of the endoreduplication cycle during Arabidopsis leaf development through inhibition of mitotic CDKA;1 kinase complexes. *Plant Cell,* **17,** 1723-1736.

Vervoort, J., Van Den Hooven, H. W., Berg, A., Vossen, P., Vogelsang, R., Joosten, M. H. A. J.*, et al.* (1997) The race-specific elicitor AVR9 of the tomato pathogen Cladosporium fulvum: A cystine knot protein. Sequence-specific 1H NMR assignments, secondary structure and global fold of the protein. *FEBS Letters,* **404,** 153-158.

Wang, X., Allen, R., Ding, X., Goellner, M., Maier, T., de Boer, J. M.*, et al.* (2001) Signal peptide-selection of cDNA cloned directly from the esophageal gland cells of the soybean cyst nematode Heterodera glycines. *Mol Plant Microbe Interact,*

14, 536-544.

Wang, X., Mitchum, M. G., Gao, B., Li, C., Diab, H., Baum, T. J.*, et al.* (2005) A parasitism gene from a plant-parasitic nematode with function similar to CLAVATA3/ESR (CLE) of Arabidopsis thaliana. *Molecular Plant Pathology,* **6,** 187-191.

Weinl, C., Marquardt, S., Kuijt, S. J. H., Nowack, M. K., Jakoby, M. J., Hu?lskamp, M.*, et al.* (2005) Novel functions of plant cyclin-dependent kinase inhibitors, lCK1/KRP1, can act non-cell-autonomously and inhibit entry into mitosis. *Plant Cell,* **17,** 1704-1722.

Westerink, N., Brandwagt, B. F., De Wit, P. J. G. M. and Joosten, M. H. A. J. (2004) Cladosporium fulvum circumvents the second functional resistance gene homologue at the Cf-4 locus (Hcr9-4E) by secretion of a stable avr4E isoform. *Molecular Microbiology,* **54,** 533-545.

Westerink, N., Roth, R., Van den Burg, H. A., De Wit, P. J. G. M. and Joosten, M. H. A. J. (2002) The AVR4 elicitor protein of Cladosporium fulvum binds to fungal components with high affinity. *Molecular Plant-Microbe Interactions,* **15,** 1219-1227.

Wevelsiep, L., Kogel, K. H. and Knogge, W. (1991) Purification and characterization of peptides from Rhynchosporium secalis inducing necrosis in barley. *Physiological and Molecular Plant Pathology,* **39,** 471-482.

Wevelsiep, L., Ru?pping, E. and Knogge, W. (1993) Stimulation of barley plasmalemma H+-ATPase by phytotoxic peptides from the fungal pathogen Rhynchosporium secalis. *Plant Physiology,* **101,** 297-301.

Whisson, S. C., Boevink, P. C., Moleleki, L., Avrova, A. O., Morales, J. G., Gilroy, E. M.*, et al.* (2007) A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature,* **450,** 115-118.

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

Win, J., Morgan, W., Bos, J., Krasileva, K. V., Cano, L. M., Chaparro-Garcia, A.*, et al.* (2007) Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Cell,* **19,** 2349-2369.

Yeaman, M. R. and Yount, N. Y. (2003) Mechanisms of antimicrobial peptide action and resistance. *Pharmacological Reviews,* **55,** 27-55.

Yeaman, M. R. and Yount, N. Y. (2007) Unifying themes in host defence effector polypeptides. *Nature Reviews Microbiology,* **5,** 727-740.

Yount, N. Y. and Yeaman, M. R. (2004) Multidimensional signatures in antimicrobial peptides. *Proceedings of the National Academy of Sciences of the United States of America,* **101,** 7363-7368.

Zeng, X. C., Corzo, G. and Hahin, R. (2005) Scorpion venom peptides without disulfide bridges. *IUBMB Life. 2005;,* **57,** 13-21.

Zuncke, U. and Eisenback, J. (1998) Morphology and ultrastructure. In: *The cyst nematodes.* (Sharma, S., ed.^eds.). Dordrecht: Kluwer Academic Publishers, pp. 31-56.

Summaries

Summary

Plant parasitic nematodes secrete a cocktail of effector molecules, which are involved in several aspects of the interaction with the host, eg. in host defense suppression, in migration and in feeding cell formation. In this thesis, we performed the first study on 10 novel peptide genes, believed to be important for parasitism of the potato cyst nematode, *Globodera rostochiensis.* Nine of the peptide genes described here belong to the *SECPEP* gene family. The *SECPEP* genes are all expressed in the dorsal esophageal gland, which is one of the main sites for the production of effector molecules. This, together with the predominant expression in preparasitic and early parasitic juvenile nematodes, makes it very likely that the *SECPEPs* code for effector peptides essential for succesful infection and feeding site formation.

In chapter 2, we show that diversifying selection is a likely driver of the molecular evolution of the *SECPEPs*. The sequences of the mature peptides appear to be highly diverse, while the non-coding 3'UTR and intronic regions as well as the region coding for the signal peptide for secretion are relatively conserved. In fact, a pairwise comparison of the SECPEPs reveals no significant sequence similarity between family members at all. In chapter 5 we further speculate on a possible role for RNA-editing as a mechanism to yield hypervariability in the *SECPEPs*, because the sequence diversity at the transcript level significantly exceeds that of the genomic locus. Chapter 5 further elaborates on the analysis of *trans*-splicing in *SECPEP1* transcripts. We show that *SECPEP1* transcripts are *trans*-spliced to a surprising diversity of novel spliced-leader sequences. The first approach to unravel the role of the members of the *SECPEP* family in plant parasitism, is described in chapter 4. We generated transgenic potato and *Arabidopsis* plants expressing *SECPEP3* while using the CaMV 35S promotor. The phenotype associated with SECPEP3 in both potato and *Arabidopsis* plants includes a reduction of root growth and an alteration of the leaf morphology. The SECPEP3 peptide harbors several sequence motifs first found in the cyclin-dependent kinase inhibitors ICK1/KRP1, SIM and Smr1. We, therefore, suggest a role for SECPEP3 in cell

cycle alteration in nematode feeding site formation. Although the SECPEP genes show only a low level of primary sequence similarity, all code for positively charged, hydrophilic peptides with a C-x-G γ-core motif (chapter 2). These are characteristics typical for host defense peptides, and in chapter 6 we investigate whether these characteristics are also found for other peptides involved in plant-parasite interactions. We show that a considerable number of these effector peptides share a positive charge, hydrophilicity and C-x-G γ-core motif with the SECPEPs, and we speculate on a role for the positive charge in peptide-ligand interaction.

In chapter 3 we describe the NEMPEP peptide, secreted by *G. rostochiensis*. NEMPEP is also a positively charged, hydrophilic peptide with a C-x-G γ-core motif, although it is genetically unrelated to the SECPEP gene family. During the life cycle of *G. rostochiensis*, the expression pattern of *NEMPEP* reveals a striking regulation. *NEMPEP* is highly expressed in preparasitic juveniles and in the parasitic life stages after initial feeding cell formation. However, *NEMPEP* expression was hardly detectable in the juveniles just after entering the plant root. Several disease resistance genes condition nematode resistance at the onset of parasitism. The downregulation of *NEMPEP* at exactly this timepoint could be a strategy to avoid recognition by the host's immune system. *In planta* expression of *NEMPEP*, as a fusion to GFP, shows that NEMPEP accumulates in the nucleolus of tobacco cells. Potato plants transformed with *35S::NEMPEP* were slow at forming roots and the internodes between the leaflets were shortened. This, together with a reduced transformation efficiency, led us to hypothesize a role for NEMPEP in cytokinin signaling (Chapter 3).

Currently, there are two models regarding the functional role of the SECPEPs and NEMPEP. The first one concerns a role as an antimicrobial peptide, which could protect the host plant against secondary infections by opportunistic microbes. As a competing hypothesis, the high hydrophilicity of the peptides may point to a role as peptide hormone. As such, they may be involved in redirecting cell cycle or hormonal regulation upon feeding cell formation.

Samenvatting

Het aardappelcystenaaltje (*Globodera rostochiensis*) injecteert plantencellen met een brede verscheidenheid aan eiwitten. Deze eiwitten vergemakkelijken de migratie van het aaltje door de plant, en zorgen ervoor dat het immuunsysteem van de plant onderdrukt wordt. Verder spelen deze eiwitten een essentiële rol in de vorming van een zogenaamde voedingscel. Dit is een actieve cel, die ontstaat na de samenvoeging van wel 200 cellen.

In dit proefschrift beschrijven we de studie van tien verschillende typen kleine eiwitten, ook wel peptiden genoemd, van het aardappelcystenaaltje. Negen van deze peptiden vertonen gelijkenis met elkaar, en worden daardoor gezien als een familie van eiwitten. Deze familie heet de SECPEP familie. In de hoodstukken 2 en 5 laten we zien dat de SECPEP familie getypeerd wordt door een enorme diversiteit, die veroorzaakt wordt door een hoge selectiedruk. In hoofdstuk 4 beschrijven we dat expressie van een van de familieleden, namelijk SECPEP3, in de aardappelplant leidt to vertraagde wortelgroei, en in sommige gevallen tot een veranderde bladvorm. SECPEP3 heeft twee sequentiemotieven die ook aanwezig zijn in ICK1/KRP1, in SIM en in Smr1. Dit zijn eiwitten die betrokken zijn bij de regulatie van de celcyclus. Mogelijkerwijs dat SECPEP3 hierin ook een rol speelt.

Het tiende peptide vertoont geen gelijkenis in samenstelling met enig ander bekend eiwit. Dit peptide is NEMPEP, en in hoofdstuk 3 laten we zien dat als de tabaksplant NEMPEP produceert, het peptide dan vooral in de nucleolus terechtkomt. NEMPEP vertoont een verrassend expressiepatroon gedurende de levenscyclus van de nematode. *NEMPEP* transcripten werden namelijk nauwelijks waargenomen in het parasitaire J2 levensstadium van het aaltje. Dit is het stadium waarin het aaltje door de plant migreert op zoek naar een geschikte plek in de wortel om een voedingscel te vormen. Het ontbreken van NEMPEP in deze beginfase van de infectie zou een strategie kunnen zijn om herkenning door het immuunsysteem van de plant te vermijden. Het is namelijk van een aantal resistentiegenen bekend dat vroege herkenning van het aaltje

essentieel is voor een adequate afweerreactie.

Zowel NEMPEP als de SECPEP peptiden hebben een positieve lading, zijn hydrofiel, en hebben een C-X-G γ-core sequentie motief. Deze eigenschappen worden ook gevonden in verschillende andere klassen van peptiden, vooral in zogenaamde antimicrobiële peptiden die een rol spelen in de bescherming tegen bacteriën (hoofdstuk 6). We kunnen niet uitsluiten dat de SECPEPs en NEMPEP ook een soortgelijke antimicrobiële activiteit bezitten, maar enige evidentie hiervoor ontbreekt vooralsnog. Bovendien wijst het fenotype van de planten die SECPEP3 en NEMPEP produceren eerder in de richting van een mogelijke rol voor deze peptiden bij de vorming van de voedingscel.

Curriculum vitae

Nikkie Elisa Marie van Bers werd op 8 juni 1979 geboren in Maastricht, waar zij in 1991 naar 't Stedelijk, nu Bonnefanten College geheten, ging. In 1997 behaalde zij daar haar atheneumdiploma, en in datzelfde jaar vertrok ze naar Nijmegen om daar Biologie te gaan studeren aan de Katholieke Universiteit Nijmegen (KUN), nu Radboud Universiteit Nijmegen. Tijdens haar biologieopleiding voltooide ze drie stage- oftewel afstudeerprojecten. De eerste betrof de functionele analyse van Pistil specific Extensin-Like proteins (PELPIII) in de tabaksplant *Nicotiana tabacum*, en vond plaats op de afdeling Celbiologie van de Plant aan de KUN. Ook haar tweede stageproject deed ze aan de KUN, namelijk bij de afdeling Experimentele Plant Ecofysiologie. Hier verdiepte ze zich in de rol van planthormonen in de "shade/avoidance response" van *Arabidopsis thaliana.* Tijdens haar laatste stage project, bij de Leerstoelgroep Moleculaire Biologie (Wageningen Universiteit), deed ze een mutationele analyse van het "Movement Protein" van het Cowpea Mosaic Virus. In oktober 2002 begon ze als AIO bij het Laboratorium voor Nematologie. Haar onderzoek vond plaats in het kader van het NWO vernieuwingsimpuls project "A unique masterpiece of parasitism: mitogenic peptides and RanBPM-two antagonistic compounds in nematode secretions- redirect the cell cycle in plants to a feeding cell", en de resultaten van dit promotieonderzoek staan beschreven in dit proefschrift. Sinds 1 maart 2008 is Nikkie werkzaam als onderzoeker (postdoc) op het NWO "Songbird Genomics" project bij de Leerstoelgroep Fokkerij en Genetica (Wageningen Universiteit).

Nawoord

Tijdens het schrijven van dit proefschrift schoten me regelmatig stellingen te binnen over het wel en (vooral) wee van een promovendus. Eén ervan was dat het schrijven van een proefschrift te vergelijken is met het aanbrengen van orde in een gigantische berg van ruim vier jaar onderzoeksresultaten. Een ander was dat het beste advies voor andere proefschriftschrijvers is: “kiezen op elkaar en doorgaan”. De kiezen kunnen inmiddels bijna van elkaar, en de data zijn ordelijk bij elkaar gebracht in een handzaam boekje. Uiteraard was dit niet mogelijk geweest zonder de hulp van alle collega’s van de leerstoelgroep nematologie, en in het bijzonder van Hein, wiens engelengeduld achter de microscoop heel wat mooie plaatjes heeft opgeleverd. Hein, bedankt! Verder wil ik Geert, Aska, Ling en Jaap bedanken voor het mogelijk maken van dit project en voor al hun ideeën en suggesties. Liesbeth, je was een ontzettend gezellige kamergenote, en ik ben blij dat we tijdens onze maandelijkse lunches rustig verder kunnen kletsen. Kamila, I´m very glad that we´ll be together on the stage on the 17th of June, although I doubt if you think the same about this. Thanks for being my friend and thanks for organizing all kind of great social events! Uiteraard wil ik ook al m´n andere collega´s en oud collega´s bedanken voor alle gezelligheid op het lab en tijdens de koffiepauzes, lunches, borrels, uitjes, barbecues en sinterklaasvieringen. Erik, bedankt voor de fantastische CLSM plaatjes. Rikus en Jan R., bedankt voor de talloze praktische adviezen en hulp bij het weer op gang brengen van dwarse apparatuur. Patrick, Anna T. and “all the other Polish”: the dress-up parties are really unforgettable. You are great fun, thanks! Anna P. thanks for your help with the layout of my thesis. Lisette en Linda, bedankt voor alle hulp, o.a. bij het vechten met de printer en het inbinden van de leesversie van dit proefschrift. Alle “nema”- en ABG AIO´s: heel veel succes met jullie onderzoek en met het schrijven van jullie proefschrift.
Erwin, je was een heel gezellige buurman, jammer dat ik je inmiddels twee keer als buurman kwijt ben, en veel succes met het vinden van een “droombaan”. Alle meiden

van 't Stedelijk (Dominique, Marike, Judith P. en Judith E., Janneke, Tessel, Anouk, Saskia, Iris en Hilde) hartstikke bedankt voor alle gezellige avondjes en weekendjes weg. Angelique, Stéphany en Karen: bedankt voor jullie luisterend oor naar alle frustraties van deze AIO tijdens onze etentjes. En Karen, gezellig dat je m'n paranimf wilt zijn, en jij natuurlijk helemaal bedankt voor de enorme interesse die je altijd toont voor het wel en wee binnen en buiten mijn onderzoek. Je bent een geweldige vriendin!
Martien, Richard, Marcel en Kees: bedankt het vertrouwen om mij aan te stellen op het "Songbird Genomics" project. We gaan er iets moois van maken!
Papa en mama, zonder jullie was ik nooit zover gekomen als ik nu ben. Jullie, Sanne en Michiel en de rest van mijn familie: bedankt voor alle steun en vertrouwen! Also Dave's family: thanks for all the fun that is always around you.
Sid, door de enorme vrolijkheid die jij altijd meebrengt is het me gelukt om alle proefschrift-stress af en toe helemaal te vergeten, en laat je me zien waar het echt om draait in het leven. Je bent een fantastisch ventje! En voor je broertje of zusje in m'n buik: kleine hummel, geweldig dat we samen op het podium zullen staan. Dave, thanks for everything you are, and thanks for all your support and help. You are doing great!
Nikkie

Education Statement of the Graduate School

Experimental Plant Sciences

The Graduate School EXPERIMENTAL PLANT SCIENCES

Issued to: **Nikkie van Bers**
Date: **17 June 2008**
Group: **Nematology, Wageningen UR**

1) Start-up phase	*date*	*cp*
► **First presentation of your project**		
Mitogenic peptides in potato cyst nematodes	Dec 10, 2002	1.5
► **Writing or rewriting a project proposal**		
► **Writing a review or book chapter**		
Peptides in pathogenicity	Jan-Mar, 2008	6.0
► **MSc courses**		
► **Laboratory use of isotopes**		
Radiation course 'Safe handling with radioactive materials and sources' level 5B	Sept, 2003	1.5
Subtotal Start-up Phase		9.0

2) Scientific Exposure	*date*	*cp*
► **EPS PhD student days**		
EPS PhD student day, Utrecht	Mar 27, 2003	0.3
EPS PhD student day, Nijmegen	Jun 02, 2005	0.3
► **EPS theme symposia**		
Theme 2 Interactions between plants and biotic agents	Jan 10, 2003	0.3
Theme 2 Interactions between plants and biotic agents	Dec 12, 2003	0.3
Theme 1 Developmental Biology of Plants	Feb 17, 2004	0.3
Theme 1 Developmental Biology of Plants	Apr 26, 2005	0.3
Theme 2 Interactions between plants and biotic agents	June 23, 2005	0.3
► **NWO Lunteren days and other National Platforms**		
NWO-ALW Experimental Plant Sciences meeting	Apr 7-8, 2003	0.6
NWO-ALW Experimental Plant Sciences meeting	Apr 5-6, 2004	0.6
NWO-ALW Experimental Plant Sciences meeting	Apr 4-5, 2005	0.6
NWO-ALW Experimental Plant Sciences meeting	Apr 3-4, 2006	0.6
NWO-ALW Experimental Plant Sciences meeting	Apr 2-3, 2007	0.6
► **Seminars (series), workshops and symposia**		
Cell cycle symposium, Gent, Belgium	Mar, 2003	0.3
Flying Seminars 2003	2003	0.3
Workshop GFP and Luc, Application of "Light"-Reporters in Biology	Apr 11-12, 2005	0.6
Real Time PCR and Gene expression analysis, seminar series from Biorad	Jun 07, 2005	0.3
► **Seminar plus**		
► **International symposia and congresses**		
XXVII Int Symp of the European Society of Nematologists, Rome	Jun 14-18, 2004	1.5
XII Int Congress on Molecular Plant-Microbe Interactions, Mérida, Mexico	Dec 14-19, 2005	1.8
XXVII Int Symp of the European Society of Nematologists, Blagoevgrad, Bulgaria	Jun 5-9, 2006	1.5
► **Presentations**		
XXVII Int Symp of the European Society of Nematologists, Rome (Poster)	Jun 14-18, 2004	0.7
NWO-ALW Experimental Plant Sciences meeting	Apr 3-4, 2006	0.7
XXVII Int Symp of the European Society of Nematologists, Blagoevgrad, Bulgaria (Oral presentation)	Jun 5-9, 2006	0.7
NWO-ALW Experimental Plant Sciences meeting	Apr 2-3, 2007	0.7
► **IAB interview**	Jun 2, 2005	0.7
► **Excursions**		
Subtotal Scientific Exposure		14.9

3) In-Depth Studies	*date*	*cp*
► **EPS courses or other PhD courses**		
EPS Autumnschool Disease Resistance in Plants	Oct 14-16, 2002	0.9
EPS Summerschool Signaling in Plant Development and defence: towards Systems Biology	Jun, 19-21, 2006	0.9
► **Journal club**		
PhD student literature discussion group Nematology	2002-2007	3.0
► **Individual research training**		
Subtotal In-Depth Studies		4.8

4) Personal development		*date*	*cp*
►	**Skill training courses**		
	Cursus Begeleiden van Onderwijsgroepen	Feb, 2005	0.6
	Cursus Afstudeervakkers begeleiden	Feb, 2004	0.6
	Cursus Culturele diversiteit in onderwijs	Apr, 2005	0.6
	Scientific Writing	Sept-Nov, 2003	1.5
►	**Organisation of PhD students day, course or conference**		
►	**Membership of Board, Committee or PhD council**		
		Subtotal Personal Development	3.3
		TOTAL NUMBER OF CREDIT POINTS*	**32.0**

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

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

This research was funded by the NWO vernieuwingimpuls program

Drukker: Print Partners Ipskamp, Enschede

Cover: Pre-parasitic juvenile of *Globodera rostochiensis.* Cover picture taken by Hein Overmars.