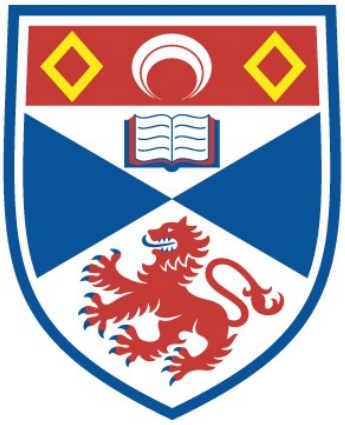
MOLECULAR ANALYSIS OF THE EGGSHELL AND CUTICLE SURFACE OF THE POTATO CYST NEMATODE, *GLOBODERA ROSTOCHIENSIS*

James Price

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



2019

Molecular analysis of the eggshell and cuticle surface of the potato cyst nematode, *Globodera rostochiensis*

James Price

University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of

Doctor of Philosophy (PhD)

at the University of St Andrews

May 2019

Declarations

Candidate's declaration

I, James Price, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 56,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree.

I was admitted as a research student at the University of St Andrews in September 2015.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date Signature of supervisor

Date Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, James Price, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

Date Signature of candidate

Date Signature of supervisor

Date Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, James Price, understand that by declaring that I have original research data or digital outputs, I should make every effort in meeting the University's and research funders' requirements on the deposit and sharing of research data or research digital outputs.

Date Signature of candidate

Permission for publication of underpinning research data or digital outputs

We understand that for any original research data or digital outputs which are deposited, we are giving permission for them to be made available for use in accordance with the requirements of the University and research funders, for the time being in force.

We also understand that the title and the description will be published, and that the underpinning research data or digital outputs will be electronically accessible for use in accordance with the license specified at the point of deposit, unless exempt by award of an embargo as requested below.

The following is an agreed request by candidate and supervisor regarding the publication of underpinning research data or digital outputs:

No embargo on underpinning research data or digital outputs.

Date Signature of candidate

Date Signature of supervisor

Date Signature of supervisor

K, K. R., E. M. & J. H.

Acknowledgments

I am not one for grand declarations and pouring my heart into words, however, although by this point my creativity lacks, my gratitude does not.

I have thoroughly enjoyed my PhD but don't think I could have if it was not for the supervision provided by both Prof. John Jones and Prof. Terry Smith. They were both always willing to listen to my concerns and questions and supported me whenever I was doubting my work. Both have taught me a lot that I hope to replicate one day and for all of this, I am eternally grateful.

I cannot thank the Perry foundation enough for the support they provided. Where others didn't think this work was worth funding, they stepped up and provided me with more than I hoped for. Not only did this allow me to carry out the work in this thesis but also enabled me to travel and present my results conferences in London, Durham, Braga (Portugal) and Ghent (Belgium).

Time management for this work was often a struggle due to actively working between the James Hutton Institute and the University of St. Andrews. However, colleagues from both sites always made me feel at home wherever I was working. I would like to extend a special thanks to friends from the NemaLab at the James Hutton Institute; Kerry, Shona, Akis and Jamie. My gratitude also goes to colleagues, past and present from the TKS lab at the University of St Andrews. Throughout my research, I have met many other PhD students and friends whom I would like to thank for always saying 'why not?' when I suggested a long weekend on a distant Scottish island. Merci aussi à Adeline, pour avoir supporté mes excentricités, m'avoir épaulé, et pour toute l'aide que tu m'as apportée au cours des deux dernières années.

Finally, I must thank my parents, Andrew and Patricia Price. I cannot put into words how grateful I am for their encouragement, their presence and their constant support.

Abstract

Plant-parasitic nematodes regulate their interactions with the external environment and their host. The surface of the nematodes, whether the cuticle or the eggshell, plays a key role in this process. Work undertaken in this project aimed to exploit new genome resources for potato cyst nematodes (PCN) to allow characterisation of these structures.

The eggshell is a key survival structure for PCN. Hatching of juvenile potato cyst nematodes occurs in response to host derived hatching factors. Little is known about the molecular mechanisms that underpin hatching in response to these host diffusates and nothing is known about the molecular composition of the PCN eggshell. Methods that allow extraction of proteins and lipids from large numbers of isolated PCN eggshells were developed. These extraction techniques permitted identification of a range of proteins present in the eggshell, including chondroitin proteoglycans and a calcium dependent phospholipid-binding annexin. The annexin was focused on due to the apparent major role of calcium in PCN hatching. This annexin was localised to the eggshells of *Globodera rostochiensis* by immunofluorescence. To our knowledge, this makes the annexin the first eggshell protein to be identified and localised in any plant-parasitic nematode. The annexin was shown to bind to phospholipids and RNAi of the annexin showed that reducing expression of the eggshell annexin alters nematode hatching patterns in response to host hatching factors.

The eggshell permeability barrier is essential for the development of a juvenile within the embryo. Eggshell lipid analysis has given insight into components of the permeability barrier. Identified lipid species highlight potential function within the nematode hatching cascade through calcium interaction or down stream signalling. Attempts to identify the presence of nematode specific glycolipids, ascarosides, within the PCN eggshell showed that ascarosides or ascarylose were not detectable within PCN eggs, cysts or juveniles.

It has previously been shown that the infective juveniles are able to alter their surface composition to avoid damage from host defence mechanisms. Here, methods that allow labelling of proteins present on the cuticle surface of PCN were developed. Three proteins; Galectin-5, Cri-2 and a nematode specific hypothetical protein that are present on the surface of PCN second stage juveniles were further investigated. These proteins were tested for their potential to suppress host defences through a variety of protein-small molecule interactions. This analysis showed that surface coat associated galectins could offer protection to an invasive juvenile through sequestering cell wall breakdown products, silencing detection of the pathogen.

Abbreviations

ALF160	Ascaroside #18 (produced in-house)
ANOVA	Analysis of variance
APN	Animal-parasitic nematode
ASCR	Ascaroside
avr	Avirulence genes
BLAST	Basic Local Allignment Search Tool
BSA	Bovine Serum Albumin
cDNA	Complementary DNA
CE	Collision energy
CpG	Chondroitin proteoglycan
cps	Counts per second
cv.	Cultivar
Da	Daltons
DAG	Diacylglycerol
DAMP	Damage Associated Molecular Pattern
DCM	Dichloromethane
DEPC	Diethyl pyrocarbonate
dH_2O	Distilled water
DIG	Digoxegenin
DNA	Deoxyribonucleic acid
DNTB	5,5'-dithiobis-2-nitrobenzoic acid
dpi	Days post infection
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ESI-MS-MS	Electrospray tandem mass spectrometry
ESI-TOF	Electrospray ionisation time of flight mass spectrometry

EV	Empty vector
FA	Fatty acid
FAMES	Fatty acid methyl esters
FAR	Fatty acid and retiol binding protein
FLN	Free living nematode
g	Relative gravitational force
GC-MS	Gas chromatography mass spectrometry
GFP	Green Fluorescent Protein
HCl	Hydrochloric
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High performance liquid chromatography
HR	Hypersensitive reaction
HRP	Horse radish peroxidase
h	Hours
ISH	*in situ* hybridisation
J2	Second stage juvenile
LB	Lysogeny broth
m/z	Mass to charge ratio
MCA	Multi channel analysis
MS	Mass spectrometry
MSTFA	N-Methyl-N-(trimethylsilyl)trifluoroacetamide
MUSCLE	Multiple sequence comparison by log-expectation
NB-LRR	Nucleotide binding and leucine rich repeat
NBT	Nitro blue tetrazolium
OD_{600}	Optical density measured at 600nm
PA	Phosphatic acid
PAMP	Pathogen Associated Molecular Pattern
PBS	Phosphate buffered saline

PC	Phosphatidylcholine
PCN	Potato cyst nematode
PCR	Polymerase chain reaction
pDNA	plasmid DNA
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PI	Phosphatidylinositol
$PI(3,4,5)P_3$	Phosphatidylinositol (3,4,5)-trisphosphate
$PI(4,5)P_2$	Phosphatidylinositol 4,5-bisphosphate
PI4P	Phosphatidylinositol 4-phosphate
PPN	Plant-parasitic nematode
PRR	Pattern recognition receptor
PS	Phosphatidylserine
Q-TOF	Quadrupole time of flight mass spectrometry
R gene	Resistance gene
r.c.f.	Relative centrifugal force
RFP	Red Fluorescent Protein
RKN	Root knot nematode
RLU	Relative light Units
RNA	Ribonucleic acid
RNAi	RNA interference
ROS	Reactive oxygen species
rpm	Revolutions per minute
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SDW	Sterile distilled water
SOC	Super optimal broth with catabolite repression
SQRT-PCR	Semiquantitative reverse transcription polymerase chain reaction

TAG	Triacylglycerol
TBE	Tris/Borate/EDTA
TMS	Trimethylsilyate
TRD/PRD	Tomato/Potato root diffusate
Tris	Tris (hydroxymethyl) aminomethane
UV	Ultra-violet
WT	Wildtype
Y2H	Yeast-two-hybrid

Table of Contents

Figures

Tables

1. Introduction

The phylum Nematoda consists of over 25,000 described species, although it is thought that several million different species of nematode may exist (Zhang, 2013). Nematodes are multicellular eukaryotes that have adapted to live in a wide range of habitats, whether free-living in aqueous or terrestrial environments or living within other organisms as parasites (McSorley, 2003). The free-living nematode *Caenorhabditis elegans* has become a model organism for study of eukaryotic biology, more is known about the biology of this organism than almost any other Eukaryote on earth. Cultures of *C. elegans* are easy and cheap to maintain, while their short generation time and genetic amenability makes this species ideal for a range of studies (Kaletta & Hengartner, 2006).

One of the major influences nematodes have on human life is due to damage caused by parasitic species (Bird & Koltai, 2000). Animal-parasitic nematodes cause loss of life and illnesses to humans as well as causing large yield losses in livestock (Stepek *et al.*, 2006; Charlier *et al.*, 2009). In addition, plant-parasitic nematodes cause considerable yield losses in crops.

1.1 Plant-Parasitic Nematodes

The first plant-parasitic nematode (PPN) was documented in 1743 in wheat (Needham, 1743). Since then, nematodes have been found to parasitise all crops including all of the world's most important staple crops (Lambert & Bekal, 2002). Most PPN are soil-borne but their feeding is not always localised to plant roots. Nematodes have been found to be adapted to feed on plant stems, leaves and buds (Powers, 1992). Due to the enormous number of nematode species and their abilities to feed on a variety of plant tissues, annual crop losses globally from PPN has been estimated at $80 billion. However, this is thought to be an underestimate due to many growers not being aware of infections particularly in developing nations (Nicol *et al.*, 2011). The damage caused by PPN makes them one of the major risks to food security.

Plant-parasitic nematodes can be categorised by their feeding behaviour and lifestyles. All phytophagous nematodes feed using a stylet. The stylet resembles a hollow needle that can be used to pierce into plant cells allowing nutrients to be drawn out of the host (Zinovieva, 2014). In some cases nematodes induce changes in host plants, including profound changes in host gene expression, that allow them to feed for a prolonged period of time from one part of the host (Sijmons *et al.*, 1994).

Ectoparasitic nematodes use their stylet to feed on plant tissues whilst remaining in the soil. Ectoparasites can be migratory or sedentary. Sedentary ectoparasites gain growth advantages from establishing permanent feeding sites. Migratory ectoparasites switch between hosts and feed at several different sites (Wyss, 2002). These nematodes are often vectors of commercially important plant viruses. Remaining outside the roots exposes ectoparasites to environmental fluctuations, pathogens and predators (Decraemer & Hunt, 2006).

Endoparasitic nematodes feed from within the plant tissue. Again, endoparasites can be migratory or sedentary. Migratory endoparasites primarily feed, moult and reproduce within the plant tissue yet they are still able to transfer to new hosts through the soil. Migratory endoparasites establish no permanent feeding site and cause considerable damage within the plant as a result of their movement whilst searching for new cells to feed on. Parasitism for these nematodes is predominantly limited to cortical cells and they are able to enter and exit the host roots at any stage within their development (Moens & Perry, 2009). This can cause extensive and non-specific damage to the plant, often resulting in secondary fungal or bacterial infections in the roots (Zunke, 1990).

Sedentary endoparasites are the most economically important of the plant-parasitic nematodes, particularly the cyst nematodes (*Globodera* spp. and *Heterodera* spp.) and the root knot nematodes (*Meloidogyne* spp.) (Wyss, 1997). Both root-knot nematodes and cyst nematodes are soil-borne and need to infect roots to allow them to create a parasitic relationship in order to complete their life cycle. Infective juveniles of these species invade host roots releasing secretions into and around cells allowing initiation of their feeding sites. The induction of a permanent feeding site allows a greater nutrient

intake. The nematodes then become sedentary due to atrophy of their somatic muscles. In addition, the plant tissues will offer the nematode some protection from pests and environmental changes (Sijmons *et al.*, 1994; Decraemer & Hunt, 2006), which is reflected in a reduced complement of genes encoding immune system components in many endoparasitic nematode genomes (Kikuchi *et al.*, 2017). However, if the host dies, or causes the feeding site to be destroyed, the nematode will also die. Although both root knot and cyst nematodes share similar phenotypes and sedentary endoparasitic lifestyles they both evolved these strategies independently. The nematodes both develop differently and have an almost complete lack of overlap in effectors used for aiding infection of their hosts.

1.2 Potato Cyst Nematodes

Cyst nematodes are pathogens of various host species including soybeans, rice and potatoes. In 1959 in order to distinguish the round-ended cysts formed by the potato cyst nematode from those produced by other cyst nematodes, Skarbilovich established the sub-genus *Globodera* (Franklin, 2012). Previously all cyst nematodes were encompassed within the genus *Heterodera.* Twelve species have been identified within the genus *Globodera* (Bohlmann, 2015). The most heavily studied of these are *G. rostochiensis* and *G. pallida*. These species feed on the roots of plants found within the genus *Solanum* which contains cultivated species such as potatoes (*S. tuberosum*), tomatoes (*S. lycopersicum*) and aubergines (*S. melongena*) (Whitehead, 1985). However, *G. rostochiensis* and *G. pallida* are principally parasites of potatoes and are therefore collectively named the potato cyst nematodes (PCN).

Although they originated in South America, PCN are now found across the globe in many potato growing countries. Introduction of PCN to Europe is thought to have occurred as a result of importing potato material in the 19th century to combat the late blight pathogen *Phytophthora infestans* (Evans *et al.*, 1975). PCN infection throughout Europe then followed from distribution of seed potatoes across the continent. Europe has also acted as a secondary distribution centre for PCN into other regions of the world, due to movement of contaminated seed potatoes (Pickup *et al.*, 2018). A survey of fields used

for growing potatoes in the U.K. showed that *Globodera* cysts could be found in 64% of samples. 67% of infestations were solely *G. pallida*, 8% were *G. rostochiensis* and the remaining 25% included both species (Minnis *et al.*, 2002). Reports on yield losses suggest that PCN infections can cause above 50% reduction to tuber formation (Trudgill, 1986). In addition, potato for seed use cannot be grown on land infested with PCN as regulations designed to prevent spread of the pathogen prohibit this.

1.3 Lifecycle of the Potato Cyst Nematode

There is relatively little difference between the life cycle of *G. pallida* and *G. rostochiensis*. The life cycles of cyst nematodes have been extensively studied and the following lifecycle information has been summarised from several key texts including Turner & Rowe (2006), Koenning & Sipes (2013) and Lee (2002), and reviewed in **Figure 1.1**.

Juvenile cyst nematodes develop within an eggshell. Post-gastrulation the embryo remains inside the egg and increases in length whilst beginning to move. The first moult occurs whilst still in the egg and after this the stylet forms at the anterior pole of what is now the second stage juvenile (J2). Hatching occurs after the stylet has formed. However, hatch depends on many environmental factors such as soil temperature, moisture and presence of host root exudates. If optimal hatching conditions are not met quiescence or diapause can be provoked, preventing hatching. Cyst nematode eggs can remain dormant inside the cyst in the soil for over 20 years. This ability to survive in a preserved state within the soil has a great impact for retaining genetic diversity as well as ensuring continuation of the population should there be no host species present. If hatching conditions are optimal then the J2 will utilize its stylet to pierce the eggshell, facilitating eclosion (Doncaster & Seymour, 1973). However, not all juveniles will hatch at the same time and some juveniles will remain unhatched in the cyst in the soil, as a further survival strategy for the population.

Irrespective of the length of time spent within the eggshell, immediately after hatching infective J2s will search for host roots using chemical cues (Robinson & Perry, 2006).

Upon identification of a host plant the infective J2 will enter the root system. Once inside, the nematode migrates through host apoplast to the pericycle, where it identifies a cell suitable for transformation into a feeding site (the initial syncytial cell). The syncytium is formed by cell wall dissolution and fusion of protoplasts. These processes are induced by secretions from the juveniles' pharyngeal gland cells (Sobczak & Golinowski, 2009). The juvenile feeds on cell contents using a feeding tube, produced during each feeding cycle, as a filter to prevent destruction of the feeding site, which the nematode must keep alive for the duration of the life cycle. Once the feeding site is established and feeding has commenced, the moult to J3 occurs.

Although PCN reproduce sexually, the sex of the nematode is not genetically determined but instead develops at the J3 stage depending on availability of nutrients. Juveniles that successfully initiate a large feeding site will develop into females. Nematodes whose nutrient intake is restricted will develop into males (Trudgill, 1967).

Both males and females enter adulthood after the fourth moult. At this stage, the female has swollen so much that her body ruptures the cortex of the root, exposing her reproductive organs to the rhizosphere. Males exit the root and are attracted to pheromones released by adult females. The swollen body of the female can retain up to 500 eggs. As the swollen female begins to senesce, her internal organs begin to deteriorate and the cuticle tans forming the toughened cyst. The developing females of *G. rostochiensis* become a golden yellow colour whereas *G. pallida* females are paler. However, cysts of the two species are identical to the naked eye. When the crop is harvested or dies, the cysts drop off the roots into the soil where infective J2s develop inside the eggs, waiting within the cyst for optimal hatching conditions.

The length of the PCN lifecycle depends greatly on environmental conditions. However, from eclosion to adulthood takes around 38 to 45 days. The lifecycle of PCN is summarised by the schematic in **Figure 1.1**.

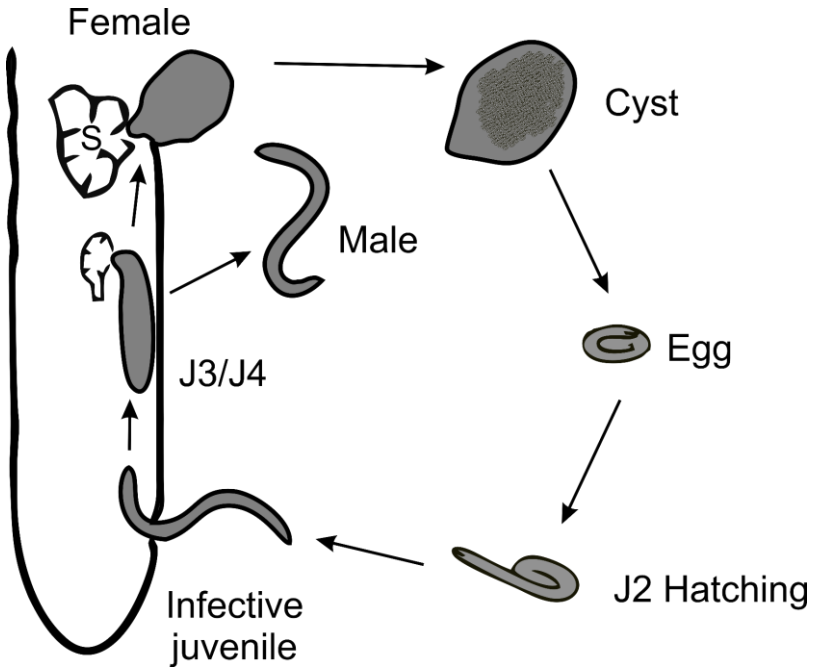


Figure 1.1 – Schematic of PCN lifecycle

PCN eggs, held within cysts in the soil, hatch in response to host root exudates. The juvenile (J) uses these chemical cues to locate the host where it moves through the root and migrates to an appropriate feeding site. If the feeding site can be sustained, the nematode will continue to feed, swelling until presenting outside the root and develop into a female. If the feeding site is restricted, the nematode will develop into a male and exit the host. After mating, eggs and juveniles begin to form within the female's body. The female's cuticle hardens to form the cyst.

Diagram modified from: https://www.plantpath.iastate.edu/scn/life_cycle

1.4 Control of Cyst Nematodes

Potatoes have become a staple food due to the crop resource potential per unit, providing more calories, protein and minerals than any other staple crops as well as being a high source of carbohydrate. An increase in the global consumption of potatoes and an increasing global population size have driven a surge in potato production (F.A.O., 2008). However, potato production is still adversely affected by pathogens, including PCN, and this needs to be controlled in order to maximise yields. A range of control methods are available for PCN, which can be separated into three main categories; biological, chemical and 'cultural' controls (S.A.S.A., 2010).

Biological control strategies involve using one living organism to control another. For PCN there are two main biological control tactics. The first technique is biocontrol though use of natural predators or pathogens. However, treatment with pathogens would require mass production and supply of the species, which remains technically challenging (Kerry & Hominick, 2002; Viaene *et al.*, 2006). In addition, although biological control has been successfully used in very specific agronomic systems, variability in environmental conditions frequently leads to variable efficacy of this control method.

Numerous chemical controls for potato cyst nematodes are used. However, many have been removed from the market due to environmental concerns (Turner & Rowe, 2006). There are two groups of chemical controls; fumigant and non-fumigant nematicides. Fumigants are centred on halogenated hydrocarbon compounds and chemicals that release methyl isothiocyanate (Haydock *et al.*, 2006). Not only are these chemicals non-specific, but their effectiveness is highly variable depending on soil conditions (Allen & Raski, 1950). Non-fumigant nematicides are either organophosphates or carbamates (Pree *et al.*, 1989). These inhibit acetylcholinesterase, a common enzyme which catalyses degradation of neurotransmitters, and so disrupts nervous system function. Both the non-fumigant and fumigant nematicides can have very broad effects. Furthermore, in order to be effective nematicides must be applied at appropriate concentrations and at a time when the nematode is in the soil.

Cultural controls attempt to regulate PCN infections through improved agricultural practices. Natural decline in PCN populations depends on numerous factors but the greatest drop in viable eggs is seen within the first six years of infection (S.A.S.A., 2010). Therefore, longer crop rotation periods are suggested to avoid growing crops at times of peak nematode soil populations. *G. pallida* exhibits a much slower rate of decline than *G. rostochiensis* meaning that crop rotation alone will not have a significant beneficial impact on reducing PCN population size (Trudgill *et al.*, 2014). PCN spread can be minimised by washing any equipment or machinery that may have been exposed to PCN before moving between different fields. Tighter control methods and testing of

potentially infected soils to reduce the potential of further dispersal should be enforced. Although these methods are useful for reducing PCN infections spreading between sites they do very little in terms of reducing or repairing damage caused by the pathogen. Currently PCN infections are treated using combinations of biological, chemical and cultural control mechanisms in an integrated pest management system. However, each year there are still substantial yield losses.

The most effective current control mechanism is through growing nematode resistant cultivars. Resistance to *G. rostochiensis* is mostly provided by cultivars containing the H1 gene. This gene confers resistance to the *G. rostochiensis* pathotypes Ro1 and Ro4 and has been mapped to chromosome V (Barone *et al.*, 1990). Upon initiation of the nurse cell for feeding, H1 brings about a hypersensitive response. This results in the feeding site becoming surrounded by necrotic cells, which degrades the site within a week (Rice *et al.*, 1985). H1 can now be found in many potato cultivars used across Europe. Although the introduction of such crops drastically reduced damage to yields caused by *G. rostochiensis*, use of these cultivars led to selection for *G. pallida,* which has subsequently become more prevalent across the U.K. (Atkinson, 2002). Loci conferring resistance to *G. pallida* have also been identified. However, due to the greater array of virulence seen in the *G. pallida* populations within Europe, resistance against this species tends to be quantitative and encoded by quantitative trait loci, rather than a single major gene, making breeding challenging (Rouppe van der Voort *et al.*, 1997; Van der Vossen *et al.*, 2000; Caromel *et al.*, 2003).

1.5 Plant defences against invasive species

Endoparasitic nematodes must surpass plant defences to successfully initiate their feeding site. Unlike animals, plants lack an adaptive immune system. Instead, they rely upon innate immunity responses to signalling from wound or infection sites (Jones & Dangl, 2006).

Native immunity in plants comes from specific resistance (R) genes. There are many different categories of R gene, some of which encode pattern recognition receptors

(PRRs). The most abundant classes contain genes encoding nucleoside-binding and leucine rich repeat (NB-LRR) proteins (Hammond-Kosack & Jones, 1997). Proteins encoded by R genes respond to specific proteins released by the nematode during infection. These proteins are translated from avirulence (avr) genes. Resistance proteins and avr proteins may not always interact directly. The 'guard hypothesis' suggests that host R proteins monitor cellular targets that parasite avr proteins are directed towards (Dangl & McDowell, 2006). This gene-for-gene relationship exhibited throughout plant-pathogen interactions results from the high selection pressures on R and avr genes consequently leading to pathogen driven evolution.

The function and evolution of plant defence systems can be described by the 'zig-zag' model. Upon infection, the host plant's PRRs allow detection of pathogen-associated molecular patterns (PAMPs). PAMPs can be a wide variety of molecules ranging from classically conserved elicitors such as glycoproteins to more unconventional non-protein targets, for example, lipopolysaccharides. These molecules are not found in the host and are essential for general fitness of the pathogen meaning they are often strongly conserved (Zipfel & Robatzek, 2010). Binding of PRRs to PAMPs results in the host activating PAMP triggered immunity (PTI). This immunity results in activation of generic antimicrobial responses such as release of reactive oxygen species (ROS) and activation of the mitogen-associated protein kinase (MAPK) pathway (Gimenez-Ibanez & Rathjen, 2010).

The high selection pressures presented by PTI has resulted in some pathogens developing counter defences known as effectors. Release of these effectors suppresses PTI resulting in effector-triggered susceptibility (ETS) allowing the pathogen to regain control within the host (Jones & Dangl, 2006).

Plants counteract these adapted pathogens using a new phase of immunity known as effector-triggered immunity (ETI). ETI is induced when the product of a resistance gene, most frequently a NB-LRR protein, recognises the presence of an effector (termed an avirulence gene). If the ETI response is strong enough the plant may force a hypersensitive response (HR) in infected cells. This response is harmful to the host,

resulting in localised cell death in infected areas to prevent any further contagion (Spoel & Dong, 2012; Feechan *et al.*, 2015). If the pathogen is further along the evolutionary arms race then it may contain additional effectors capable of reinitiating ETS and preventing the HR.

Although the zig-zag model is able to convey an understating of how plant immunity to pathogens functions it is simply a model. Realistically, all portions of this model whether PTI, ETS or ETI are more likely to occur simultaneously (Pritchard & Birch, 2014).

1.5.1 Plant defences against cyst nematodes

The zig-zag model can be applied to nematode infections. Upon infection, the nematode will begin to move through the apoplast before locating a potential feeding site. Plant PRRs presented into the apoplast will recognise nematode PAMPs. Cell wall break down products (known as Damage Associated Molecular Patterns - DAMPs) will also be recognised via PRRs initiating an invasion triggered immunity (Malinovsky *et al.*, 2014). Further PAMPs come in the form of small organic molecules instead of proteins. Several PPN (*Meloidogyne*) were investigated for PAMPs formed from ascarylose containing glycolipids, ascarosides. Ascaroside #18 was shown to induce local defence responses, increasing expression of PTI genes (Manosalva *et al.*, 2015). Receptors within the plant for many of these nematode associated molecular patterns remain unknown.

Mitogen activated protein kinases (MAPKs) have been shown to be implemented in response to cyst nematodes (Sidonskaya *et al.*, 2016), however, exact MAPK signalling occurring throughout nematode infection remains unknown. Further plant defences such as ROS, protease inhibitors and proteolytic enzymes are also released. ETS is triggered by nematode proteins released into the apoplast to supress host immune responses. For example, a *Globodera rostochiensis* venom allergen like protein (Gr-VAP-1) has been shown to aid infection of unrelated pathogens when transiently expressed in the apoplast (Lozano-Torres *et al.*, 2014).

Numerous R genes providing resistance to PCN have been mapped and a few have been cloned. However, the only example of an R/avr pairing in PCN is with Gpa2 and RBP1.

RBP-1 is a *G. pallida* secreted protein which induces HR mediated cell death via binding to the NB-LRR, Gpa2 (Sacco *et al.*, 2009). Here, the cytoplasmic Gpa2 recognises RBP-1 after secretion from the dorsal oesophageal gland, through the stylet into the host cell.

1.6 The nematode cuticle

Ultimately it is the nematode's body wall that is the main contact barrier between host and pathogen. If the nematode is to reach and remain active at an appropriate feeding site this barrier must be adaptive and help prevent recognition by the host or suppress responses if recognition does occur.

The cuticle is reasonably conserved in overall structure and composition between nematode species, originating from the hypodermis and seam cells which surround the body. At each life stage the cuticle is shed and a new one must be formed. The section below describing the nematode cuticle and associated tissues has been summarised using information from Lee, (2002a); Lints & Hall, (2009); Baldwin & Handoo, (2018). A schematic for the PCN cuticle has been produced (**Figure 1.2**).

Lying beneath the body wall structure is a region of somatic muscles. During the J2 life stage these are key for motility enabling the nematode to move toward its host plant. Furthermore, the somatic muscles support maintenance of internal body pressure which is key for the nematode hydrostatic skeleton. Once the nematode becomes sedentary (as a female) the somatic muscles deteriorate in the body wall, remaining in the head to allow movement at the feeding site.

Overlaying the somatic muscles is the basal lamina. Fibrils anchor between the somatic muscle and hypodermis through the basal lamina to allow muscle attachment to the cuticle. In areas lacking somatic muscle the basal lamina encloses the nematode pseudocoelom (body cavity).

The hypodermis, also called the epidermis, lies beneath the layers of cuticle. Although in some species of nematode the hypodermis is cellular, in PCN it is syncytial. The hypodermis is responsible for secreting a new cuticle at every moult. Additionally, the

hypodermis is associated with cuticular ducts such as the secretory-excretory duct, which could allow for secretion of surface coat associated proteins onto the epicuticle. In juveniles, the hypodermis is thin, expanding in regions around the nuclei. However, at later moults the hypodermis swells, presumably a reflection of the increased activity required of this tissue for production of the cuticles synthesised at each moult. The hypodermis stretches from the anterior end where it is responsible for creation of the cephalic framework to the posterior end where it stops short of the tail creating the hyaline section.

The cuticle itself is made up of three sections, the first of these is the basal zone which is connected to the hypodermis. The basal zone primarily consists of collagen giving it a striated appearance. In PCN the basal zone remains striated in males, but in females it deteriorates, losing striations. This may be indicative of the need for locomotion at these life stages (Jones *et al.*, 1993).

Above the basal zone is the median zone, a zone which varies the most across nematode species. In PPN the zone is semi-fluid or fluid rather than being made of plates. The median zone of PCN is supported by small columns surrounded by fluid. The material within this region, although unknown, appears to respond to host root diffusates (Jones *et al.,* 1993).

The cortical zone lies above the median zone and below the epicuticle. This zone is a tough outer area that provides rigidity to the cuticle. The main component of the cortex is non-collagenous cuticulin. In *C. elegans*, collagens have also been shown to localise to the cortical zone. Cuticles of PCN cysts differ from the regular three-layer (basal, median, cortical) structure. Once eggs have formed within the female, the cyst wall begins to form, strengthening the cuticle with two additional collagenous layers. These layers help contain internal pressure in the female caused by expansion of the eggs as the juveniles develop.

Surrounding the cuticle is the epicuticle. Similarly to the cortical zone, the epicuticle is made of non-collagenous, cross-linked proteins.

A surface coat covers the epicuticle of the nematode. This covering is the only portion of the nematode that is directly exposed to the local environment. The surface coat is comprised of carbohydrates, lipids and proteins (Spiegel & McClure, 1995). Although it is still unclear where the surface coat originates from, there is evidence that gene expression for surface coat proteins can be localised to the hypodermis (Jones *et al.*, 2004).

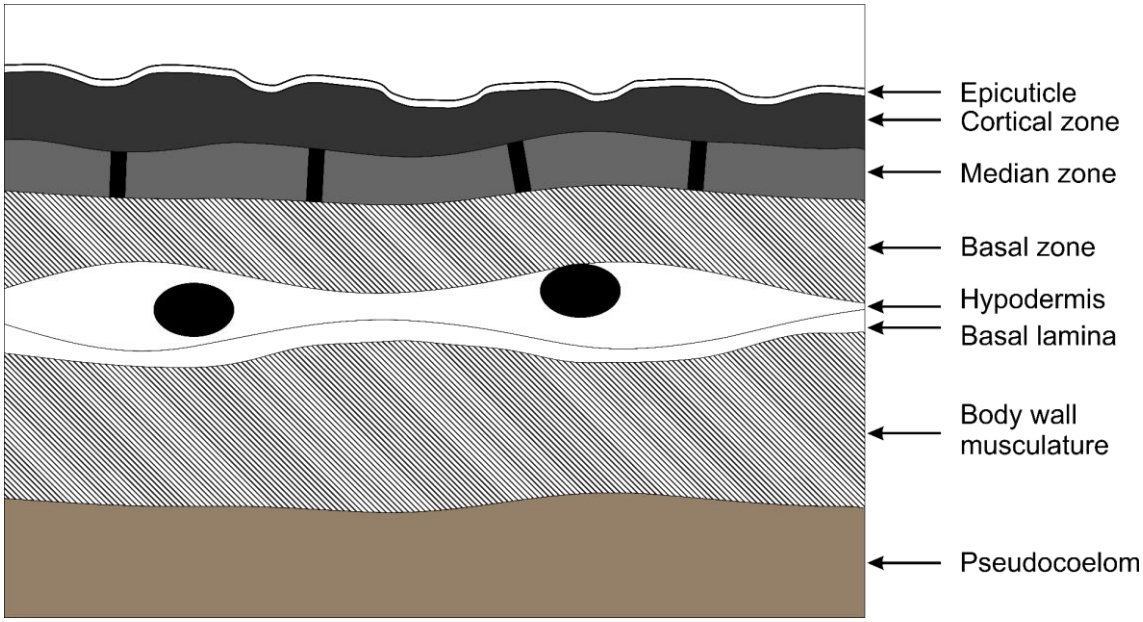


Figure 1.2 – Schematic of generic nematode body wall
The nematode cuticle is formed of cortical, median and basal zones beneath the epicuticle. The syncytial hypodermis is thought to be the origin of proteins destined for the surface coat. The basal lamina separates the hypodermis from the muscle sarcomere. Not to scale.

The surface of PCN responds to host cues. When exposed to root diffusate collected from host plants, the surface coat of cyst nematodes changes to allow higher insertion of lipid probes prompted by changes to the surface lipophilicity (Akhkha *et al.*, 2002). *G. rostochiensis* exhibited surface coat changes when exposed to root exudate from potato and tomato plants. However, *Meloidogyne incognita* (an extremely polyphagous root-knot nematode) has been shown to allow surface lipophilicity changes in response to common phytohormones Indole-3 acetic acid (IAA) and kinetin, which had no effect on *G. rostochiensis* (Akhkha *et al.*, 2002; Akhkha *et al.*, 2004). It is possible that these changes are related to differences in host range of these nematodes.

Although the receptors that detect host cues remain unknown, signalling pathways activated by host interaction have been identified. Akhkha *et al.* (2004) demonstrated that in response to the host, the phospholipase C (PLC) and Phosphatidylinsitol-3 kinase (PI3-kinase) pathways are stimulated. Consequently, this synthesises inositol trisphosphate (IP3) and phosphatidylinositol (3,4,5) trisphosphate (PIP3), which are thought to regulate the surface coat through sequential activation of protein kinases.

Some known PCN surface proteins have been identified and localised. Although these will be discussed in a later chapter, briefly these proteins are a glutathione peroxidase, a fatty acid and retinol binding protein and a peroxiredoxin (Robertson *et al.*, 2000; Prior *et al.*, 2001; Jones *et al.*, 2004). These proteins act to detoxify reactive oxygen species released by the host upon detecting pathogen infection.

Genomic and transcriptomic information has been produced for PCN (Cotton *et al.*, 2014; Eves-van den Akker *et al.*, 2016). These data showed that surface associated genes are continuously expressed throughout the life cycle of the nematode allowing the parasite to respond continuously to changes in the hosts' internal environment.

After infection of the host, cuticular secretions can be detected surrounding the nurse cell and adult females, potentially shielding the nematodes from host detection (Lopez de Mendoza *et al.*, 2002). Previous work with antibodies suggested that the surface coat of PPN mimicked components of host tissues after host infection, camouflaging the pathogen (Bellafiore *et al.*, 2008). Information on the nature of the PCN surface coat composition once in the host is still very limited. After infection, modifications to the cuticle of the nematode are hard to distinguish as differences could be due to the moulting process. Similarly, surface associated structures could be of plant or nematode origin (Jones *et al.*, 1993).

1.7 Eggshell formation

Eggshells provide a range of functions in nematodes. The eggshell is formed to control sperm binding and avoid polyspermy (Johnston *et al.*, 2010). Once fertilisation has

occurred, the eggshell provides protection from abiotic and biotic stresses to the developing juvenile (Perry, 2002).

Most eggshell formation research has been carried out in *C. elegans* and due to the similarities in eggshell structures it would be understandable if most nematode eggshells were formed in similar ways. Eggs are formed of 1-5 layers depending on the species. In *C. elegans,* these layers are built in a hierarchical order, with the outer layer produced first allowing retention of inner layers upon formation (Olson *et al.*, 2012) **Figure 1.3**. As the oocyte passes through the oviduct it begins forming the outer vitelline membrane layer (Foor, 1967), which is formed from the oocytes' oolemma. The vitelline membrane prevents polyspermy upon exposure to sperm in the spermatheca (Stein & Golden, 2015).

After fertilisation, a calcium spike is detectable in the oocyte at the site of sperm entry (Samuel *et al.*, 2001). As the egg passes through the spermatheca to the uterus chitin synthesis and deposition begins (Maruyama *et al.*, 2007). It is believed that chitin synthase is induced by the sperm induced calcium wave (Stein & Golden, 2015). The chitin layer provides the eggshell with structural tensile strength, collagenous proteins within this layer give further rigidity to the eggshell (Perry, 2002; Zhang *et al.*, 2005). No non-chitinous nematode eggshells have yet been identified.

Inside the chitinous layer are the lipid bilayers that provide a permeability barrier for the developing juvenile. Although the number of lipid layers here can vary between nematode species, their function and origin are likely the same across species. This inner layer allows osmotic regulation, as even without the other eggshell layers, the juvenile can develop to completion if this permeability barrier is maintained (Schierenberg & Junkersdorf, 1992).

1.7.1 PCN eggshell composition

To further understand PCN hatching, studies have attempted to identify structural components of the eggshell. Initially the chemical composition of the eggshell of *G. rostochiensis* was investigated. Although relatively few specific compounds were

named, percentages were given describing the proportions of proteins, lipids and other molecules in the egg (Clarke *et al.*, 1967). Lipoprotein membranes were discovered on the inner layer of eggshells of other Tylenchida nematodes (Bird & McClure, 1976). Perry *et al.*, (1982) investigated similar structures in *G. rostochiensis* concluding that the eggshell consisted of a vitelline layer surrounding a chitinous layer that encompassed an inner lipid layer. No surface specialisations were seen. Similar features have been seen in all studied eggshells of PPN. **Figure 1.3** represents the structure of a tylenchid eggshell. No structural schematic has been specifically produced for *G. pallida* or *G. rostochiensis* eggs. Proteins and lipids previously identified from nematode eggshells will be covered in later chapters.

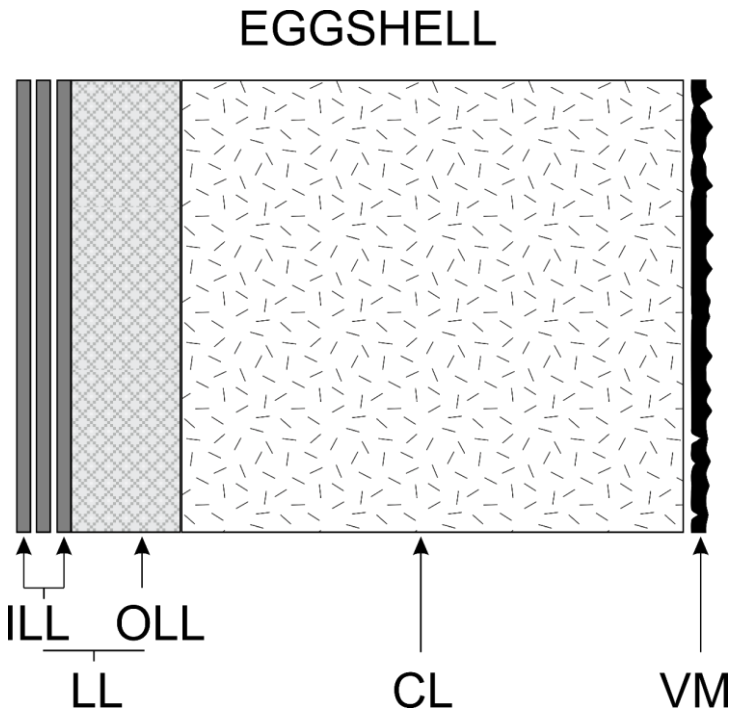


Figure 1.3 – Schematic of a general tylenchid eggshell

Graphic of tylenchid eggshell structure. The eggshell is surrounded by a vitelline membrane (VM). Inside this is the chitinous layer (CL), comprised of microfibrils held in a protein coat. This surrounds the lipid layer (LL), again containing inner (ILL) and outer (OLL) segments. The number of lipid layers and chitin thickness varies between species. Diagram modified from Perry & Trett, (1986). Not to scale.

1.8 PCN hatching

Unlike hatching of broad host range nematodes, hatching of PCN and other PPN with restricted host range is triggered by host cues. To synchronise with host cropping seasons and to aid survival over winter periods, cyst nematode juveniles exhibit a diapause within their eggshells. There are two forms of diapause, obligate diapause is initiated endogenously and can be stopped by the juvenile upon receiving stimuli. Facultative diapause is also seen in PCN. This type of diapause is initiated by exogenous factors and terminated by internal factors. In the case of PCN, facultative diapause is affected by both temperature and length of day (Salazar & Rjrrer, 1993).

PCN are generally a temperate species and therefore need to have a strategy to delay development until winter is complete and favourable conditions return. After the moult from J1 to J2, juvenile *G. rostochiensis* therefore enter an immediate obligate diapause as soil temperatures drop. Following this, the unhatched juvenile enters quiescence and waits until detection of a signal/cue from the host. If no such stimuli are detected the nematode remains in quiescence until such a signal/cue is detected. Using diapause and quiescence, PCN can survive in dormant state in the soil for up to 20 years waiting for stimuli from a host (Masler & Perry, 2018).

The stimulus for exit from quiescence and the onset of hatching is found in host root exudates (Forrest & Perry, 1980; Perry & Beane, 1982). Identifying the compounds in root exudates that stimulate hatch has been a topic of research for some decades. However, the exact composition of all hatching factors found within root diffusate is still unknown. Root exudate components that stimulate hatch are present in trace amounts (Devine & Jones, 2000a). In addition, root exudates contain hatching stimulants and hatching inhibitors. Early developing plant roots produce inhibitors (Byrne *et al.*, 1998). Timing J2 hatching to occur when root systems are further developed increases the chances of initiating a successful feeding site and therefore increases the female egg producing population.

Several plant root exudate components have now been identified that stimulate hatch. Solanoeclepin A, a *G. pallida* and *G. rostochiensis* hatching stimulant, was first identified

and isolated by Mulder *et al.* (1996). Functional synthesis of this stimulant has now been achieved (Tanino *et al.*, 2011), yet it is still unclear how hatching is activated by this factor. Exploiting naturally occurring hatching stimulants (and synthetic alternatives) has been tested for use in nematode control to bring about a 'suicide hatch' (Clarke & Shepherd, 1966; Devine & Jones, 2000b). Although they do cause hatch and therefore could reduce PCN populations, mass synthesis of such specific chemicals is currently too expensive to allow field scale use (Perry, 2002). Although full synthesis techniques for solanoeclepin A have been published and the chemical is claimed to be a major hatching stimulant, little additional work has been carried out using synthesised solanoeclepin A since the original publication. Other hatching factors for PCN include α-chaconine and α-solanine both of which are glycoalkaloids. Hatching in response to these glycoalkaloids appeared faster in *G. rostochiensis* (Byrne *et al.*, 2001). Interestingly, this could reflect the preference of *G. rostochiensis* for species of potatoes rich in glycoalkaloids found in the Andes (Masler & Perry, 2018). However, the heightened presence of glycoalkaloids has also been associated with resistance in species such as *S. canasense*, possibly reflecting the toxicity of glycoalkaloids to higher animals (Castelli *et al.*, 2006).

Analysing and classifying hatching factors within root diffusates is challenging. The chemical complexity of compounds that induce hatching makes their identification problematic. However, with improvements in small molecule mass spectrometry techniques, identifying these compounds may be more feasible in the near future.

1.8.1 The physiology of the hatching process

The earliest detectable change induced by root diffusates during the hatching process is a change in the permeability of the eggshell. The unhatched nematode has a reduced water content and is contained in the perivitelline fluid within the egg. This perivitelline fluid contains an elevated level of trehalose, which is thought to act as a cryoprotectant and desiccation protectant (Wharton *et al.*, 2000). After exposure to hatching stimulants, trehalose can be detected outside the eggshell (Ellenby & Perry, 1976) indicating that a change in nematode eggshell permeability occurs upon exposure to root exudates.

The change in permeability of the PCN eggshell in response to root exudates is thought to involve Ca^{2+} binding (Atkinson & Taylor, 1980). It is still unclear as to whether Ca^{2+} transport systems or facilitated structural changes due to Ca^{2+} are being exploited. Original work suggested that free Ca^{2+} is essential for hatching. This was disputed by Clarke & Hennessy (1983) who instead proposed that root diffusate affects eggshell bound Ca^{2+}, resulting in permeability changes. This was supported by earlier work that showed one of the major inorganic components of the eggshell is calcium, which was found in 'considerable amounts' (Clarke *et al.*, 1967). There have been three forms of Ca^{2+} binding sites suggested to be present within the eggshell of *G. rostochiensis*, those that bind Ca^{2+} on the outer layers without being involved in hatch, those where root exudate displaces Ca^{2+} already associated with the site and those that can bind additional Ca^{2+} when exposed to root diffusate (Clarke & Perry, 1985). A biological nematicide originating from fermentation with hyphomycete fungi (DiTera®) significantly reduced PCN hatching by irreversibly preventing permeability changes usually caused by root diffusate, and it is possible that this may involve blocking of Ca^{2+} binding sites (Twomey *et al.*, 2000). It should be noted that although a key role for Ca^{2+} binding sites in the eggshell of PCN has been proposed, no specific proteins have been isolated that could be involved in this process.

After a change in permeability of the lipid barrier the quiescent nematode begins to rehydrate as water moves into the egg. After metabolic activation the juvenile begins slight locomotion within the eggshell (Blair *et al.*, 1999). When ready for eclosion, the juvenile starts a series of stylet thrusts through the eggshell to form a slit from which it can hatch (Doncaster & Seymour, 1973).

Despite the work on hatching done to date, little is known about the nature of the lipids in the eggshells of *G. rostochiensis*, or how these change in response to hatching stimuli. Similarly, no proteins present within the eggshell have been identified. However, isolating and classifying proteins within the eggshell structure could allow a greater understanding of the effect root diffusates have on the eggshell. Monitoring changes in

extracted peptide abundance before and after initiation of the hatching cascade also has the potential to allow identification of proteins interacting with hatching factors.

1.9 Conclusions

The economic importance and threat to food security of the potato cyst nematode has made it a key organism for research. Although many control mechanisms exist for limiting the pathogen, none of them give full protection.

To further understand the pathogenic lifestyle of these nematodes it is important to appreciate all aspects of their lifecycle. Previous research on PCN eggshells has focused on compounds that cause or inhibit hatching rather than the hatching mechanism itself. Research is still required to ascertain the role(s) of root diffusate as a hatching stimulus. The exact changes that host diffusate causes in the eggshell including binding sites and progression of permeability and consequently hatching remain unknown.

Research into host-parasite interactions is largely based around effector proteins originating from pharyngeal glands. However, research into proteins released from the nematode's cuticle may offer a greater understanding as to how the nematode can be protected from host defences.

1.9.1 Project aims

This thesis aims to cover some of the gaps identified in our knowledge of structures offering protection to the nematode. Primarily this will be done by characterising the PCN eggshell. Identifying and localising proteins within the eggshell will help develop a better understanding of the PCN hatching mechanisms. Lipid extractions from eggshells could offer further information into how the permeability barrier is formed and how this could relate in response to host cues. Research will also be carried out to identify proteins presented on the surface of the juvenile that defend the nematode during infection of the host.

1.10 References

AKHKHA, A., KUSEL, J., KENNEDY, M. & CURTIS, R. 2002. Effects of phytohormones on the surfaces of plant-parasitic nematodes. *Parasitology*, **125**, 165–175.

AKHKHA, A., CURTIS, R., KENNEDY, M. & KUSEL, J. 2004. The potential signalling pathways which regulate surface changes induced by phytohormones in the potato cyst nematode (Globodera rostochiensis). *Parasitology*, **128**, 533–539.

ALLEN, M.W. & RASKI, D.J. 1950. The effect of soil type on the dispersion of soil fumigants. *Phytopathology*, **40**, 1043–1053 .

ATKINSON, H.J. 2002. Molecular Approaches to Novel Crop Resistance Against Nematodes. *In* Lee, D., *ed. The Biology of Nematodes*. Taylor & Francis, 569–592.

ATKINSON, H.J. & TAYLOR, J.D. 1980. Evidence for a calcium-binding site on the eggshell of Globodera rostochiensis with a role in hatching. *Annals of Applied Biology*, **96**, 307–315.

BALDWIN, J.G. & HANDOO, Z.A. 2018. General morphology of cyst nematodes. *In* Perry, R.N., Moens, M. & Jones, J.T., *eds. Cyst Nematodes*. CABI, 337–364.

BARONE, A., RITTER, E., SCHACHTSCHABEL, U., DEBENER, T., SALAMINI, F. & GEBHARDT, C. 1990. Localization by restriction fragment length polymorphism mapping in potato of a major dominant gene conferring resistance to the potato cyst nematode Globodera rostochiensis. *Molecular & general genetics : MGG*, **224**, 177–182.

BELLAFIORE, S., SHEN, Z., ROSSO, M., ABAD, P., SHIH, P. & BRIGGS, S.P. 2008. Direct identification of the Meloidogyne incognita secretome reveals proteins with host cell reprogramming potential. *PLoS pathogens*, **4**.

BIRD, A.F. & MCCLURE, M.A. 1976. The tylenchid (Nematoda) egg shell: structure, composition and permeability. *Parasitology*, **72**, 19.

BIRD, D.M. & KOLTAI, H. 2000. Plant Parasitic Nematodes: Habitats, Hormones, and Horizontally-Acquired Genes. *Journal of plant growth regulation*, **19**, 183–194.

BLAIR, L., OPARKA, K., JONES, J. & PERRY, R. 1999. Activation of transcription during the hatching process of the potato cyst nematode Globodera rostochiensis. *Nematology*, **1**, 103–111.

BOHLMANN, H. 2015. Basic Biology of Cyst Nematodes. *In* Escobar, C. & Fenoll, C., *eds.*

Advances in Botanical Research: Plant Nematode Interactions - A View on Compatible Interrelationships. 34–59.

BYRNE, J.T., TWOMEY, U., MAHER, N.J., DEVINE, K.J. & JONES, P.W. 1998. Detection of hatching inhibitors and hatching factor stimulants for golden potato cyst nematode, Globodera rostochiensis, in potato root leachate. *Annals of Applied Biology*, **132**, 463–472.

BYRNE, J.T., MAHER, N.J. & JONES, P.W. 2001. Comparative Responses of Globodera rostochiensis and G. pallida to Hatching Chemicals. *Journal of nematology*, **33**, 195–202.

CAROMEL, B., MUGNIÉRY, D., LEFEBVRE, V., ANDRZEJEWSKI, S., ELLISSÈCHE, D., KERLAN, M.C., ROUSSELLE, P. & ROUSSELLE-BOURGEOIS, F. 2003. Mapping QTLs for resistance against Globodera pallida (Stone) Pa2/3 in a diploid potato progeny originating from Solanum spegazzinii. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, **106**, 1517–1523.

CASTELLI, L., BRYAN, G., BLOK, V.C., RAMSAY, G., SOBCZAK, M., GILLESPIE, T. & PHILLIPS, M.S. 2006. Investigations of Globodera pallida invasion and syncytia formation within roots of the susceptible potato cultivar Désirée and resistant species Solanum canasense. *Nematology*, **8**, 103–110.

CHARLIER, J., HÖGLUND, J., VON SAMSON-HIMMELSTJERNA, G., DORNY, P. & VERCRUYSSE, J. 2009. Gastrointestinal nematode infections in adult dairy cattle: impact on production, diagnosis and control. *Veterinary parasitology*, **164**, 70–79.

CLARKE, A.J., COX, P.M. & SHEPHERD, A.M. 1967. The chemical composition of the egg shells of the potato cyst-nematode, Heterodera rostochiensis Woll. *The Biochemical Journal*, **104**, 1056–1060.

CLARKE, A.J. & HENNESSY, J. 1983. The role of calcium in the hatching of Globodera rostochiensis. *Revue de Nématologie*, **6**, 247–255.

CLARKE, A.J. & PERRY, R.N. 1985. Egg-shell calcium and the hatching of Globodera rostochiensis. *International Journal for Parasitology*, **15**, 511–516.

CLARKE, A.J. & SHEPHERD, A.M. 1966. Picrolonic Acid as a Hatching Agent for the Potato Cyst Nematode, Heterodera rostochiensis Woll. Nature. *Nature*, **211**, 546–546.

COTTON, J.A., LILLEY, C.J., JONES, L.M., KIKUCHI, T., REID, A.J., THORPE, P., TSAI, I.J., … URWIN, P.E. 2014. The genome and life-stage specific transcriptomes of Globodera pallida elucidate key aspects of plant parasitism by a cyst nematode. *Genome Biology*, **15**, R43, 10.1186/gb-2014-15-3-r43.

DANGL, J.L. & MCDOWELL, J.M. 2006. Two modes of pathogen recognition by plants. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 8575–8576.

DECRAEMER, W. & HUNT, D.J. 2006. Structure and Classification. *In* Perry, R.N. & Moens, M., *eds. Plant Nematology*. Wallingford: CABI, 26–27.

DEVINE, K.J. & JONES, P.W. 2000a. Purification and partial characterisation of hatching factors for the potato cyst nematode Globodera rostochiensis from potato root leachate. *Nematology*, **2**, 231–236.

DEVINE, K.J. & JONES, P.W. 2000b. Response of Globodera rostochiensis to exogenously applied hatching factors in soil. *Annals of Applied Biology*, **137**, 21–29.

DONCASTER, C.C. & SEYMOUR, M.K. 1973. Exploration and Selection of Penetration Site By Tylenchida. *Nematologica*, **19**, 137–145.

ELLENBY, C. & PERRY, R.N. 1976. The Influence of the Hatching Factor on the Water Uptake of the Second Stage Larva of the Potato Cyst Nematode Heterodera Rostochiensis. *J. Exp. Biol.*, **64**, 141–147.

EVANS, K., FRANCO, J. & DE SCURRAH, M.M. 1975. Distribution of Species of Potato Cyst-Nematodes in South America. *Nematologica*, **21**, 365–369.

EVES-VAN DEN AKKER, S., LAETSCH, D.R., THORPE, P., LILLEY, C.J., DANCHIN, E.G.J., DA ROCHA, M., RANCUREL, C., … JONES, J.T. 2016. The genome of the yellow potato cyst nematode, Globodera rostochiensis, reveals insights into the basis of parasitism and virulence. *Genome Biology*, **17**, 124.

F.A.O. 2008. World Potato Production. *Potato world: Production and Consumption - International Year of the Potato 2008*.

FEECHAN, A., TURNBULL, D., STEVENS, L.J., ENGELHARDT, S., BIRCH, P.R.J., HEIN, I. & GILROY, E.M. 2015. The Hypersensitive Response in PAMP- and Effector-Triggered Immune Responses. *In Plant Programmed Cell Death*. Cham: Springer International

Publishing, 235–268.

FOOR, W.E. 1967. Ultrastructural aspects of oocyte development and shell formation in Ascaris lumbricoides. *The Journal of Parasitology*, **53**, 1245–1261.

FORREST, J.M.S. & PERRY, R.N. 1980. Hatching of Globodera pallida eggs after brief exposures to potato root diffusate. *Nematologica*, **26**, 130–132.

FRANKLIN, M.T. 2012. Taxonomy of Heteroderidae. *In* Zuckerman, B.M., Mai, W.F. & Rohde, R.A., *eds. Plant Parasitic Nematodes - Volume 1, Morphology, Anatomy, Taxonomy, and Ecology*. Elsevier Ltd, 152–153.

GIMENEZ-IBANEZ, S. & RATHJEN, J.P. 2010. The case for the defense: plants versus Pseudomonas syringae. *Microbes and Infection*, **12**, 428–437.

HAMMOND-KOSACK, K.E. & JONES, J.D.G. 1997. Plant Diesease Resistance Genes. *Annual Review of Plant Physiology and Plant Molecular Biology*, **48**, 575–607.

HAYDOCK, P.P.J., WOODS, S.R., GROVE, I.G. & HARE, M.C. 2006. Chemical Control of Nematodes. *In* Perry, R.N. & Moens, M., *eds. Plant Nematology*. Wallingford: CABI, 392–410.

JOHNSTON, W.L., KRIZUS, A. & DENNIS, J.W. 2010. Eggshell Chitin and Chitin-Interacting Proteins Prevent Polyspermy in C. elegans. *Current Biology*, **20**, 1932–1937.

JONES, J.D.G. & DANGL, J.L. 2006. The plant immune system. *Nature*, **444**, 323–329.

JONES, J.T., PERRY, R.N. & JOHNSTON, M.R.L. 1993. Changes in the ultrastructure of the cuticle of the potato cyst nematode, Globodera rostochiensis, during development and infection. *Fundamental and Applied Nematology*, **16**, 433–445.

JONES, J.T., REAVY, B., SMANT, G. & PRIOR, A.E. 2004. Glutathione peroxidases of the potato cyst nematode Globodera Rostochiensis. *Gene*, **324**, 47–54.

KALETTA, T. & HENGARTNER, M.O. 2006. Finding function in novel targets: C. elegans as a model organism. *Nature reviews. Drug discovery*, **5**, 387–398.

KERRY, B.R. & HOMINICK, W.M. 2002. Biological Control. *In* Lee, D., *ed. The Biology of Nematodes*. Taylor & Francis, 483–504.

KIKUCHI, T., EVES-VAN DEN AKKER, S. & JONES, J.T. 2017. Genome Evolution of Plant-Parasitic Nematodes. *Annual Review of Phytopathology*, **55**, 333–354.

KOENNING, S.R. & SIPES, B.S. 2013. Biology. *In* Sharma, S.B., *ed. The Cyst Nematodes*.

Springer Science & Business Media, 159–161.

LAMBERT, K. & BEKAL, S. 2002. Introduction to Plant-Parasitic Nematodes. *The Plant Health Instructor*.

LEE, D.L. 2002a. Cuticle, Moulting and Exsheathment. *In* Lee, D.L., *ed. The Biology of Nematodes*. Taylor & Francis, 171–209.

LEE, D.L. 2002b. Life Cycles. *In* Lee, D.L., *ed. The Biology of Nematodes*. Taylor & Francis, 62–64.

LINTS, R. & HALL, D.H. 2009. The Cuticle. *In* Herndon, L.A., *ed. WormAtlas*. Accessed from: https://www.wormatlas.org/hermaphrodite/cuticle/mainframe.htm

LOPEZ DE MENDOZA, M.E., ABRANTES, I.M.O., ROWE, J.A., GOWEN, S. & CURTIS, R. 2002. Immunolocalisation in planta of secretions from parasitic stages of Meloidogyne incognita and M. hispanica. *International Journal of Nematology*, **12**, 149–154.

LOZANO-TORRES, J.L., WILBERS, R.H.P., WARMERDAM, S., FINKERS-TOMCZAK, A., DIAZ-GRANADOS, A., VAN SCHAIK, C.C., HELDER, J., … SMANT, G. 2014. Apoplastic Venom Allergen-like Proteins of Cyst Nematodes Modulate the Activation of Basal Plant Innate Immunity by Cell Surface Receptors. Kaloshian, I., *ed. PLoS Pathogens*, **10**.

MALINOVSKY, F.G., FANGEL, J.U. & WILLATS, W.G.T. 2014. The role of the cell wall in plant immunity. *Frontiers in Plant Science*, **5**, 178.

MANOSALVA, P., MANOHAR, M., VON REUSS, S.H., CHEN, S., KOCH, A., KAPLAN, F., CHOE, A., … KLESSIG, D.F. 2015. Conserved nematode signalling molecules elicit plant defenses and pathogen resistance. *Nature Communications*, **6**, 7795.

MARUYAMA, R., VELARDE, N. V., KLANCER, R., GORDON, S., KADANDALE, P., PARRY, J.M., HANG, J.S., … SINGSON, A. 2007. EGG-3 Regulates Cell-Surface and Cortex Rearrangements during Egg Activation in Caenorhabditis elegans. *Current Biology*, **17**, 1555–1560.

MASLER, E.P. & PERRY, R.N. 2018. Hatch, survival and sensory perception. *In* Perry, R.N., Moens, M. & Jones, J.T., *eds. Cyst Nematodes*. CABI, 44–73.

MCSORLEY, R. 2003. Adaptations of nematodes to environmental extremes. *Florida Entomologist*, **86**, 138–142.

MINNIS, S.T., HAYDOCK, P.P.J., IBRAHIM, S.K., GROVE, I.G., EVANS, K. & RUSSELL, M.D. 2002. Potato cyst nematodes in England and Wales - occurrence and distribution. *Annals*

of Applied Biology, **140**, 187–195.

MOENS, M. & PERRY, R.N. 2009. Migratory plant endoparasitic nematodes: a group rich in contrasts and divergence. *Annual review of phytopathology*, **47**, 313–332.

MULDER, J.G., DIEPENHORST, P., PLIEGER, P. & BRUGGEMANN-ROTGANS, I.E.M. 1996. Hatching agent for the potato cyst nematode. *Official gazette of the United States Patent and Trademark Office. Patents*, **1193**, 2131.

NEEDHAM, T. 1743. A letter concerning certain chalky tubulous concretions, called malm; with some microscopical observations on the farina of the red lilly and worms discovered in smutty corn. *Philosophical Transactions of the Royal Society, London*, **42**, 634–641.

NICOL, J.M., TURNER, S.J., COYNE, D.L., NIJS, L. DEN, HOCKLAND, S. & TAHNA MAAFI, Z. 2011. Current Nematode Threats to World Agriculture. *In* Jones, J., Gheysen, G. & Fenoll, C., *eds. Genomics and Molecular Genetics of Plant-Nematode Interactions*. Springer Netherlands, 347–367.

OLSON, S.K., GREENAN, G., DESAI, A., MÜLLER-REICHERT, T. & OEGEMA, K. 2012. Hierarchical assembly of the eggshell and permeability barrier in C. Elegans. *Journal of Cell Biology*, **198**, 731–748.

PERRY, R.N. 2002a. Hatching. *In* Lee, D.L., *ed. The Biology of Nematodes*. Taylor & Francis, 147–169.

PERRY, R.N. & BEANE, J. 1982. The effects of brief exposures to potato root diffusate on the hatching of Globodera rostochiensisec. *Revue de Nématologie*, **5**, 221–224.

PERRY, R.N. & TRETT, M.W. 1986. Ultrastructure of the eggshell of Heterodera schachtii and H. glycines (Nematoda: Tylenchida). *Revue de nematologie*.

PERRY, R.N., WHARTON, D.A. & CLARKE, A.J. 1982. The structure of the egg-shell of Globodera rostochiensis (Nematoda: Tylenchida). *International Journal for Parasitology*, **12**, 481–485.

PICKUP, J., ROBERTS, A.M.I. & NIJS, L. DEN. 2018. Quarantine, Distribution Patterns and Sampling. *In* Perry, R.N., Moens, M. & Jones, J.T., *eds. Cyst Nematodes*. CABI, 128–153.

POWERS, T. 1992. Biological control of plant parasitic nematodes: Progress, problems and

prospects. *Parasitology Today*, **8**, 320.

PREE, D.J., TOWNSHEND, J.L. & ARCHIBALD, D.E. 1989. Carbamate and Organophosphorus Nematicides: Acetylcholinesterase inhibition and Effects on Dispersal. *Journal of nematology*, **21**, 483–489.

PRIOR, A., JONES, J.T., BLOK, V.C., BEAUCHAMP, J., MCDERMOTT, L., COOPER, A. & KENNEDY, M.W. 2001. A surface-associated retinol- and fatty acid-binding protein (Gp-FAR-1) from the potato cyst nematode Globodera pallida: lipid binding activities, structural analysis and expression pattern. *The Biochemical journal*, **356**, 387–394.

PRITCHARD, L. & BIRCH, P.R.J. 2014. The zigzag model of plant-microbe interactions: is it time to move on? *Molecular Plant Pathology*, **15**, 865–870.

RICE, S.L., LEADBEATER, B.S.C. & STONE, A.R. 1985. Changes in cell structure in roots of resistant potatoes parasitized by potato cyst-nematodes. I. Potatoes with resistance gene H1 derived from Solanum tuberosum ssp. andigena. *Physiological Plant Pathology*, **27**, 219–234.

ROBERTSON, L., ROBERTSON, W.M., SOBCZAK, M., HELDER, J., TETAUD, E., ARIYANAYAGAM, M.R., FERGUSON, M.A., FAIRLAMB, A. & JONES, J.T. 2000. Cloning, expression and functional characterisation of a peroxiredoxin from the potato cyst nematode Globodera rostochiensis. *Molecular and Biochemical Parasitology*, **111**, 41–49.

ROBINSON, A.F. & PERRY, R.N. 2006. Behaviour and Sensory Perception. *In* Perry, R.N. & Moens, M., *eds. Plant Nematology*. Wallingford: CABI, 211–214.

ROUPPE VAN DER VOORT, J., WOLTERS, P., FOLKERTSMA, R., HUTTEN, R., VAN ZANDVOORT, P., VINKE, H., KANYUKA, K., … BAKKER, J. 1997. Mapping of the cyst nematode resistance locus Gpa2 in potato using a strategy based on comigrating AFLP markers. *TAG Theoretical and Applied Genetics*, **95**, 874–880.

S.A.S.A. 2010. Control of Potato Cyst Nematodes - Official Control Programme for Sampled Units Recorded as Infested - Guidance for Growers and Landowners. *Control of Potato Cyst Nematodes - Official Control Programme for Sampled Units Recorded as Infested - Guidance for Growers and Landowners*. Edinburgh.

SACCO, M.A., KOROPACKA, K., GRENIER, E., JAUBERT, M.J., BLANCHARD, A., GOVERSE, A., SMANT, G. & MOFFETT, P. 2009. The Cyst Nematode SPRYSEC Protein RBP-1 Elicits Gpa2- and

RanGAP2-Dependent Plant Cell Death. Opperman, C., *ed. PLoS Pathogens*, **5**, e1000564.

SALAZAR, A. & RJRRER, E. 1993. Effects of daylength during cyst formation, storage time and temperature of cysts on the in vitro hatching of Globodera rostochiensis and G. pallida. *Fundam. appl. Nematol*, **16**, 567–572.

SAMUEL, A.D., MURTHY, V.N. & HENGARTNER, M.O. 2001. Calcium dynamics during fertilization in C. elegans. *BMC developmental biology*, **1**, 8.

SCHIERENBERG, E. & JUNKERSDORF, B. 1992. The role of eggshell and underlying vitelline membrane for normal pattern formation in the early C. elegans embryo. *Roux's Archives of Developmental Biology*, **202**, 10–16.

SIDONSKAYA, E., SCHWEIGHOFER, A., SHUBCHYNSKYY, V., KAMMERHOFER, N., HOFMANN, J., WIECZOREK, K. & MESKIENE, I. 2016. Plant resistance against the parasitic nematode Heterodera schachtii is mediated by MPK3 and MPK6 kinases, which are controlled by the MAPK phosphatase AP2C1 in Arabidopsis. *Journal of Experimental Botany*, **67**, 107–118.

SIJMONS, P.C., ATKINSON, H.J. & WYSS, U. 1994. Parasitic Strategies of Root Nematodes and Associated Host Cell Responses. *Annual Review of Phytopathology*, **32**, 235–259.

SOBCZAK, M. & GOLINOWSKI, W. 2009. Cell Biology of Plant Nematode Parasitism. *Cell Biology of Plant Nematode Parasitism*. Berg, R.H. & Taylor, C.G., *eds.* Berlin, Heidelberg: Springer Berlin Heidelberg.

SPIEGEL, Y. & MCCLURE, M.A. 1995. The surface coat of plant-parasitic nematodes: chemical composition, origin, and biological role-a review. *Journal of nematology*, **27**, 127–134.

SPOEL, S.H. & DONG, X. 2012. How do plants achieve immunity? Defence without specialized immune cells. *Nature Reviews Immunology*, **12**, 89–100.

STEIN, K.K. & GOLDEN, A. 2015. The C. elegans eggshell. *In WormBook: the online review of C. elegans biology*. 1–36. Accessed from: http://www.wormbook.org/chapters/www_eggshell/eggshell.html

STEPEK, G., BUTTLE, D.J., DUCE, I.R. & BEHNKE, J.M. 2006. Human gastrointestinal nematode infections: are new control methods required? *International journal of*

experimental pathology, **87**, 325–341.

TANINO, K., TAKAHASHI, M., TOMATA, Y., TOKURA, H., UEHARA, T., NARABU, T. & MIYASHITA, M. 2011. Total synthesis of solanoeclepin A. *Nature Chemistry*, **3**, 484–488.

TRUDGILL, D.L. 1967. The Effect of Environment On Sex Determination in Heterodera Rostochiensis. *Nematologica*, **13**, 263–272.

TRUDGILL, D.L. 1986. Yield losses caused by potato cyst nematodes: a review of the current position in Britain and prospects for improvements. *Annals of Applied Biology*, **108**, 181–198.

TRUDGILL, D.L., PHILLIPS, M.S. & ELLIOTT, M.J. 2014. Dynamics and management of the white potato cyst nematode Globodera pallida in commercial potato crops. *Annals of Applied Biology*, **164**, 18–34.

TURNER, S.J. & ROWE, J.A. 2006. Cyst Nematodes. *In* Perry, R. & Moens, M., *eds. Plant Nematology*. Wallingford: CABI, 122.

TWOMEY, U., WARRIOR, P., KERRY, B.R. & PERRY, R.N. 2000. Effects of the biological nematicide, DiTera , on hatching of Globodera rostochiensis and G. pallida. *Nematology*, **2**, 355–362.

VAN DER VOSSEN, E.A.G., JEROEN, N.A.M., ROUPPE VAN DER VOORT, J., KANYUKA, K., BENDAHMANE, A., SANDBRINK, H., BAULCOMBE, D.C., BAKKER, J., STIEKEMA, W.J. & KLEIN-LANKHORST, R.M. 2000. Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: a virus and a nematode. *The Plant Journal*, **23**, 567–576.

VIAENE, N., COYNE, D.L. & KERRY, B.R. 2006. Biological and Cultural Management. *In* Perry, R. & Moens, M., *eds. Plant Nematology*. Wallingford: CABI, 347–349.

WHARTON, D.A., JUDGE, K.F. & WORLAND, M.R. 2000. Cold acclimation and cryoprotectants in a freeze-tolerant Antarctic nematode, Panagrolaimus davidi. *Journal of comparative physiology. B, Biochemical, systemic, and environmental physiology*, **170**, 321–327.

WHITEHEAD, A.G. 1985. The potential value of British wild Solanum spp. as trap crops for potato cyst nematodes, Globodera rostochiensis and G. pallida. *Plant Pathology*, **34**, 105–107.

WYSS, U. 2002. Feeding Behavior of Plant-Parasitic Nematodes. *In* Lee, D.E., *ed. The*

Biology of Nematodes. Taylor & Francis, 233–260.

WYSS, U. 1997. Root Parasitic Nematodes: An Overview. *In* Fenoll, C., Grundler, F. & Ohl, S., *eds. Cellular and Molecular Aspects of Plant-Nematode Interactions*. Springer Science & Business Media, 5–25.

ZHANG, Y., FOSTER, J.M., NELSON, L.S., MA, D. & CARLOW, C.K.S. 2005. The chitin synthase genes chs-1 and chs-2 are essential for C. elegans development and responsible for chitin deposition in the eggshell and pharynx, respectively. *Developmental Biology*, **285**, 330–339.

ZHANG, Z.-Q. 2013. Animal biodiversity: An update of classification and diversity in 2013. In: Zhang, Z.-Q. (Ed.) Animal Biodiversity: An Outline of Higher-level. *Zootaxa*, **3703**, 5.

ZINOVIEVA, S. V. 2014. Co-adaptation mechanisms in plant-nematode systems. *Parazitologiia*, **48**, 110–130.

ZIPFEL, C. & ROBATZEK, S. 2010. Pathogen-associated molecular pattern-triggered immunity: veni, vidi...? *Plant Physiology*, **154**, 551–554.

ZUNKE, U. 1990. Observations on the Invasion and Endoparasitic Behavior of the Root Lesion Nematode Pratylenchus penetrans. *Journal of Nematology*, **22**, 309–320.

2. General Materials & Methods

2.1 Biological Materials

2.1.1 Tomato/potato root diffusate collection

Root diffusate (RD) was collected from roots of 4 week old plants ('Moneymaker' for TRD, 'Desiree' for PRD). Soil was washed from the roots with water before the plant was placed in 500 mL dH_2O for 24 h. This water, now containing root diffusates, was filtered through Whatman filter paper before being stored at 4°C. Root diffusate was used within 4 weeks of being collected.

2.1.2 Collection of Juveniles

G. rostochiensis cysts from a 2012 population were placed on a 200 µm sieve, which was seated inside a large Petri dish. An abundance of root diffusate was poured over the cysts and any air bubbles between the sieve and the Petri dish were removed. The apparatus was wrapped in cling film and tin foil then placed in an incubator set to 18 ˚C. Large numbers of juveniles could usually be collected from the Petri dish after 9-12 days. As the juveniles hatched from the cysts they were able to migrate through the sieve and collect in the Petri dish. The root diffusate and nematodes were removed and replaced with fresh root diffusate at regular intervals. Fungal contaminants were often found in the Petri dish in which juvenile nematodes were being hatched. These contaminants were removed by sucrose floatation. An equal amount of nematode solution and 50% sucrose stock (50 g sucrose, 50 mL dH_2O) were combined and overlaid with 500 µL of sterile distilled water (SDW), this was then centrifuged at 5200 x g. for 10 minutes. Fungal contaminants collected in the base of the tube and juveniles were removed from the interphase between the sucrose and water. Juveniles were washed twice with SDW to remove any remaining sucrose. Juveniles were used within 2 days of collection.

2.1.3 Collection of parasitic nematode life stages

Chitting potatoes ('Desiree') were warmed in the glass house for four days before being planted in a 50:50 mix of sterile sand:loam. Approximately 50 *G. rostochiensis* cysts per plant were incorporated into the soil. Plants were grown in controlled greenhouse conditions with 16-hour days and 8-hour nights and were watered lightly from above daily.

After three weeks the infected roots were removed from the plant and rinsed free of any soil. The root systems were cut into roughly one-inch sections before being blended in a food processor with a small amount of sterile water. Blended roots were passed through a series of sieves decreasing in gap diameter, 2.98 mm, 500 µm, 250 µm, 90 µm, 25 µm. Nematodes were collected on the smallest sieve and purified using sucrose floatation as described in section **2.1.2**. The top 5 mL from the sucrose float was removed and placed onto a 38 µm sieve. This allowed nematodes to be collected and allowed the sucrose solution to be washed off. Nematodes were then rinsed into a protein LoBind Eppendorf tube in sterile water and stored at 4 °C.

2.1.4 Collection of Eggshells

G. rostochiensis cysts from a 2012 population were placed in 1 mL of dH_2O before being transferred to a general-purpose Potter-Elvehjem homogeniser (GPE Scientific). After crushing the cysts with the pestle, the nematode solution was transferred to a 15 ml Falcon tube. To remove fragments of cysts the nematode solution was poured over two mesh sieves with gap widths of 50 µm and 38 µm. Cyst fragments were trapped on the 50 µm sieve allowing juveniles and eggshells to pass through. The 38 µm sieve trapped the eggshells and juveniles. Eggshells and juveniles were washed from the sieve into a 15 mL Falcon tube where they settled to the bottom over the period of an hour. The supernatant was removed and the juvenile and egg pellet was resuspended in 10 mL of sterile water. The eggs were then evenly distributed between 10 1.5 mL Eppendorf microcentrifuge tubes.

An ultrasonic water bath was used to break open the eggshells (Grant ultrasonic bath MXB22). The bath was cooled to 4 °C to avoid overheating. Eppendorf tubes containing nematode/egg solutions were floated in the bath. Sonication was run for 30 minutes, and eggshells were observed under the microscope after each 30 min run. If enough eggs had been burst open the eggshells were purified by density centrifugation in sodium-potassium tartrate tetrahydrate (Sigma) solutions at varying concentrations (Table 2.1). Most eggs burst within 90 minutes of sonication. This method was partially modified from the method used by Clarke *et al.* (1967), increasing the centrifugation time from 1 to 2 minutes and increasing the sonication time from 10-15 minutes depending on the number of visualised burst eggshells. Figure 2.1 shows separation of the juveniles, juveniles in eggshells and sheared eggshells by the sodium-potassium tartrate gradient.

Table 2.1 – Sodium-potassium tartrate gradient for eggshell purification

Sodium-potassium tartrate concentration	Volume used per purification
1 g/mL	1.5 mL
0.75 g/mL	1.0 mL
0.625 g/mL	1.0 mL
0.5 g/mL	1.0 mL
0.375 g/mL	1.0 mL
0.25 g/mL	1.0 mL

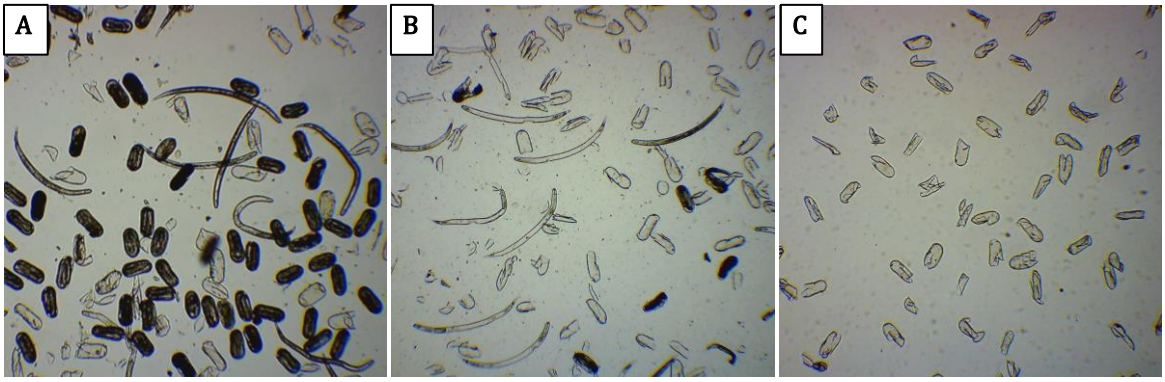


Figure 2.1 – Eggshell purification.
A – Eggshell/nematode sample after cysts are crushed with a tissue homogeniser. B – Eggshell lysis after a period in the ultrasonic water bath. C – purified eggshells after sodium-potassium tartrate gradient density centrifugation. Images taken on a moticam microscope camera attached by a 0.5x magnification mount seated above a 40x objective lens.

2.3 Molecular Biology

2.3.1 Nematode RNA Extraction

Nematodes were centrifuged at 5200 x g to form a pellet. The supernatant was removed and the microcentrifuge tube containing the nematode pellet was placed in liquid nitrogen. The pellet was transferred to a sterile, chilled mortar and crushed to a fine powder. In the mortar, 1 mL of TRIzol reagent (Ambion Life Sciences, Thermo-Fisher Scientific) and 0.2 mL of chloroform was added to the nematode powder. After the sample had defrosted, the resulting mixture was transferred to a microcentrifuge tube and vortexed before being incubated at room temperature for 5 minutes with regular shaking. The sample was pelleted by centrifuging at 12,000 x g for 15 minutes at 4 ˚C. The upper aqueous phase was removed and transferred to a fresh tube containing 0.5 mL of 100% isopropanol. After incubating at room temperature for 10 minutes the sample was centrifuged at 12,000 x g for 10 minutes (4 ˚C). The supernatant was removed and replaced with 1 mL of 75% ethanol to wash the pellet. The sample was vortexed then centrifuged at 7500 x g for 5 minutes (4 ˚C). The supernatant was removed and discarded and the pellet was left to air dry. RNase-free water was added to resuspend the pellet in a 20 µL volume before incubation for 10 minutes at 60 ˚C to allow dissolution of the RNA.

2.3.2 DNase Treatment of RNA Extraction

DNase digestion was used to remove any DNA in the RNA sample. This was achieved using an RQ1 RNase-free DNase kit from Promega. The instructions were followed as per the manufacturer's specifications. Briefly, 8 µL of RNA was added to 1 µL RQ1 10 x reaction buffer and 1 µL RQ1 RNA-free DNase and incubated at 37 ˚C for 30 minutes. 1 µL RQ1 DNase stop solution was added and the mixture was incubated for a further 10 minutes at 65 ˚C to terminate the digestion. DNA free RNA was used to create cDNA immediately.

2.3.3 cDNA Preparation from DNase Treated RNA

DNase treated RNA was added to 1 µL of 50 µM oligo$(dT)_{20}$ (Eurofins) and 1 µL 10 mM dNTP mix. The solution was heated to 65 ˚C for 5 minutes before being immediately transferred to ice to snap chill. Next, 4 µL of 5 x first strand buffer, 1 µL 0.1 M DTT, 1 µL RNaseOUT and 1 µL Superscript III reverse transcriptase (Invitrogen) were added to the RNA sample. The mixture was incubated at 50 ˚C for 60 minutes before the temperature was raised to 70 ˚C for 15 minutes to inactivate the reaction.

2.3.4 Polymerase Chain Reactions (PCR)

PCR was used to amplify sequences of interest. GoTaq (Promega) was used for diagnostic PCRs. KOD polymerase (MERCK) or Q5 polymerase (N.E.B) were used when proofreading was required.

Reactions using GoTaq polymerase had a total volume of 25 µL which was made up of 5 µL 5 x GoTaq buffer, 1.5 µL 25 mM $MgCl_2$, 2.5 µL 2 mM dNTPs, 2.5 µL 10 µM sequence specific forward primer, 2.5 µL 10 µM sequence specific reverse primer, 1 µL template DNA, 0.2 µL (1 unit) GoTaq polymerase and 9.8 µL dH_2O. General PCR cycling conditions consisted of an initial denaturing hold stage for 5 minutes at 95 ˚C followed by 35 cycles of denaturing at 95 ˚C for 30 seconds, annealing at 53 ˚C for 30 seconds and elongation at 72 ˚C for 1 minute. These cycles were followed by a final hold stage at 72 ˚C for 5 minutes, before decreasing sample temperature to 4 ˚C until being retrieved. The annealing temperature and elongation time (underlined above) varied depending on the

sequence specific primers being used and the expected size of the PCR product. For more accurate guides on annealing temperatures used see **7.1 Primer table, use and Tm**.

Reactions using KOD polymerase had a total volume of 50 µL. This consisted of 5 µL 10x KOD buffer, 3 µL of 25 mM $MgSO_4$, 5 µL of 2 mM dNTPs, 1.5 µL of the 10 µM specific forward and reverse primers, 1 µL of 5 µg/µL KOD polymerase, 1 µL template DNA and 32 µL dH_2O. General PCR cycling conditions consisted of an initial hold phase at 95 ˚C followed by 30 cycles of denaturing at 95 ˚C for 20 seconds, annealing at 60˚C for 15 seconds and elongation at 70 ˚C for 1 minute. This was followed by a final hold stage at 70 ˚C for 3 minutes before decreasing the sample temperature to 4 ˚C. The annealing temperature and elongation time used varied depending on the sequence specific primers being used and the target sequence length. Annealing temperatures used can be found in **7.1 Primer table, use and Tm**.

Reactions using Q5 polymerase also had a total volume of 50 µL. This consisted of 10 µL 5x Q5 buffer, 1 µL of 10 mM dNTPs, 2.5 µL of both gene specific forward and gene specific reverse primer at 10 µM, 1 µL of template DNA, 1 µL of Q5 polymerase and 32 µL of dH_2O PCR cycling conditions when using Q5 polymerase were similar to those used for KOD polymerase. Annealing temperatures varied for primers used in Q5 polymerase reactions, these can be found in **7.1 Primer table, use and Tm**.

2.3.5 DNA Gel Electrophoresis

PCR products were analysed by loading 5 µL of sample into wells in a 1-2% agarose gel with 0.7 µL SYBR Safe DNA gel stain per 50 mL of gel, run in 1 x TBE. Electrophoresis was typically carried out for 20-30 minutes at 75 volts. Gels were imaged using UV transillumination where PCR product sizes could be compared next to a 100 base pair or 1 kilobase pair ladder.

2.3.6 PCR Purification

DNA selected for purification following analysis by gel electrophoresis was purified using a QIAquick PCR purification kit (Qiagen) or alternative DNA column clean up kits. When

purification of specific DNA bands in the agarose gel was required, gel band extraction kits were used (QIAquick Gel Extraction). These kits were used following the manufacturer's protocol. Purified DNA samples were quantified using a NanoDrop spectrophotometer.

2.3.7 pGEM-T easy Cloning

pGEM-T easy (Promega) was used for cloning of *in situ* hybridisation probe templates. Purified DNA was ligated into the pGEM-T easy vector (Promega) by combining 3 µL of purified DNA with 5 µL of 2 x ligation buffer, 1 µL pGEM-T easy plasmid and 1 µL T4 DNA ligase. The ligation reaction was incubated at 14 ˚C overnight. As GoTaq polymerase had been used for these PCR amplifications there was no need to add 3' A-overhangs to the PCR product before ligation.

Electroporation was used to insert the pGEM-T easy plasmid (1 µL) into *E. coli* DH5-α cells (100 µL) (BioRad MicroPulser, standard bacteria EC1 setting). After electroporation, cells were diluted in 1 mL SOC. Transformed cells were incubated for 1 hour at 37 ˚C before being spread onto LB AIX (ampicillin, isopropyl β-D-1-thiogalactopyranoside, X-gal) plates. The plates were left to dry and were incubated at 37 ˚C overnight.

The combination of pGEM-T easy and LBAIX plates allowed for blue/white colony screens to be carried out. White colonies, which contained the DNA insert in the plasmid, were selected and removed from the plates and placed in 20 µL of dH_2O. Colony PCR was carried out using 1 µL of bacteria suspension as template DNA and M13 forward and reverse primers to amplify the desired DNA. Amplifications were migrated on an agarose gel to determine their size.

White bacterial colonies containing a PCR product of the expected length were transferred to a fresh LBAIX plate and incubated at 37 ˚C to propagate overnight.

2.3.8 pCR8/GW/TOPO Cloning

The pCR8/GW/TOPO vector (Invitrogen) was used when cloning for Gateway purposes.

Cloning into the TOPO vector required purified PCR products to have 3' A-overhangs. Where this was not the case (*e.g.* when a proof-reading polymerase had been used), GoTaq polymerase was used to add these. The reaction was set up with purified PCR product to give a final concentration of DNA between 10-30 ng/µL, 2 µL of 5 x GoTaq buffer, 0.8 µL of 25 mM $MgCl_2$, 0.4 µL of 10 mM dATP, 0.1 µL of GoTaq polymerase, the volume was adjusted to 10 µL with dH_2O. The reaction was incubated in a PCR machine at 72˚C for 10 minutes. TA-cloning was carried out immediately after 3' A-overhang had been added. 4 µL fresh A-overhang PCR reaction mixture was added to 1 µL of 4 x diluted salt solution, 0.5 µL of dH_2O and 0.5 µL of TOPO vector. The reaction was mixed gently and stored at room temperature overnight.

Electroporation or heat shock techniques were used to insert the plasmid into DH5-α or TOP10 cells depending on availability. For heat shock transformations, 100 µL of chemically competent cells and 1 µL of ligation mixture were mixed, incubated on ice for 30 min and subjected to 45 seconds in a water bath at 42 ˚C before being returned to the ice bath and diluted in 1 mL LB. Transformed cells were incubated for 1 hour at 37 ˚C before being spread onto LB plates containing the appropriate antibiotic resistance. The plates were left to dry and were incubated at 37 ˚C overnight.

TA-cloning can result in the PCR fragment being inserted in an incorrect orientation. Therefore, two colony PCR reactions were carried out for each colony obtained in this process using GoTaq polymerase. In one reaction M13 forward and M13 reverse primers were used. This confirmed presence and size of the insert. In a separate PCR reaction M13 forward and the gene specific reverse primer were used. This only resulted in amplification of the gene product if it was inserted in the correct orientation.

2.3.9 pET15b Cloning

The pET15b expression vector was used for expression of candidate proteins amplified from nematode cDNA. The expression vector was obtained from a stock in the TKS lab group in St. Andrews. This version of pET15b contained a modified linker region with additional restriction enzyme cutting sites (**7.3 pET15b with linker vector**).

The vector was cut with *Nde*1 and *Bam*H1 (Fast Digest, ThermoFisher). Forward and reverse primers were synthesised with corresponding *Nde*1 and *Bam*H1 restriction sites (**7.1 Primer table, use and Tm**) to allow amplification of PCR products with corresponding restriction sites.

Digests were carried out in a 30 µL reaction containing 1 µg of target DNA (vector or PCR product) combined with 3 µL of 10 x Fast Digest buffer and 1 µL of both *Nde*1 and *Bam*H1. The reaction was incubated at 37 ˚C for 1 hour. 1 µL of alkaline phosphatase was subsequently added to pET15b digests before further incubation at 37 ˚C for 1 hour to inhibit re-ligation of the backbone.

Digested vector and digested PCR product inserts were ligated together using T4 DNA ligase (ThermoFisher). A molar ratio of 4:1 insert:vector was calculated using 50 ng of digested vector per insert. The total ligation reaction volume was 10 µL, which included 1 µL of T4 DNA ligase, 1 µL of 10 x DNA ligase buffer and the calculated quantities of digested and cleaned vector and PCR inserts. The reaction was left for 1 hour at room temperature before being used to transform DH5-α cells (chemically competent). Cells were plated onto LB-ampicillin plates and incubated at 37 ˚C overnight.

2.3.10 Cloning into Gateway Vectors pK7FWG2, pK7WGF2 & ApSiPR

Gateway vectors pK7FWG2, pK7WGF2 & ApSiPR were used in plant-based studies further described in section **2.5**. ApSiPR was created to allow apoplastic targeting of recombinantly expressed proteins expressed via *Agrobacterium* infiltration of *Nicotiana benthamiana* plants. These proteins could be tagged with RFP for localisation or pull downs (**7.2 ApSiPR vector**).

TOPO entry clones were digested with *Xba*1 and *Hpa*1 before carrying out the LxR reaction. In a 30 µL reaction 2 µg of TOPO pDNA was combined with 3 µL of 10 x Multicore buffer (Promega), 0.3 µL of 100 x BSA, 0.4 µL of both *Xba*1 and *Hpa*1. The reaction was left to incubate at 37 ˚C for 3 hours. After this incubation a further 0.4 µL of each restriction enzyme was added again and the reaction was incubated for a further

2 hours. 1 µL of alkaline phosphatase was added before incubation for another 45 minutes at 37 ˚C.

Digested products were purified using a QIAquick PCR purification kit. Cleaned PCR products were run on a 1-2% agarose gel to check the digestion. Successfully digested products were quantified using a NanoDrop spectrophotometer.

LxR reactions were set up with a ratio of 3:1 entry clone:destination vector. The reaction consisted of 120-140 ng digested TOPO entry vectors, 0.5 µL LR clonase, 150 ng Gateway destination vector and 10 mM Tris-HCl (pH 8.0) to complete the reaction to 10 µL. The reaction was incubated for 2 hours at room temperature before adding 1 µL of proteinase K (2 µg/µl) after which the reaction was incubated at 37 ˚C for 10 minutes. DH5α cells were transformed by electroporation, incubated for 1 hour at 37 ˚C and spread out onto LB-spectinomycin plates. Plates were then incubated at 37 ˚C overnight to allow colonies to grow.

2.3.11 Cloning into pEHISTEV

pEHISTEV is an expression vector obtained from the Naismith group in St. Andrews (Liu & Naismith, 2009). Where cloning into pET15b had been an issue or expression was not very clear the pEHISTEV vector was used for expressing candidate proteins amplified from nematode cDNA.

The vector was cut with *Nco*1 and *Bam*H1 (Fast Digest, ThermoFisher) as described above. New nematode gene fragments were amplified from cDNA using forward primers with *Pag*1 (*BspH*1) restriction sites. The same *Bam*H1 site reverse primers as used in **2.3.9** could be used for amplification. PCR products were digested with the appropriate enzymes. *Pag*1 and *Nco*1 cut sites are compatible. However, enzyme attachment sites differ meaning these enzymes can be used in place of each other when genes contain one or other cut site within the sequence. Digestion, ligation, transformation into bacteria and amplification of pEHISTEV vectors followed the same methods as those used for pET15b expression vectors.

2.3.12 Colony PCR

Colony PCR was used to determine whether correct insertion of target DNA had occurred from methods described in sections **2.3.7 – 2.3.11**. Colony PCRs used universal primers such as M13 or T7 to avoid false positives that may occur from residual, non-ligated, insert DNA. Gene specific reverse primers could be used to determine orientation of the inserted DNA. Incorrect orientations would occasionally occur when using a TA-cloning system.

Colonies were removed from plates and placed into 20 µL of dH_2O. 1 µL of this solution was used as the PCR template. Reactions were carried out in 10 µL volumes containing 1 µL template bacterial suspension, 0.5 mL of 0.5 µM forward primer, 0.5 mL of 0.5 µM reverse primer, 5 µL of GoTaq green Hot Start mix (containing dNTPs and $MgCl_2$), 3 µL of dH_2O.

General PCR cycling conditions consisted of an initial denaturing hold stage for 10 minutes at 95 ˚C followed by 25 cycles of denaturing at 95 ˚C for 30 seconds, annealing at 53 ˚C for 30 seconds and elongation at 72 ˚C for 2 minutes. These cycles were followed by a final hold stage at 72 ˚C for 5 minutes before decreasing sample temperature to 4 ˚C. The annealing temperature (underline above) varied depending on the primer sets being used.

2.3.13 Bacterial Colony Propagation and DNA extraction

Bacterial colonies deemed to contain the correct insert sequence by colony PCR were collected and transferred to liquid culture. Stock and working concentrations of antibiotics can be seen in **Table 2.2** Propagation in liquid culture was carried out at 37 ˚C overnight, the samples were placed on a shaking platform to ensure aeration of bacteria. Plasmids were purified using a GeneJet plasmid purification kit (ThermoFisher Scientific) following the manufacturer's instructions. Plasmid concentration was determined using a NanoDrop spectrophotometer.

Table 2.2 – Antibiotic concentrations for bacterial liquid cultures

Antibiotic	Stock concentration (mg/mL)	Working concentration (µg/mL)	Stock:LB ratio
Ampicillin	100	100	1:1000
Spectinomycin	100	100	1:1000
Kanamycin	50	50	1:1000
Chloramphenicol	30	7.5	0.25:1000
Gentamycin	25	6.25	0.25:1000
Rifampicin	25	6.25	0.25:1000

2.3.14 Sanger Sequencing

DNA sequencing was either performed by GATC Biotech using the GATC SUPREME run service or carried out by the sequencing lab located within the James Hutton Institute. In both cases, in-house sequencing primers were used. Universal primers (M13 or T7) were used to confirm that the insert was within the correct reading frame. Sequencing results were analysed in BIOEDIT software (Hall, 1999) before being aligned to the expected sequence using the online tool, MultAlin (Corpet, 1988).

2.4 Recombinant Protein Expression

Functional studies of identified proteins required collection of correctly sequenced, soluble protein. Genes transcribing identified surface proteins were amplified using Q5 polymerase reactions as described in section **2.3.4**. Proteins were N-terminally 6x His tagged by either the pET15b or pEHISTEV expression vectors, therefore primers included a stop codon.

2.4.1 Selecting Expression Cells

Genetic sequences in addition to the 6x His tag, were entered into the 'Rare codon calculator' (http://svkdipmpps-rarecodon-analysis.blogspot.co.uk/) that highlighted codons that are rare or unusual for expression within *E. coli*.

pET15b and pEHISTEV plasmids containing the correct insert in frame with the 6x His tag, were transformed into Rosetta 2 cells (Merck) due to high levels of rare codon usage in selected nematode genes favouring arginine, glycine, threonine, leucine, isoleucine and proline amino acids.

2.4.2 Expression Trials

Transformed Rosetta 2 cells were spread onto plates containing ampicillin (kanamycin for pEHISTEV) and chloramphenicol and incubated at 37 ˚C overnight. Rosetta 2 colonies were tested for the expression vector using colony PCR (**2.3.12**), colonies with the insert were propagated in 5 mL LB with 5 µL of 100mg/mL ampicillin, 1.25 µL of 30 mg/mL chloramphenicol in a shaking incubator at 37 ˚C overnight to produce a starter culture. A non-induced cell sample was taken for later comparison against induced samples. 300 µL of starter culture was removed, cells were pelleted and resuspended in 100 µL sample buffer (**7.12 Buffer recipes**). Cells were boiled at 95 ˚C for 10 minutes to lyse the cells, samples were then kept at -20 ˚C.

Starter cultures were used to inoculate 10 mL fresh LB or auto-induction media (AI media). Growth was continued until a cell density reading OD_{600} of 0.6 was reached. OD_{600} readings were measured using a spectrophotometer. At the optimal density cells were cooled to trial expression temperatures of 18 ˚C or 25 ˚C. If the culture was to be grown in LB then expression was induced using 1 mM IPTG. Cultures were grown at expression temperatures overnight.

Similarly to the non-induced sample, an induced, whole cell sample of 300 µL was taken after 24 hours at expression temperature. This was prepared in the same way as the non-induced sample.

Remaining cells were pelleted and resuspended in 700 µL resuspension buffer (10 mM imidazole, 50 mM Tris pH 7.5, 300 mM NaCl). Cells were lysed using an ultrasonic probe cycling between 15 seconds at 12 kHz and 15 seconds of rest. Cells were kept on ice at all times. Suspensions were pelleted leaving all soluble proteins in the supernatant.

Supernatants were removed and added to 2x sample buffer before being boiled at 95 ˚C, samples were stored at -20 ˚C.

2.4.3 SDS-PAGE

Sodium dodecyl sulphate polyacrylamide gel electrophoresis was carried out to separate proteins from un-induced, whole cell and soluble fractions. Gels were made in house consisting of a separating and a stacking portion. Separating portions were made to contain 16% acrylamide. **Table 2.3** shows quantities of reagents for making 4 gels. Lab-made gels were not used for any samples that would ultimately be sent to mass spectrometry for protein identification. If protein samples were to be submitted for mass spectrometry analysis, pre-cast gels were purchased (NuPAGE, Thermofisher) to reduce the presence of contaminating proteins (**2.6**).

Gels were run at 140 V until the dye front reached the base of the gel (approximately 90 minutes). Coomassie staining allowed visualisation within the gel. If gels were used for Western Blots then they were left unstained.

Table 2.3 – Polyacrylamide gel casting reagents

	Separating Gel	Stacking Gel
30% Acrylamide (ProtoGel)	10.6 mL	1.3 mL
1.5M Tris pH 8.8	5 mL	/
0.5M Tris pH 6.8	/	2.5 mL
10% SDS	200 µL	100 µL
H_2O	3.8 mL	6 mL
TEMED	30 µL	20 µL
10% APS	200 µL	150 µL

2.4.4 Western Blotting

Western blots were carried out using a semi-dry transfer apparatus. Transfers to nitrocellulose membranes followed a ‘semi-dry’ protocol and took 30 minutes. The

number of milliamps varied depending on the number of gels being transferred, 80 mA per gel.

Once proteins were transferred the membrane was blocked for 1 hour in blocking buffer (PBS supplemented with 5% milk powder). Blocking buffer was removed by washing twice in PBS. Proteins were detected with appropriate primary antibodies. Primary antibody solution was made up with PBS, 1% milk powder and primary antibodies at an appropriate dilution (1:20,000 for anti-HIS). Membranes were left for 1 hour at room temperature with end to end agitation. After an hour, the primary antibody solution was removed and the blot was washed 6 times in PBS over 1 hour.

Primary antibodies were detected with secondary antibodies coupled to a fluorescent tag for visualisation. Membranes were submerged in a solution of PBS, 1% milk powder and secondary antibodies at 1:15,000 dilution. Membranes were protected from direct light by covering with tin foil and were incubated at room temperature for an hour with end to end agitation. The secondary antibody solution was removed and the blot was washed as previously described.

Western blots were imaged on a LiCor Odyssey imaging system. This system scanned for fluorophores emitting at either 700 or 800 nm wavelengths. Exposure and contrast were automatically adjusted by the system to capture the best image.

2.4.5 Protein Purification

Once optimal expression conditions had been determined (**2.4.2**) large quantities of cells could be cultured for increased protein recovery. 1 L of cell culture was pelleted, resuspended in lysis buffer (50 mM Tris pH 7.5, 0.3 M NaCl, 10 mM imidazole) and lysed using the same sonication conditions as in **2.4.2**. Lysed cells were separated from soluble proteins by centrifuging at 20,000 rpm for 30 minutes. Soluble proteins were removed and the cell pellet was discarded. Recombinant proteins were tagged with a 6xHis tag and so could be purified using a nickel column (GE Life Sciences).

The nickel column was equilibrated by running 5 column volumes of lysis buffer through the column. Solutions were run through the column with a peristaltic pump. The pump

was set to flow at 1 mL/min to load soluble protein samples. Once the protein sample was loaded the column was washed with at least 5 column volumes of wash buffer (lysis buffer, 20-50 mM imidazole). Protein coming off the column was monitored with a Quick-start Bradford assay (BioRad) (40 µL Bradford reagent + 10 µL column flow through). The column was washed until there was almost no protein coming off the column. Proteins bound to the column were eluted using elution buffer (lysis buffer, 100 mM imidazole). Fractions from column purifications were monitored via SDS-PAGE.

2.4.6 Dialysis/buffer exchange

Protein buffer exchange was carried out using P25 desalting columns (EMP BioTech CentriPure). Columns were equilibrated with 25 mL destination buffer. 2.5 mL of sample was loaded and eluted with 3.5 mL destination buffer.

2.4.7 TEV cleavage of 6x His tags

In general, bacterial proteins that were His-rich also bound and eluted to the nickel column with His-tagged recombinant proteins. Removal of the His tag and re-running the sample through the nickel column corrected this as recombinant proteins would no longer bind to the column once the tag was removed. This was only an option for recombinant proteins expressed using the pEHISTEV vector.

Protein samples were exchanged into TEV cleavage buffer (150 mM KCl, 20 mM Tris-HCl pH 7.7, 3 mM $MgCl_2$, 0.5 mM EDTA pH 8.0, 0.1% TWEEN, 1 mM DTT). TEV protease was produced in house. TEV protease was added to recombinant protein at a ratio of 1:10. Samples were left incubating at 4 ˚C overnight with gentle agitation.

The TEV cleavage reaction was run back through the nickel column. After cleavage recombinant proteins no longer bound to the column and so were collected in the flow through. HIS-tagged TEV protease was not present in the flow through as it had bound to the nickel column. Cleavage was monitored using SDS-PAGE and western blots.

2.4.8 Recombinant protein storage

Recombinant proteins were exchanged into storage buffer (50 mM HEPES (pH 7.8), 0.1 M NaCl, 10% glycerol). Aliquots of recombinant protein were snap frozen in liquid nitrogen before being stored at -80 ˚C.

2.5 Immunofluorescence

Nematode samples were collected for immunofluorescence following sections **2.1.2-2.1.4**.

Samples were washed twice in 1 mL of M9 buffer (**7.12 Buffer recipes**). Next, samples were centrifuged out of buffer (5200 x g. x 10 mins) and incubated in 1 mL 1:500 antiserum, rotating for 1 hour at room temperature. Pre-immune serum and no anti serum controls were carried out following the same method. These confirm that secondary antibodies were not binding to other native nematode proteins ahead of the primary antibody and that the antibody present in the antiserum was the reason for antibody binding.

Samples were washed out of primary serums into 1 mL M9 buffer three times. 5 µg/mL dilutions of anti-rabbit secondary antibodies conjugated to a fluorophore were made from a 2 mg/mL stock. All samples were incubated in secondary antibodies, rotating at room temperature for 1 hour in the dark to avoid bleaching of the fluorophore.

Samples were washed a further 3 times in M9 buffer to remove any background staining and unbound antibodies. Microscope slides were prepared and samples were visualised using a fluorescence microscope.

2.6 Protein Mass spectrometry

Protein mass spectrometry was carried out by the in-house facility at the University of St Andrews. Samples were reduced with DTT, alkylated with Iodoacetamide and digested at 37 ˚C with trypsin. Samples in gels were extracted from diced gel pieces in 5% formic acid to an extraction volume of 20 µL. A proportion of the sample was injected onto an eksigent nanoLC set up in ‘trap elute’ configuration using a Pepmap column and

trap (ThermoScientific). Peptides were eluted in a linear gradient over 180 minutes and flowed directly into the Sciex 5600+ Q-Tof mass spectrometer. Survey scans were carried out between 400-1200 m/z and the strongest 15 peptides from each scan were fragmented to give MS-MS spectra from 100-2000 m/z. Spectra were extracted using the mgf generator script from Sciex and the data searched using the mascot search algorithm against *Globodera_rostochiensis_*(Nematode) (14,309 sequences) downloaded March 2016 from the '*Globodera* genomes project' website (globodera.bio.ed.ac.uk). The fasta file containing mRNA sequences for each gene was gathered in the same way.

2.7 Bioinformatics

2.7.1 BLAST searches

The basic local alignment search tool (BLAST) was used to identify characterised homologues of genes or proteins in other species. BLAST finds regions of similarity between given sequences and a database calculating significance scores. For searching against a non-redundant database the NCBI website was used (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul *et al.*, 1997). For searches against specific nematode databases BLAST was used through WormBase ParaSite (https://parasite.wormbase.org/Tools/Blast?db=core) (Howe *et al.*, 2017).

2.7.2 Expression data

Globodera rostochiensis normalised expression data was obtained from the *G. rostochiensis* genome publication and downloaded from the supplementary data (Additional file 7, file S1) (Eves-van den Akker *et al.*, 2016). *G. pallida* lifestage specific normalised gene expression data was obtained from the James Hutton Institute intranet following publication of the transcriptome data in 2014 (Cotton *et al.*, 2014).

2.7.3 Sequence alignments

Sequence alignments for comparing BLAST results or checking identified DNA sequences returned from sequencing (**2.3.14**) were aligned using MultAlin (Corpet, 1988).

Pairwise alignment of larger numbers of sequences was carried out using MUSCLE (Edgar, 2004; McWilliam *et al.*, 2013). MUSCLE was either accessed online or alignments using MUSLCE default settings were achieved through Jalview which simultaneously allowed identification of other sequence parameters (Waterhouse *et al.*, 2009).

2.7.4 Protein characterisation

2.7.4.1 ExPASy tools

ExPASy is a collection of bioinformatics tools and databases provided by the Swiss Institute of Bioinformatics (SIB). ProtParam computes various protein parameters from a given peptide sequence. Primarily ProtParam was used to aid protein expression by evaluating expressed sequences for amino acid composition, molecular weight and theoretical pI. The ExPASy translate tool was used to translate nucleotide sequences into amino acid sequences (Gasteiger *et al.*, 2005).

2.7.4.2 DTU Bioinformatics tools

DTU bioinformatics (Technical University of Denmark) offer bioinformatics tools capable of predicting protein features from a given sequence. SignalP (versions 4.0 and 4.1) was used to identify the presence of a predictable signal peptide for classical secretion within a sequence (Petersen *et al.*, 2011; Nielsen, 2017). Default settings for eukaryotic species were used for nematode sequences.

The TMHMM server (2.0) was used to predict transmembrane helices within a given protein sequence (Krogh *et al.*, 2001).

2.7.5 Data handling software

2.9.5.1 Vector mapping

SnapGene Viewer (GSL Biotech) was used to map cloning and expression vectors. Predictions of common sequences and restriction enzyme cutting sites were carried out by the software. Maps for the novel ApSiPR vector were produced in SnapGene (**7.2 ApSiPR vector**).

2.7.5.2 Scaffold viewer

Scaffold viewer software (Proteome Software) was used to handle protein mass spectrometry data. This software allowed easy comparison of samples submitted for ESI-TOF analysis and gave visual representations of identified peptides against the given database. The software calculated numbers of identified peptides for a given protein and determined the number of peptides that were unique to proteins within the database (Searle, 2010).

2.8 References

ALTSCHUL, S.F., MADDEN, T.L., SCHÄFFER, A.A., ZHANG, J., ZHANG, Z., MILLER, W. & LIPMAN, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **25**, 3389–3402.

CLARKE, A.J., COX, P.M. & SHEPHERD, A.M. 1967. The chemical composition of the egg shells of the potato cyst-nematode, Heterodera rostochiensis Woll. *The Biochemical Journal*, **104**, 1056–1060.

CORPET, F. 1988. Nucleic Acids Research Multiple sequence alignment with hierarchical clustering. *Nucl. Acids Research,* **16**, 10881-10890.

COTTON, J.A., LILLEY, C.J., JONES, L.M., KIKUCHI, T., REID, A.J., THORPE, P., TSAI, I.J., … URWIN, P.E. 2014. The genome and life-stage specific transcriptomes of Globodera pallida elucidate key aspects of plant parasitism by a cyst nematode. *Genome Biology*, **15**, R43.

EDGAR, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**, 1792–1797.

EVES-VAN DEN AKKER, S., LAETSCH, D.R., THORPE, P., LILLEY, C.J., DANCHIN, E.G.J., DA ROCHA, M., RANCUREL, C., … JONES, J.T. 2016. The genome of the yellow potato cyst nematode, Globodera rostochiensis, reveals insights into the basis of parasitism and virulence. *Genome Biology*, **17**, 124.

GASTEIGER, E., HOOGLAND, C., GATTIKER, A., DUVAUD, S., WILKINS, M.R., APPEL, R.D. & BAIROCH, A. 2005. Protein Identification and Analysis Tools on the ExPASy Server. *In The Proteomics Protocols Handbook*. Totowa, NJ: Humana Press, 571–607.

HALL, T. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series 41*.

HOWE, K.L., BOLT, B.J., SHAFIE, M., KERSEY, P. & BERRIMAN, M. 2017. WormBase ParaSite − a comprehensive resource for helminth genomics. *Molecular and Biochemical Parasitology*, **215**, 2–10.

KROGH, A., LARSSON, B., VON HEIJNE, G. & SONNHAMMER, E.L.L. 2001. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes, Edited by F. Cohen. *Journal of Molecular Biology*, **305**, 567–

580.

LIU, H. & NAISMITH, J.H. 2009. A simple and efficient expression and purification system using two newly constructed vectors. *Protein expression and purification*, **63**, 102–111.

MCWILLIAM, H., LI, W., ULUDAG, M., SQUIZZATO, S., PARK, Y.M., BUSO, N., COWLEY, A.P. & LOPEZ, R. 2013. Analysis Tool Web Services from the EMBL-EBI. *Nucleic acids research*, **41**, W597-600.

NIELSEN, H. 2017. Predicting Secretory Proteins with SignalP. *In Methods in molecular biology (Clifton, N.J.)*. 59–73.

PETERSEN, T.N., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods*, **8**, 785–786.

SEARLE, B.C. 2010. Scaffold: A bioinformatic tool for validating MS/MS-based proteomic studies. *PROTEOMICS*, **10**, 1265–1269.

WATERHOUSE, A.M., PROCTER, J.B., MARTIN, D.M.A., CLAMP, M. & BARTON, G.J. 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189–1191.

3. Identification and characterisation of PCN eggshell proteins

3.1 Introduction

Although hatching of plant-parasitic nematodes has been well described at a physiological level, the molecular mechanisms that underpin the hatching process remain relatively underexplored. In the case of PCN, it is well known that hatching occurs as a response to exposure to root diffusates (RD) from host plants (Perry, 2002). This allows synchronisation with the host plants' lifecycle and gives the juvenile the best chance of successfully infecting the plant. If there is no signal from the roots of a host plant the majority of the population will not hatch. However, in some field conditions a small proportion of PCN will hatch spontaneously, likely due to a physiological accident such as an fault in eggshell permeability (Devine & Ryan, 2005; Kerry *et al.*, 2009).

As discussed in chapter 1 (**1.8.1**), a key stage of the PCN hatching cascade is the permeabilisation of the lipid layer. This process allows influx of water into the eggshell and rehydration of the nematode after its quiescent period. Lipid barrier permeabilisation is a process that must be highly regulated for the nematode to be able to synchronise its lifecycle with that of its host. Hatching too soon or too late will result in the nematode being unable to fulfil its lifecycle due to lack of a viable host. Similarly, if the nematode hatched without being in close proximity to a host this would increase the lipid reserves used whilst traveling to the host. Having reduced lipid availability in the infective juvenile then reduces the chances of successful infection (Atkinson *et al.*, 1985).

Unlike some species of plant-parasitic nematode, metabolic activation of PCN occurs after the permeability change in the eggshell (Perry, 2002). This would suggest that there is a component in the eggshell that regulates changes to the lipid layer. Likewise,

there must also be a related system that responds to hatching factors from the host plant. The hatching process is summarised in Figure 3.1.

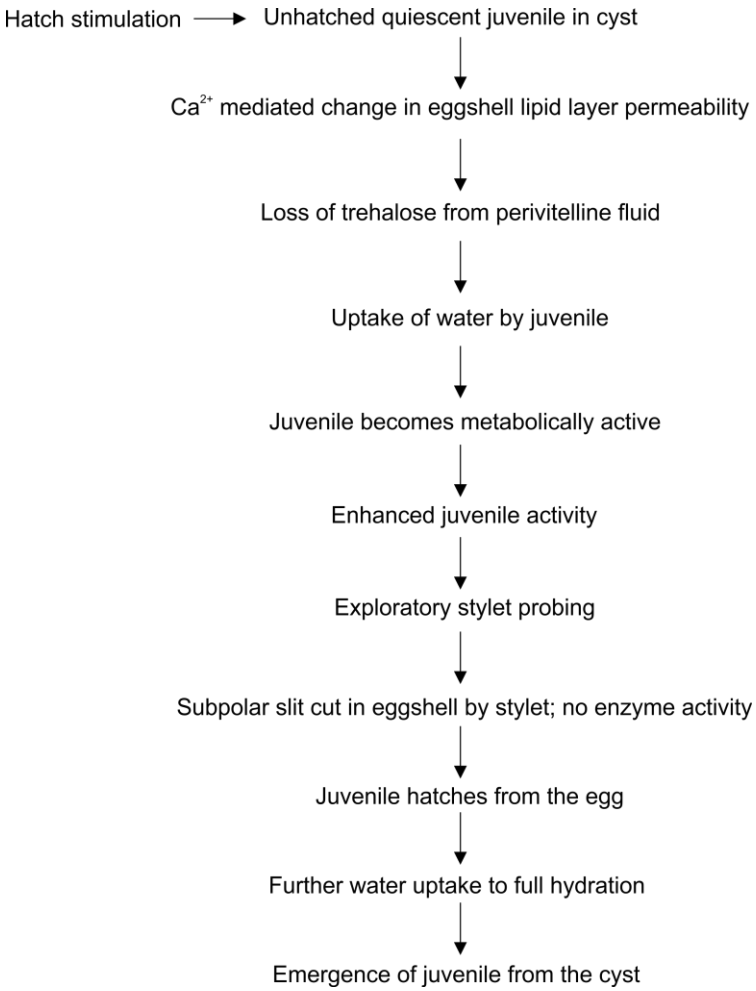


Figure 3.1 – Summary of processes involved in hatching of PCN
Flow chart stating the main steps a juvenile PCN must go through before hatching. Flow chart originally from (Masler & Perry, 2018).

3.1.1 Calcium binding in PCN eggshells

Calcium is thought to have a key role in the PCN hatching cascade. Three categories of calcium binding sites have been suggested as being present in eggshells of *G. rostochiensis*. The first two are sites located within the lipoprotein layer where calcium can either be released or bound in the presence of hatching factors. Another

suggested site is found on the outer eggshell layers but this is unlikely to be related to hatching (Clarke & Perry, 1985).

Calcium will not stably bind to lipids for long-term storage. Consequently, it is likely that any eggshell associated calcium would be bound to a protein. An eggshell associated calcium binding sialoglycoprotein was inferred by Atkinson & Taylor (1983). However, the properties of this protein were only inferred based on experimental data gathered using ruthenium red and lanthanum to block calcium binding. The presence of the sialoglycoprotein in the eggshell was never confirmed through use of antibodies or any other protein analysis.

Ions play an important role in hatching of many nematode species and not just PCN. Zinc is often used to induce hatch of *Heterodera glycines in vitro* and *Ascaris suum* will readily hatch in the presence of sodium (Clarke & Perry, 1980; Tefft & Bone, 1984). However, in both of the above cases free ions can induce hatch of the nematodes whereas free calcium does not induce PCN hatching (Clarke & Hennessy, 1983). Interestingly, where eggshell calcium presence is reduced after PCN hatching (Clarke & Perry, 1985), *Ascaris* eggshells hatched artificially contain substantially more sodium after hatching (Clarke & Perry, 1988).

3.1.2 Receptor-ligand hatching factor interaction

Parasitic nematodes exhibiting a broad host range are not obliged to coordinate their lifecycle with their host. When they hatch, it is likely that they will be able to feed on something nearby. However, host-specific parasites, such as PCN, must be synchronised with the host plant to complete their lifecycle. Although ions of calcium have been shown to be involved in hatching of PCN, it is unlikely that these ions alone are the initiators of the hatching cascade *in vivo*. Synchronisation of the parasite life cycle to that of its host would suggest that another factor, specific to the host, is used to initiate the hatching cascade. For PCN, these factors are known to be the root diffusates originating from the host plant. However, a receptor for these root diffusates has not yet been identified. In fact, there is debate as to whether the root diffusates are detected by receptors at all.

A significant number of PCN can be hatched after just 5 minutes exposure to RD at weekly intervals (Forrest & Perry, 1980). This indicates a rapid and selective system, suggestive of a ligand-receptor mediated interaction. However, if PCN hatching was reliant upon an eggshell originating receptor responding directly to RD then 100% egg hatching after a single exposure to RD would be possible, however, this does not happen. Characteristically, some eggs hatch after initial exposure, but most eggs need re-stimulating with RD. Even then some eggs will remain unhatched, but viable, potentially as a survival strategy (Perry & Clarke, 1981).

3.1.3 Known nematode eggshell proteins

Early work, based upon analysis of nitrogen content, suggested that PCN eggshells consisted of up to 59% protein, making it the major component of the eggshell (Clarke *et al.*, 1967). However, to date, none of these proteins have been identified or localised to the eggshell.

Research originally designed to identify the number of eggs per soil for diagnostic purposes gave some insights into eggshell protein similarities and differences between species of nematodes. Kennedy *et al.* (1997), used crushed eggshells as the immunogen to produce polyclonal antibodies against *H. glycines* eggs. Western blots showed that eggs within the genus (*Heterodera*) had more similarity than between genera (*Heterodera* vs *Meloidogyne*). Additional differences in detected proteins were seen between species within the genus (*H. glycines* vs *H. schachtii*). Very few of the detected proteins were also present within juveniles of the same species. This work highlights the variability of eggshells between PPN species. No proteins bound by the polyclonal antibodies were identified.

The presence of enzymes in eggshells of PPN is discussed later (**Chapter 4 – Identification and characterisation of PCN eggshell lipids**).

More information is available about the eggshell of the model nematode species *Caenorhabditis elegans*. *C. elegans* is a free-living nematode and so does not rely upon

a host to initiate hatching. However, there are likely to be structural eggshell proteins that are common between *C. elegans* and species of PPN.

Most proteins that have been identified in the *C. elegans* eggshell are proteins associated with forming the eggshell. These proteins are not necessarily ones that are retained within the formed eggshell or that are involved in a hatching cascade. For example, chitin synthase 1 (CHS-1) is responsible for chitin synthesis in the eggshell, but once the chitin layer is formed, CHS-1 is not held within the structure. However, one group of identified proteins retained within the eggshell are the chondroitin proteoglycans (CpGs) (Stein & Golden, 2015). The chondroitin proteoglycan layer in *C elegans* sits between the chitin layer and the inner lipid layers of the eggshell. For many years this layer was thought to be made up of lipids (Foor, 1967; Perry & Trett, 1986). This proteoglycan layer exhibits chitin binding domains, potentially linking it to the chitin layer above. Unlike glycoproteins, proteoglycans are covalently bonded to carbohydrate chains containing amino sugars and therefore, nitrogen.

The function of the chondroitin proteoglycan layer in *C. elegans* eggshells remains unknown. Due to the hierarchical formation of the eggshell, disruption of this layer also disrupts formation of the inner lipid layer, disturbing the permeability barrier producing embryonic lethality. However, mutants removing proteoglycan functionality produce viable eggshells suggesting that although the CpG core protein is required for eggshell formation, the chondroitin attachments are not necessary (Olson *et al.*, 2012; Stein & Golden, 2015).

3.1.4 Chapter 3 aims

Due to the current lack of knowledge relating to any PCN eggshell proteins, the aim of this chapter was to identify any proteinaceous components that could be extracted from the eggshell. Identification of proteins that have functions potentially related to the hatching cascade or eggshell structure will be investigated further.

3.2 Materials & Methods

3.2.1 Biological materials

Eggshells were collected using the techniques described in **2.1.4**. Collected materials were used immediately after collection and were never stored for longer than 24 hours. To make the most of stocks of PCN, eggshells collected for protein extractions were also used for lipid extractions.

3.2.2 Eggshell protein extraction

Eggshells were incubated in 750 µL 2:1 methanol:chloroform overnight at 4 °C. Solvents were of analytical grade. Samples were kept in glass vials and extractions were carried out on a vibrating platform. The next day 250 µL chloroform was added to create a biphasic solution. The lower chloroform layer was removed (and retained for analysis of lipids). The remaining methanol:water was dried under a stream of nitrogen gas. Eggshells were resuspended in a methanol:water (80:20) mix and placed back on the vibrating platform for a further 5 hours. If the proteins were to be run into a polyacrylamide gel, the eggshells were left in solution as they would be filtered out by the gel. If the eggshells needed to be removed this was achieved using the sieving technique described in **2.1.4**. Protein mass spectrometry was carried out as described in **2.6**.

3.2.3 Cloning and expressing recombinant eggshell proteins

Genes encoding proteins of interest as identified by mass spectrometry were amplified from *G. rostochiensis* cDNA. PCR for amplification of the target gene used Q5 polymerase (**2.3.5**). Primers were designed to amplify the full coding region of the gene with the addition of appropriate restriction enzyme sites (**7.1 Primer table, use and Tm**) allowing cloning into the pEHISTEV vector (**2.3.11**). Plasmids containing desired genes were transferred to *E. coli* (Rosetta 2) (**2.3.8**), after reaching an optimal density expression was induced and carried out at 18 °C overnight. Protein purification was carried out as described in **2.4.5**.

3.2.4 Peptide synthesis and production of antisera

Peptides were synthesised and used for production of antisera in rabbits by Eurogentec. The 28 day 'Speedy' protocol allowed two antisera to be raised against regions of the protein of interest. Immunogenic regions of the protein were determined by Eurogentec, these regions were checked for similarities with other proteins by BLAST searching against the *G. rostochiensis* genome to rule out any potential off-target hits. After completion of the program, Eurogentec provided pre-immune serum, large bleed samples, final bleed serum, remaining coupled peptide and purified antibodies.

3.2.5 Immunofluorescence of proteins in PCN eggshells

Eggshell samples were fixed in 4% paraformaldehyde (PFA) overnight at 4 °C. PFA was removed by washing twice with PBS. Eggshell material was permeabilised by incubating in PBS + 0.1% Triton-X100 followed by two more PBS washes to remove remaining Triton-X100, centrifuging eggshells at 2600 x g between each wash. Non-specific binding was blocked by incubating in 2% BSA for 1 hour at room temperature with gentle agitation. This was followed by 3 further washes in 0.1% BSA. Primary antibody solution (1:1000) in 0.1% BSA was applied to eggshells for 2 hours at room temperature; samples were left rotating during this incubation. Primary antibodies were removed by washing samples in 0.1% BSA 6 times, changing the solution every 10 minutes. Fluorescent anti rabbit secondary antibodies were used to localise primary antibodies. Secondary antibodies (Goat anti-rabbit Alexa Fluor Plus 488, ThermoFisher) were incubated with eggshells for 1 hour at room temperature while rotating. 6 further 10 minutes washes in PBS were used to remove any un-bound secondary antibody. A control using rabbit pre-immune serum was used to highlight any non-specific interactions.

Eggshells that had not been exposed to any antibodies were first imaged to reduce autofluorescence. Computational gain in channel 1 (associated with emission at 488 nm) was lowered to a level where autofluorescence was no longer visible, effectively acting as a 'zero' setting on the microscope. Microscope conditions were not altered after this optimisation (**Table 3.1**).

Imaging of eggshells was carried out using a Zeiss LSM710 laser scanning microscope mounted on an AxioImager z2 motorised upright microscope.

Table 3.1 – Imaging settings for detection of fluorescence in PCN eggshells.
Track 1 (488 nm) detected fluorescence from AlexaPlus488 anti rabbit. Track 2 (561 nm) detects auto fluorescence levels of the eggshell to check for consistency between samples. These settings were retained as standard for all eggshell fluorescence imaging experiments.

Track	Master gain	Digital gain	Digital offset	Pinhole	Filters	Lasers
1	600	1.00	0.00	59 µm	500-530	488 nm
2	800	1.00	0.00	59 µm	591-630	561 nm

3.2.6 Creation of transgenic lines for RNAi

Hairpin RNA was created by cloning a short, gene-specific sequence into vectors pHannibal and later, pART-27. Four primers were synthesised to amplify two complementary gene sequences. Sequences were amplified using Q5 polymerase following the protocol in **2.3.4** from nematode cDNA. The forward to reverse fragment using primers xho1_3104RNAi_F and kpn1_3104RNAi_R (**7.1 Primer table, use and Tm**) was amplified and the product was subsequently purified using a PCR purification kit (**2.3.6**). The complementary reverse to forward fragment was amplified using primers hindIII_3104RNAi_R, xba1_3104RNAi_F and purified in the same way.

Amplified PCR fragments were digested using restriction enzymes corresponding to the primer name. Initially digestion was carried out on the forward to reverse fragment and the pHannibal vector. In both cases, 500 ng of DNA was added to 1 µL corresponding fast digest enzymes (*Xho*1 and *Kpn*1 for forward to reverse fragment) (ThermoFisher), 2 µL fast digest buffer and made up to 20 µL with SDW. Digestion was carried out for 1 hour at 37 °C.

Digested vector and digested PCR product inserts were ligated together using T4 DNA ligase (ThermoFisher). A molar ratio of 4:1 insert:vector was calculated using 50 ng of digested vector per insert. The total ligation reaction volume was 10 µL which included 1 µL of T4 DNA ligase, 1 µL of 10 x DNA ligase buffer and the calculated quantities of

digest and cleaned vector and PCR inserts. The reaction was left for 1 hour at room temperature before being used to transform DH5-α cells (chemically competent). Cells were incubated for 1 hour at 37 °C to allow expression of the antibiotic resistance gene before being spread onto LB-Ampicillin plates. Plates were incubated overnight at 37 °C to allow colonies to grow.

Colony PCR was carried out as described in section **2.3.12** to identify colonies that contained the pHannibal and forward to reverse fragment insert. Colony PCR used the xho1_3104RNAi_F primer and FRAG_1_R primers (**7.1 Primer table, use and Tm**). Colonies containing the correct vector were propagated and subsequently sent for sequencing (**2.3.13, 2.3.14**).

Upon obtaining the required results from sequencing, pHannibal with the forward to reverse fragment was digested with *Hin*dIII and *Xba*I. The complementary reverse to forward fragment amplified using primers hindIII_3104RNAi_R and xba1_3104RNAi_F was digested using the same enzymes. Vector and digested reverse to forward fragment were ligated, transformed into *E. coli*, propagated and sequenced as described above. However, colony PCR to confirm insertion of the reverse to forwards fragment used primers FRAG_2_F and xba1_3104RNAi_F. The resulting pHannibal vector contained forward to reverse and complementary reverse to forward fragments flanking a PDK intron (Wesley *et al.*, 2001).

Both modified pHannibal and pART-27 were digested with *Not*I. The cut pHannibal fragment containing forward to reverse and complementary reverse to forward fragments flanking a PDK spacer intron and recipient pART-27 vector were purified, ligated together and transformed into *E. coli* as before (Gleave, 1992). Cells were plated onto LB-Spectinomycin plates. Colony PCR was carried out as per **2.3.6** using xho1_3104RNAi_F primer and FRAG_1_R primers. Colonies containing the correct sequence were propagated and send for sequencing (**2.3.13**, **2.3.14**).

pART-27 containing the GROS_g03104 FR-PDK intron-RF sequence was electroporated (**2.3.8**) into competent AGL1 *Agrobacterium tumefaciens* before adding 1 mL of SOC and

incubating for 2 hrs at 28 °C. The *Agrobacterium* culture was then plated out onto LB-spectinomycin plates and incubated at 28 °C for 48 hours. Colony PCR was carried out as described in section **2.3.12** using xho1_3104RNAi_F primer and FRAG_1_R primers to confirm presence of the RNAi construct in the *Agrobacterium* colonies. Colonies were grown in 5 mL liquid culture at 28 °C overnight.

3.2.6.1 Transformation of 'Desiree'

Sterilised internodes were cut from 4-6 week old 'Desiree' plantlets (provided by S.A.S.A.) and placed onto an agar plate. Transformed *Agrobacterium* (**3.2.6**) solution was placed onto the same agar plate and the plate was gently rocked at room temperature to ensure even coverage of bacteria onto plant tissue. Internodes were transferred onto callus induction media plates (1 L MS30 media with 2.5 mg Zeatin Riboside, 0.2 mg naphthalene acetic acid and 0.02 mg gibberellic acid). No antibiotic was present on these plates to avoid further stress. After 2-3 days internodes were transferred to agar plates containing Timentin® and kanamycin for 2 weeks.

At this point some internodes had formed calli. Live calli were transferred to shoot promotion media plates (1 L MS30 media with 2 mg Zeatin Riboside, 0.02 mg naphthalene acetic acid, 0.02 mg gibberellic acid and kanamycin at 50 mg/mL). After 2 more weeks calli shoots were excised and transferred to MS30 root induction media plates containing the same antibiotics. As roots began to form, leaf material was removed to check for the presence of the GROS_g03104 FR-PDK intron-RF insert. gDNA was extracted from this material using the DNeasy Plant Mini Kit (Qiagen) using the manufacturer's guidelines. Conditions for a colony PCR (**2.3.12**) were used along with gDNA extractions and primers Kana_ntp_II_F and Kana_ntpII_R (**7.1 Primer table, use and Tm**) to identify successfully transformed plants on the basis of the presence of the kanamycin resistance gene. Stem tips were cut from plants identified as carrying the transgene and were planted into insecticide-free compost. Control plants for this process included calli originally formed through transformation with *Agrobacterium* that do not contain the modified pART-27 vector.

3.2.7 Semi quantitative reverse transcription (SQRT) PCR to assess transgene expression in transformed 'Desiree'

Transformed plants were left to grow in 3.5-inch pots for several weeks. Leaf material was taken from each plant and flash frozen in liquid nitrogen for RNA extraction. 7 autoclaved 5 mm metal beads were added to a 1.5 mL Eppendorf tube containing frozen leaf material. The microfuge tube was vortexed vigorously until the leaf material was ground into a fine powder, refreezing in liquid nitrogen where necessary. RNA was extracted using the RNeasy Plant Mini Kit from Qiagen, following the manufacturer's instructions including the optional step of on column DNase treatment using RNase free DNase (Qiagen). Purity and concentration of samples was checked using a NanoDrop™ measurement of the A_{260}:A_{280} ratio.

PCR was used to check for DNA contamination of RNA. Two PCRs were simultaneously run, one using the newly extracted RNA as a template and one using gDNA (**3.2.6.1**). Colony PCR conditions were used (**2.3.12**) using Xho1_3104RNAi_F and Kpn1_3104RNAi_R primers. PCRs were run for 40 cycles to ensure no DNA was present in RNA samples. cDNA was synthesised as per **2.3.3**. Equal amounts of RNA were added to each cDNA synthesis reaction.

SQRT-PCR allowed comparison of expression of RNAi construct in each plant line. A control PCR using housekeeping gene Elongation Factor- 1 alpha (EF-1 α) allowed comparison of samples to check presence of equal amounts of cDNA. To ensure there was no oversaturation of PCR product, the amplification of EF-1 α was carried out using 25 cycles. The PCR using RNAi-insert specific primers was carried out for 31 cycles. To increase quantification power of this experiment the agarose gel was imaged using a G:Box F3 gel imager (Syngene). This allowed quantification of gel band images using the associated GeneTools software from Syngene.

After semi-quantification of expression, 20 cuttings were taken from each mother plant. Four plant lines were used, two expressing the construct at the highest levels, one expressing the construct at the lowest and one that expressed the construct at a level

between the two. These lines would be inoculated alongside a plant line transformed with empty *Agrobacterium*. Cuttings were taken, the wounded end was dipped in rooting powder and placed into a peat pellet (Jiffy). Cuttings were left well-watered for 2 weeks until roots appeared from the base of the pod.

3.2.8 PCN infections of RNAi plants

Plant inoculations were carried out using *G. rostochiensis* cysts from a 2012 population. 20 cysts of uniform shape and size were handpicked per plant to be infected to ensure that no empty cysts were present (2000 cysts in total).

Rootrainers™ (https://www.rootrainers.co.uk/rootrainers/) were filled with 50:50 autoclaved sand:loam and watered. A hole was formed in the sand:loam using a 5 mL pipette tip dipped in up to the 3 mL mark. 20 cysts were dropped into the hole using a clean 5 mL pipette tip as a funnel. The hole was filled in and a peat pod (Jiffy) containing the growing cutting was placed on top of the sand:loam mix, allowing roots to grow through the Rootrainer™. Before planting each pod, the root system was scored based on visibility of roots.

Seven weeks after inoculation the Rootrainers™ were opened and females present on the surface of the roots were counted.

3.2.9 Quantitative PCR (qPCR) in *G. rostochiensis*

Levels of RNAi in *G. rostochiensis* females grown on the transgenic Desiree lines were quantified using qPCR.

Seven weeks after inoculation, 5 females were removed from the roots of 3 daughter plants per plant line and frozen in liquid nitrogen. Nematode RNA was extracted using an RNeasy Plant Micro kit (Qiagen) using the manufacturer's protocol.

cDNA synthesis was done as described in section **2.3.3**., with 80 ng of RNA converted into cDNA for each sample. A standard PCR was carried out to confirm that there was no amplification from RNA using the 3104RNAi_F/R primer pair. Similarly, a PCR was

carried out on newly synthesised cDNA to check that the primer pairs amplified the correct sized fragment.

qPCR reactions consisted of 2 µL template cDNA, 6 µL UV treated SDW, 1 µL of both 3104RNAi_F and 3104RNAi_R primers and 10 µL 2x LightCycler FastStart DNA master mix SYBR green I (Roche). For each sample an additional reaction was carried out using primers for the housekeeping Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene to confirm that cDNA quantity did not significantly vary between samples. After an initial denaturing step at 95 °C for 10 minutes, cycling conditions repeated 40 cycles of denaturing at 95 °C for 10 seconds, annealing at 54 °C for 10 seconds and elongating at 72 °C for 30 seconds. qPCR was carried out using an LC480 LightCycler (Roche) in LightCycler 480 multiwell 96 white plates.

qPCR data was normalised against the housekeeping GAPDH gene and fold change was represented as ΔCT. This was calculated by subtracting the CT value (threshold cycle) for GAPDH amplification from the sample from the CT value of 3104RNAi_F/R amplification from the same sample.

3.2.10 Extracting cysts from plant roots

Ten weeks after inoculation, above ground plant material was removed and the sand:loam mix containing root systems was left to dry completely. Once dry, the peat pod was removed and the remaining soil and roots were broken up into a large bucket. The bucket was filled with water and left to stand. Cysts floated to the top of the water where they were poured through a 710 µm sieve and collected on a 250 µm sieve. Floating of cysts in the bucket was repeated 3 times before the contents of the 250 µm sieve were collected onto a fine mesh and left to dry for 48 hours.

Once dry, cysts were transferred from the mesh to an A3 sheet of printer paper. Cysts were rolled from the paper onto a separate sheet before being collected in a watch glass. Cysts were separated from any other debris by hand under a microscope.

3.2.11 Analysis of RNAi phenotypes - Nematode and eggshell morphology

Due to the unknown consequences of knocking down the eggshell annexin a variety of basic morphological tests were carried out.

Twenty cysts from populations collected from either knock-down or negative control plants were imaged under a dissecting microscope. ImageJ software was used to measure the diameter of cysts (Schneider *et al.*, 2012). Statistical significance between sizes of cyst populations was calculated using ANOVA tests assuming significance of $P<0.05$.

Similarly, eggshells removed from cysts were imaged under a stage microscope and the length and width of the eggshells was measured using ImageJ. Statistical T-tests were carried out to identify significant differences between negative control and knock-down eggshells. Statistical significance was assumed if $P<0.05$.

3.2.12 Analysis of RNAi phenotypes - Hatching assays

Hatching assays were used to identify differences in responses to host root diffusates after eggshell annexin knock-down. Hatching of juveniles was tested pre-diapause.

Fifteen cysts from each nematode population were placed into 500 µL SDW and left at room temperature overnight. Three cysts were placed into a single well of a 24-well cell culture plate, replicated 5 times. 500 µL of freshly produced TRD was aliquoted into each well containing cysts, the plate was sealed, wrapped in tin foil and placed at 18 C for a week. Juveniles present in the wells were counted at weekly intervals. After counting, the old TRD was replaced with fresh TRD and the plate was returned to 18 °C. Hatched juveniles were counted over a 4 week period.

After 4 weeks of exposure to TRD, cysts were removed from the plate and dissected to release the eggs. The number of eggs per well was counted allowing for a percentage hatch to be calculated. Accuracy was increased by imaging the number of eggs per well under a dissecting microscope. Eggs were then counted manually using the cell counter plugin for ImageJ software. Significance between the average percentage of juveniles

hatched in RNAi or negative control populations was calculated using T-tests. ANOVA tests were used to test significance of percentage hatch within replicates of the same cyst population. Statistical significance was assumed if $P<0.05$ for both ANOVA and T-tests.

3.2.13 Predicting protein structure

Protein sequences were submitted to the web based Phyre2 server which predicts protein structure from amino acid sequences based upon solved crystal structures of orthologous proteins (Kelley *et al.*, 2015). Orthologous protein structures were downloaded from the RCSB protein databank (https://www.rcsb.org/). Protein structures were handled in PyMol 2.3.0.

Following structural prediction, the protein model was submitted to 3D LigandSite. Similarly to Phyre2, 3D ligand site uses known ligand binding sites on proteins with solved crystal structure to allow identification of ligand sites in the predicted protein structure (Wass *et al.*, 2010).

3.3 Results

3.3.1 Identifying eggshell proteins by mass spectrometry

Proteins were extracted from eggshell samples using solvents. Polar solvents promote disruption of chitin layers and so are sufficient for extraction of proteins that could be held within this chitin rich layer (**3.2.2**).

Peptide hits from mass spectrometry data were collected from across 5 separate extractions. Hits occurring in 3 or more experiments with more than 2 unique peptide matches per protein are shown in **table 3.2**.

Table 3.2 – Proteins identified from eggshell protein extractions.
Proteins in the table appeared in 3 or more extractions. The peptide match confidence limit was set to 95% and the protein match confidence limits were set to 99%. The average number of unique peptides has been rounded to the nearest whole number.

Gene name	Protein function	Molecular weight (kDa)	Average number of unique peptides identified	Average percentage coverage (%)
GROS_g01373	Chondroitin proteoglycan	31	6	16.3
GROS_g13541	Flavin adenine dinucleotide binding domain containing protein	96	6	11.0
GROS_g14162	Flavin adenine dinucleotide binding domain containing protein	71	13	19.6
GROS_g03104	Annexin	37	6	19.8
GROS_g06666	Chondroitin proteoglycan	35	10	23.2
GROS_g00671	Actin	55	8	21.5
GROS_g13621	Glycosyl hydrolase family 30	55	5	8.0
GROS_g11707	Lipase	27	7	30.5
GROS_g06834	Tyrosinase copper binding protein	86	6	13.3
GROS_g02583	Hypothetical protein	23	6	35.3
GROS_g06701	Transthyretin-like	17	6	38.3
GROS_g02714	Vitellogenin	226	8	5.4

GROS_g12524	Glycosyl hydrolase family 27	48	4	11.5
GROS_g06331	Hypothetical protein	16	4	21.8
GROS_g12038	Neprilysin	90	4	5.5
GROS_g08138	Hypothetical protein	12	3	27.7
GROS_g03996	Strictosidine synthase	61	5	10.0
GROS_g14202	Fatty acid and retinol binding protein	22	5	21.8
GROS_g11876	Haem peroxidase	76	5	11.3
GROS_g10871	Superoxide dismutase	23	4	21.0

Identification of proteins such as chondroitin proteoglycans provides reassurance that the extraction method used was capable of extracting genuine eggshell proteins, as these orthologous proteins have been previously discovered in *C. elegans* eggshells. Identification of a glycosyl hydrolase and a lipase could reflect a function for the use of eggshell based enzymes in degradation of the eggshell to aid eclosion. The association of enzymes and PCN eggshells is discussed in **Chapter 4 – Identification and characterisation of PCN eggshell lipids**. As a protein associated with egg yolk, identification of vitellogenin is unsurprising here. However, vitellogenin is a protein that is found within the egg but is unlikely to be a part of the eggshell itself. Similarly, additional proteins identified here (Fatty acid and retinol binding protein, transthyretins and superoxide dismutase) represent proteins exported to the surface of the juvenile nematode and were likely trapped within the eggshell during eggshell collection and therefore cross-contamination.

The annexin GROS_g03104 was chosen for further study due to the predicted function of annexins as calcium dependent lipid binding proteins and the known important role for calcium within the hatching cascade of PCN. A previous study on the PCN annexin (Gp-nex, orthologue of GROS_g03104) suggested that this protein has effector like properties and is secreted into the plant host (Fioretti *et al.*, 2001). However, the expression pattern for this gene (**Figure 3.2**) shows no upregulation at parasitic life-stages. Additionally, the polyclonal antibody produced for this work was not specific to only Gp-nex and is likely to bind to other annexins. It is likely that the antisera could detect highly conserved domains found in all annexins. Interestingly, although not

shown by the authors, it is mentioned that the antisera had a high affinity for eggs and adult females.

Expression data from previously acquired RNA-seq data shows little to zero expression throughout the parasitic life stages of both *G. rostochiensis* and *G pallida* but show that the gene encoding this annexin is highly upregulated at the nematode life stage containing the eggs (Cotton *et al.*, 2014; Eves-van den Akker *et al.*, 2016). New RNA-seq data from the root knot nematode (RKN) *M. incognita* (E. Danchin, 2019, personal communication, 15/01/2019) shows that the three RKN sequences most similar to GROS_g03104 are not up-regulated at egg-producing life stages (**Figure 3.2**). Given that RKN hatch via a different mechanism to PCN and do not link their life cycle to a specific host through detection of host sepcific root diffusates, this provides further evidence that the annexin identified here may play a role in control of hatching in PCN.

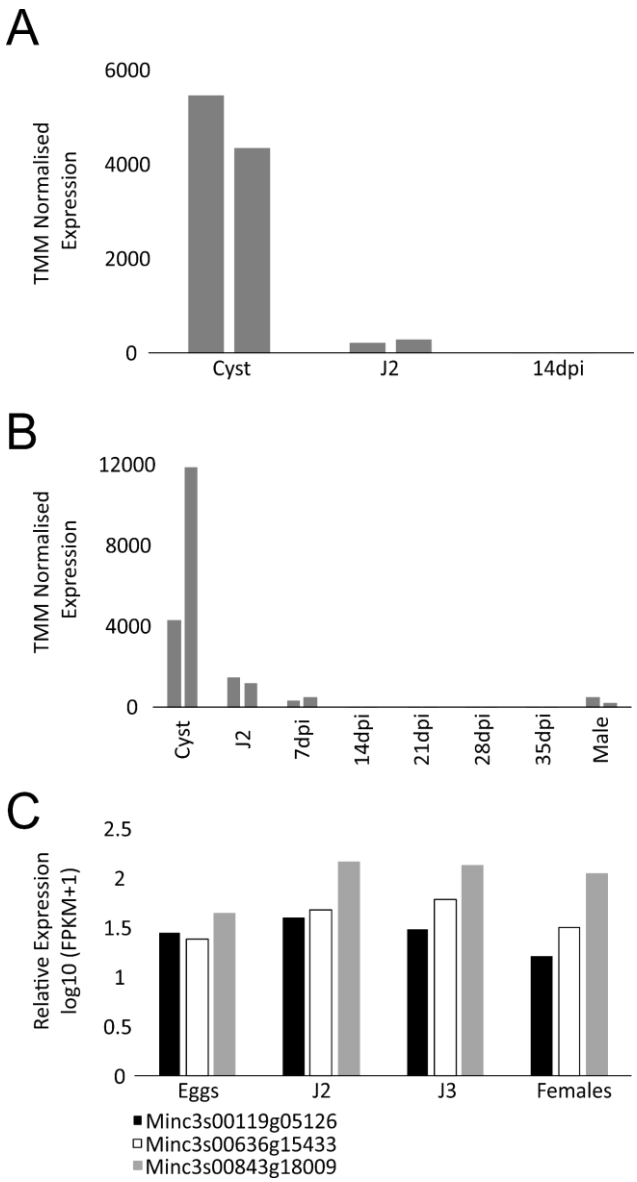


Figure 3.2 – Eggshell annexin life-stage specific expression

A) GROS_g03104 expression in *G. rostochiensis*. **B)** GPLIN_000171600 (ortholog of GROS_g03104) expression in *G. pallida*. Both **A** and **B** are generated from RNAseq data where relative expression was calculated with two replicates per given life stage. Dpi = days post-infection. **C)** Expression data for three of the closest matches to GROS_g03104 in *M. incognita*.

Other PCN annexins were identified using BLAST against *G. rostochiensis* transcriptome and genome data. The annexin family is abundant in *G. rostochiensis* with many of the proteins sharing a highly conserved sequence. **Figure 3.3** shows the peptides identified in GROS_g03104 by mass spectrometry (full alignment in **7.4 Alignment of annexins from *G. rostochiensis***). A unique peptide -FFGIGNLGI- was identified in eggshell protein extractions and is specific to this annexin. This confirms that the annexin identified in

these protein extractions was GROS_g03104. When comparing GROS_g03104 homologues between a variety of parasitic and free-living nematodes it became apparent that this unique peptide is specific to the genus *Globodera* (full alignment in **7.5 GROS_g03104 alignment against other nematode annexin orthologues**).

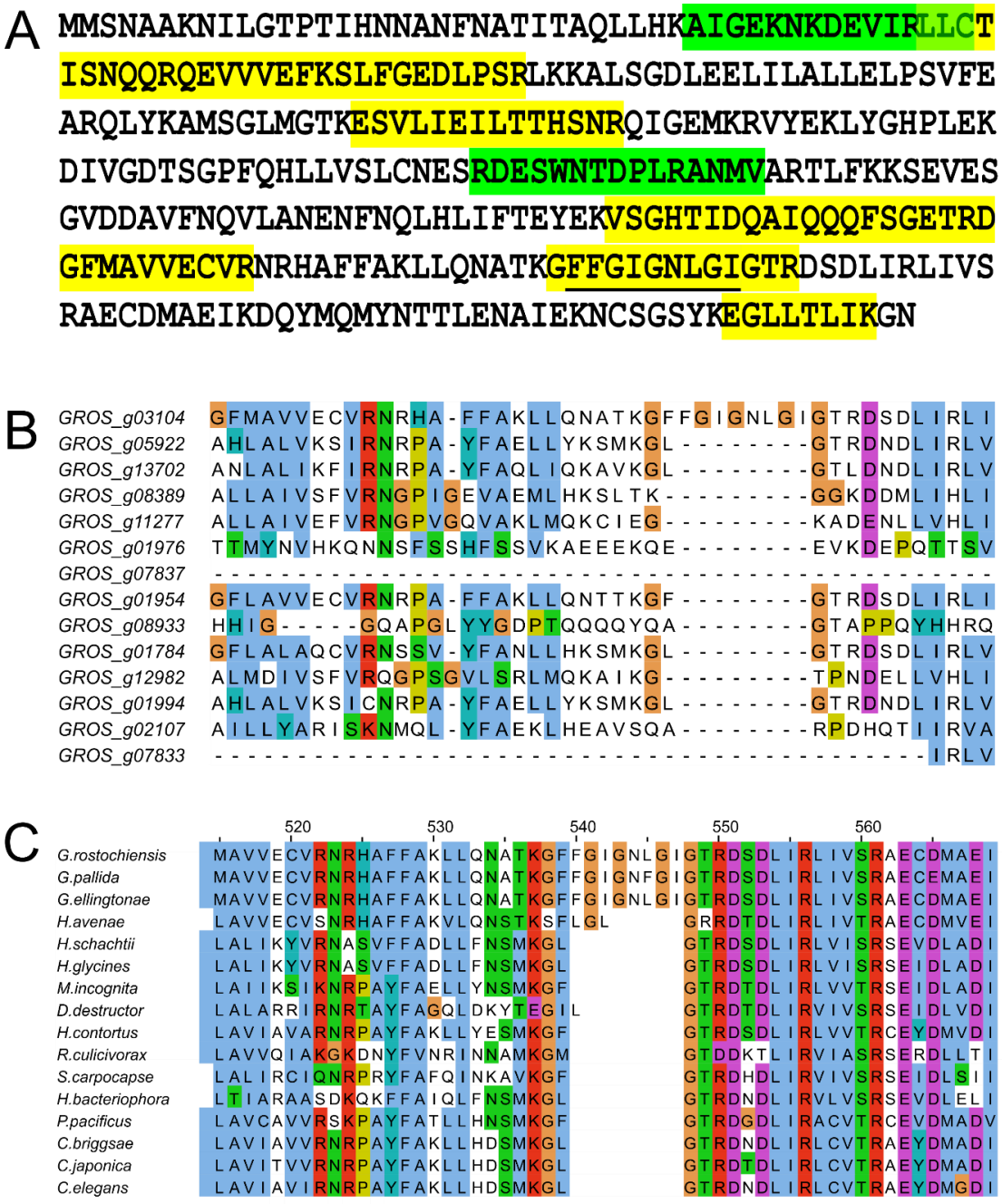


Figure 3.3 – Annexin sequence identification and alignment

A) Protein sequence for GROS_g03104 with peptides identified by mass spectrometry highlighted in yellow. Peptides used for antibody production are highlighted in green. **B)** section of sequence alignment comparing GROS_g03104 to other annexins identified using BLAST against *G. rostochiensis* proteome data. The unique sequence FFGIGNLGI visible here (underlined in **(A)**) was an identified peptide from the mass spectrometry data. **C)** section of sequence alignment comparing GROS_g03104 orthologues between various nematode species. The same unique peptide appears specific to the *Globodera* genus. Amino acid colouring in both **B** and **C** is a representation of the ClustalX standard colour scheme.

3.3.2 Localisation of GROS_g03104 to the eggshell

The eggshell is an extracellular entity meaning techniques such as *in situ* hybridisation to localise gene expression are not appropriate. This makes identifying sites of expression of eggshell protein candidates more of a challenge. Therefore antibodies raised against the recombinant annexin were used to localise this protein.

Recombinant annexin was obtained as described in section **3.2.3** and antisera raised against peptides synthesised by Eurogentec (**3.2.4**) were tested using a western blot (**Figure 3.4**). Two peptides were synthesised and used for antiserum production. A1 (RDESWNTDPLRANMV) was raised against an immunogenic peptide found between two annexin domains, A2 (KAIGEKNKDEVIRLLC) was synthesised against a peptide within an annexin domain. Although both antisera were specific to GROS_g03104, being outside the highly conserved annexin domain meant A1 had the potential to be more specialised for GROS_g03104. Antibody A1 bound to recombinant annexin whereas A2 did not.
Once working concentrations of GROS_g03104 antibodies had been identified (1:1000) they were used to localise GROS_g03104 to the PCN eggshell.

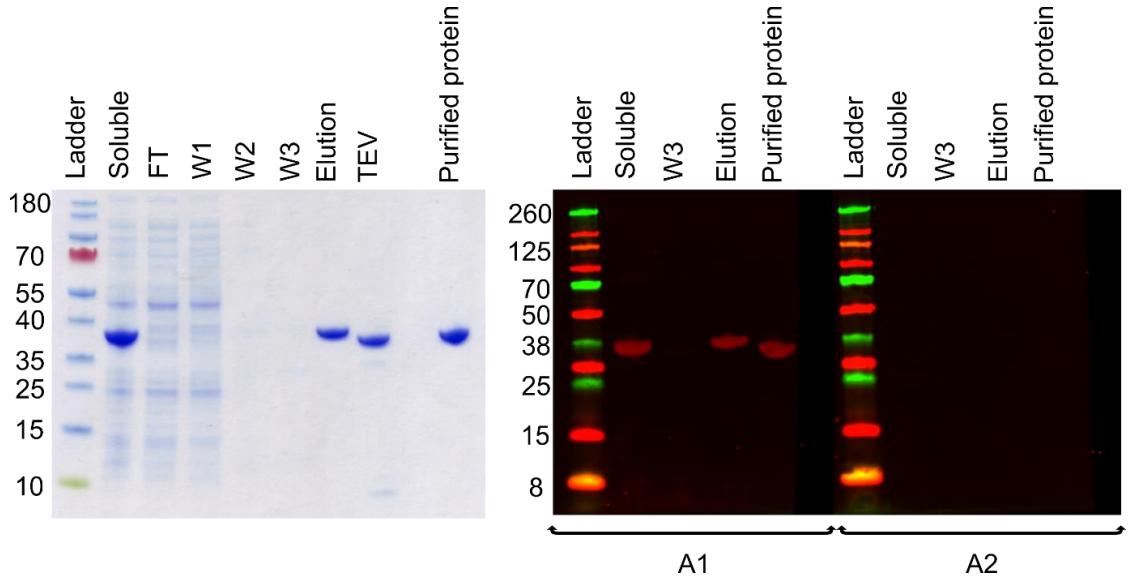


Figure 3.4 – Recombinant annexin and antibody testing.

Left) SDS-PAGE showing expression and purification of recombinant GROS_g03104. Recombinant protein was purified with a 6xHIS tag which was later cleaved using TEV. Right) western blots showing binding of anti-peptides A1 and A2 to recombinant GROS_g03104. Anti-peptide A2 did not show binding to recombinant GROS_g03104. W = column wash, FT = column flow through

Auto-fluorescence is a problem when imaging nematode eggshells. For this reason an Alexa secondary antibody (AlexaPlus) was used (Invitrogen). These fluorophores are reportedly 4.2x brighter than their predecessors making them ideal for imaging when computational fluorescence gain levels are set to low levels, as was required here. Eggshells that had not been exposed to any antibodies were first observed under the microscope. The microscope was then adjusted to reduce the computational gain to a level where eggshell auto-fluorescence at 488 nm was only just visible effectively zeroing the microscope to allow any additional fluorescence, resulting from secondary antibodies (AlexaPlus 488), to be seen. Auto-fluorescence levels were still imaged at 516 nm showing consistent levels between all eggshells whether incubated in pre-immune or anti-serum. The same microscope settings were used for imaging all eggshells (**3.2.5**). A clear increase in fluorescence was visible at 488 nm reflecting binding of primary antibodies (A1) and the secondary antibody, therefore, localising GROS_g03104 to the PCN eggshell (Figure 3.5).

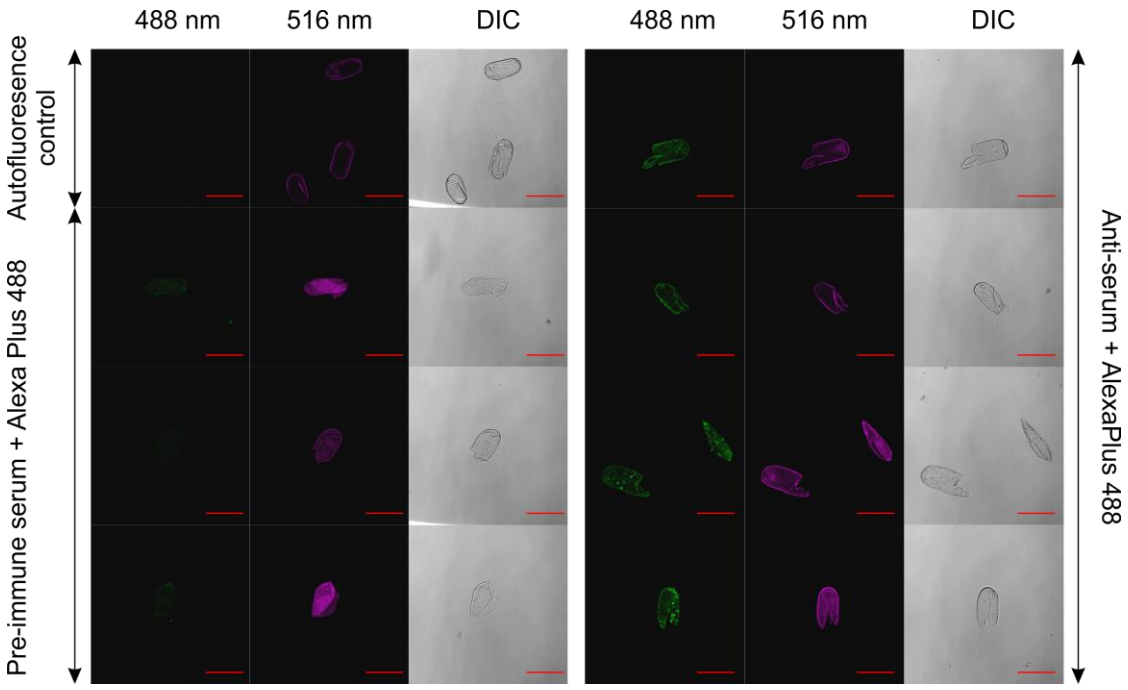


Figure 3.5 – Immunolocalisation of GROS_g03104

Left) Auto-fluorescence control used to minimise computational gain at 488 nm, reducing auto-fluorescence in this channel to just visible. Pre-immune serum shows no binding of rabbit antibodies and no additional fluorescence from the secondary antibody. **Right)** Binding of anti-serum, and therefore the secondary antibody, shows localisation of GROS_g03104 to the eggshell. The brightness and contrast for all the images has been increased 20% and 30% respectively for printing. Scale bars represent 100 µm

3.3.3 Structural projection supporting predicted GROS_g03104 function

Due to the highly-conserved structures of annexins a model for GROS_g03104 can be inferred based on solved crystal structures of orthologous annexins (**Figure 3.6**). The Phyre2 server modelled 98% of GROS_g03104 at >90% accuracy. This predicted structure shares a large amount of structural similarity compared to reference structures. Most structural alignments were completed against bovine annexin VI due to sequence similarity and 3D structure availability (Avila-Sakar *et al.*, 1998). Bovine annexin VI is a dimer. The sequence for GROS_g03104 has only been compared to one of the strands in this dimer.

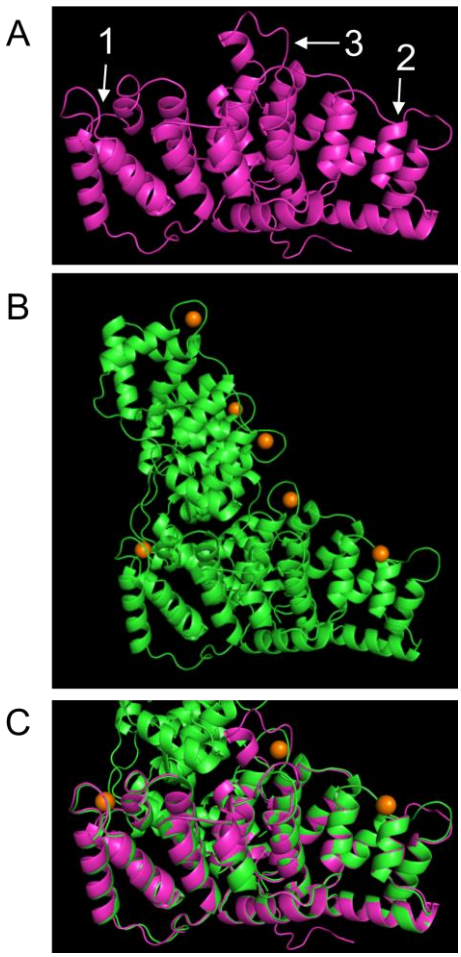


Figure 3.6 – Predicted structure of GROS_g03104

A) Predicted structure for GROS_g03104 based on sequence similarity of annexins with solved 3D structures. Sites 1, 2 and 3 highlighted here are predicted calcium binding sites discussed in later figures. **B)** Solved 3D structure of bovine annexin VI with calcium ions highlighted in orange. **C)** Overlay of GROS_g03104 prediction and bovine annexin showing large amounts of homology between the two sequences

From this structure, calcium binding sites were also predicted by 3D LigandSite software. Sites predicted to occur less than 5 times were discarded leaving 5 further sites to be investigated. Binding sites were compared to calcium binding sites for similarities in orthologous proteins. The first of these predicted sites is suggested to form a pocket between glutamate, serine, glycine and glutamine at positions 191, 192, 193 and 233 respectively. Bovine annexin VI shows a similar structure at these positions and is known to have calcium binding capabilities at this location (Figure 3.7). The location on the whole GROS_g03104 structure prediction is marked as number 1 on Figure 3.6 A.

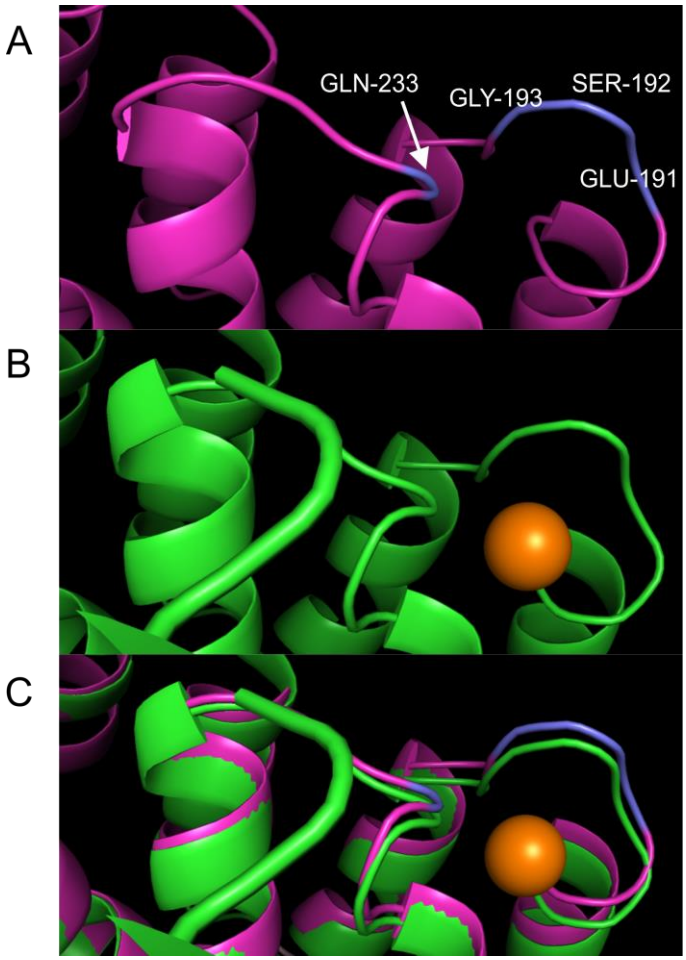


Figure 3.7 – GROS_g03104 predicted calcium binding site 1
A) Predicted binding site for calcium in GROS_g03104. Interacting amino acids have been highlighted in blue. **B)** The same calcium binding site in bovine annexin VI. **C)** Overlay of A and B showing structural similarity between this position in the two sequences. Calcium ions are highlighted in orange.

The next predicted calcium binding site is at leucine 77. There is also a predicted calcium binding site at the same leucine on bovine annexin VI. Calcium has a coordination number of 6 meaning that the calcium ion preferably holds on to 6 neighbouring atoms in a complex (Katz *et al.,* 1996). Therefore, it seems unlikely that calcium would interact with a single amino acid in an exterior position (**Figure 3.8**). The location on the whole GROS_g03104 structure prediction is marked as number 2 on **Figure 3.6 A**. Two more similar single amino acid calcium binding locations were suggested at lysine 37 and arginine 239. There was no calcium binding site at these positions in modelled annexins and so these sites were omitted.

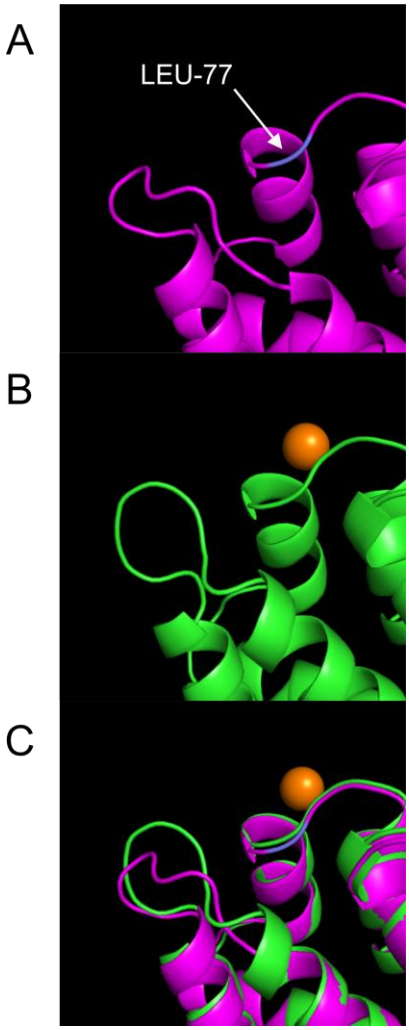


Figure 3.8 – GROS_g03104 predicted calcium binding site 2
A) Predicted binding site for calcium in GROS_g03104. The interacting amino acid has been highlighted in blue. **B)** The same calcium binding site in bovine annexin VI. **C)** Overlay of A and B showing structural similarity between this position in the two sequences. Calcium ions are highlighted in orange.

The final calcium binding site is predicted at asparagine 316 in the GROS_g03104 peptide sequence. Calcium also binds to this location in bovine annexin VI. However, in bovine annexin VI there is a clear pocket for calcium formed, as with the first calcium binding site. In GROS_g03104, this pocket appears expanded (Figure 3.9). From N to C terminus the peptide sequence disforming the calcium binding site is N-FFGIGNLGI-C, the same motif that is seemingly unique to the *Globodera* genus as identified in Figure 3.3 C. The location on the whole GROS_g03104 structure prediction is marked as number 3 on Figure 3.6 A.

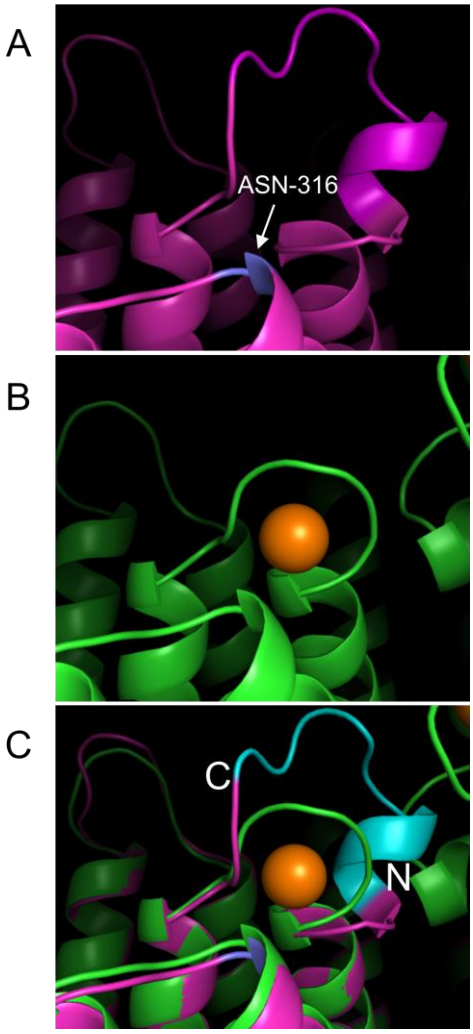


Figure 3.9 – GROS_g03104 predicted calcium binding site 3

A) Predicted binding site for calcium in GROS_g03104. The interacting amino acid has been highlighted in blue. **B)** The same calcium binding site in bovine annexin VI. **C)** Overlay of A and B showing structural difference between this position in the two sequences. The *Globodera* specific motif found in GROS_g03104 has been highlighted in light blue. Calcium ions are highlighted in orange.

Due to the differing structure at predicted calcium binding site 3 in GROS_g03104 it is unclear whether this area still functions in calcium binding.

Another known annexin crystal structure comes from the fresh-water polyp *Hydra vulgaris*. This annexin (XII) forms a homo-hexamer resulting in a pore-like structure (Cartailler *et al.*, 2000) (**Figure 3.10 A**).

Aligning strands A and strand E of the *H. vulgaris* annexin XII with two copies of the GROS_g03104 structural prediction shows that the interacting surface between the two strands of GROS_g03104 contains the unique *Globodera* motif (N-FFGIGNLGI-C). Converting the structure in this region to a space filling model shows potential for this region to act as an embracing arm between two GROS_G03104 monomers (**Figure 3.10 C**).

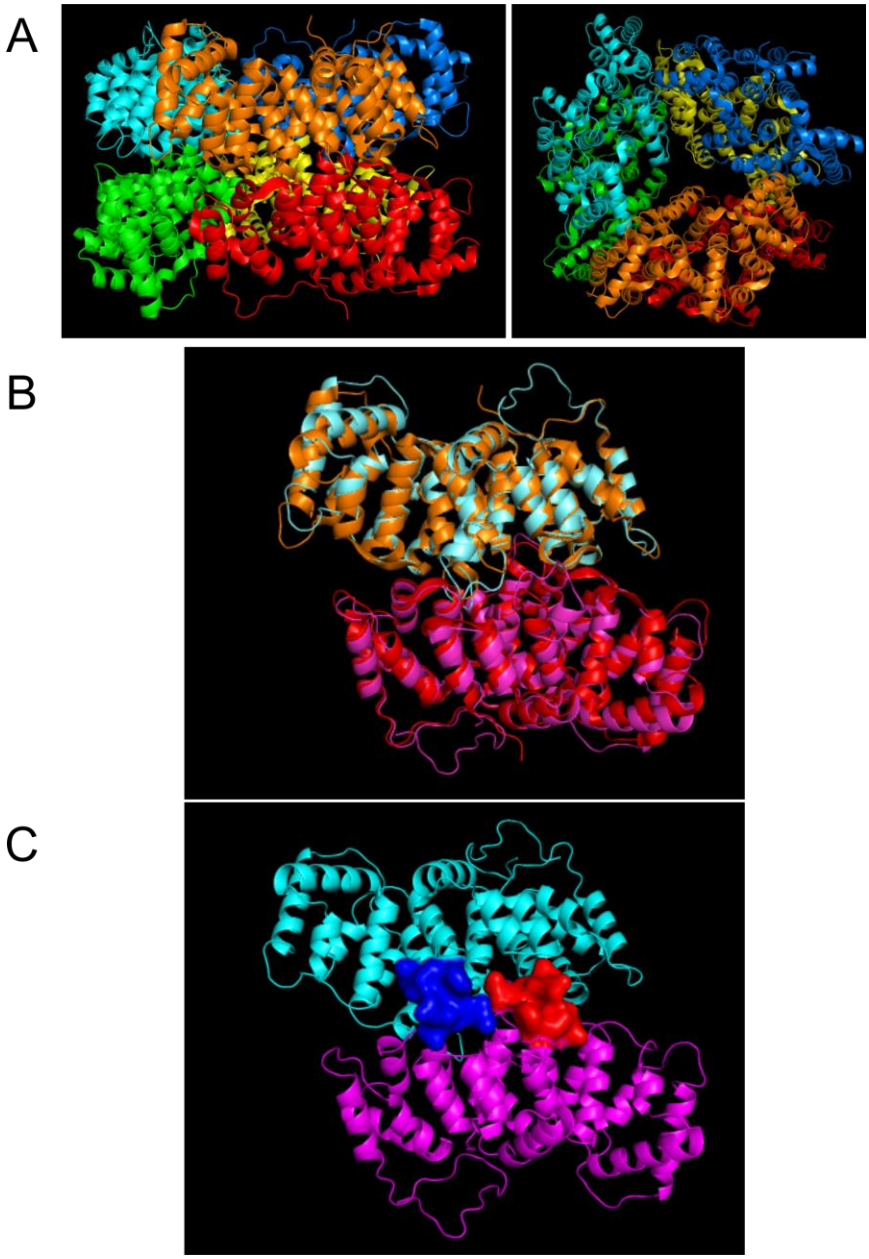


Figure 3.10 – GROS_g03104 multimeric potential

A) *H. vulgaris* annexin XII homo hexamer in different orientations to show pore-like structure. Monomeric units have been coloured differently for clarification. **B)** Overlay of *H. vulgaris* annexin XII monomer strands A and E with two separate units of the predicted structure for GROS_g03104. There is a large amount of overlap between the two structures. **C)** Monomers of GROS_g03104 previously aligned to *H. vulgaris* annexin XII. The unique motif associated with the annexin from the *Globodera* genus have been highlighted in red or blue. Using a space filling model shows the potential for these regions to interact between GROS_g03104 monomers.

3.3.4 Eggshell annexin RNAi

The function of the eggshell annexin was investigated using RNAi. This was achieved by creating transgenic potato plants ('Desiree') that express a short hairpin of RNA which, when consumed by female PCN, would target the GROS_g03104 mRNA.

Cuttings from WT Desiree were exposed to *Agrobacterium* containing the modified pART-27 vector that contained GROS_g03104-specific forward and reverse regions, capable of making an RNA hairpin when expressed (**3.2.6**). Presence of the transgenic construct was tested using semi-quantitative PCR (**Figure 3.11**). Leaf material was excised from developing shoots from 5 separate plant lines (2, 6, 7, 14 & 20) and a control line (EV) that had been transformed by *Agrobacterium* in the same way but the *Agrobacterium* contained no pART-27 vector. This line was used to check that any changes seen in the next generation of nematodes was due to the RNA hairpin knocking-down annexin and not because of the plant transformation process. RNA was extracted from leaf material and tested for the presence of contaminating DNA by PCR. RNA was then quantified and the same amount of RNA for each sample was converted into cDNA. Two PCRs were then carried out on cDNA using separate primer sets. The first using primers for a control gene showed similar amplification in all samples. The second reaction, using primers specific to the GROS_g03104 only allowed amplification if the DNA encoding the hairpin was present. Although all reactions were carried out to the same number of cycles, more amplification was visible in plant lines expressing the construct to a higher level. This semi-quantitative method was taken one step closer towards quantification by using a UV gel imager that allowed for quantification luminescence of DNA bands (**Figure 3.11 C**).

RNAi lines #2 and #7 showed the highest expression of the RNA interfering hairpin. This was followed by plant line #20, #14 and #6 expressing the construct at lower levels. The negative control line expressing the unmodified pART-27 vector showed no expression of the RNA interfering hairpin.

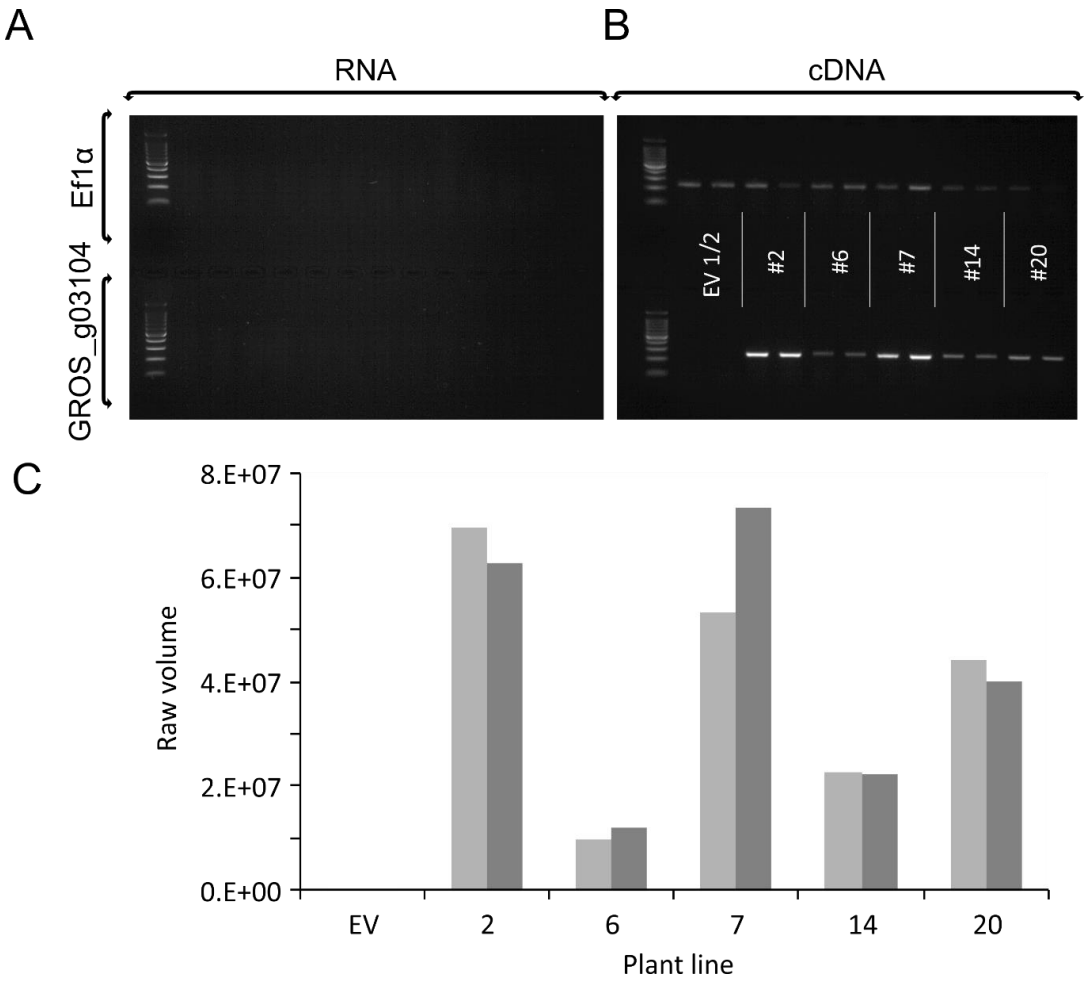


Figure 3.11 – Semi-quantitative PCR of transgenic Desiree cDNA

A) Amplification of DNA present in RNA extractions from transgenic Desiree leaves. No gel bands were visible after 40 cycles suggesting there was no DNA in the RNA extractions. **B)** Semi-quantitative PCR of cDNA created from RNA extractions. Top lane shows near equal amplification of the control gene EF-1 α after 25 cycles, reflecting equal amounts of cDNA per sample. The bottom lane shows a lack of amplification in response to GROS_g03104 specific RNAi hairpin primers in the negative control (EV) plant line. Varying levels of amplification are displayed by the brightness of gel bands in RNAi lines. **C)** Quantification of luminescence from gel bands read by GeneTools software. Two replicates were carried out for each plant line.

From this experiment, lines EV, 2, 6, 7 and 20 were grown further and each plant was split into 20 cuttings. Cuttings were transferred to Rootrainer™ pots containing 50:50 sand:loam mixture. 20 hand-picked cysts of equal size were added to each section of the Rootrainer™. Although it could not be guaranteed that the same number of juveniles would infect each plant, using cysts reflects natural infection conditions better. Nematodes were left to infect the host plant for 7 weeks.

5 females were removed from roots of 3 random lines of the 20 cuttings to be tested for levels of GROS_g03104 knock-down. RNA was extracted from the female nematodes and converted into cDNA which was used to assess levels of eggshell annexin knock-down compared to expression of the unaffected control gene, GAPDH, using qPCR (**Figure 3.12**).

Levels of annexin knock-down were not consistent within cuttings of the same line. Additionally, increased presence of the RNAi hairpin in the transgenic line did not correlate with increased knock-down of the annexin. Due to the variation in levels of knock-down, only cyst lines that had their knock-down measured by qPCR were used from this point forwards. Cyst lines were assigned a number. The value before the decimal corresponds to the plant line that cysts were collected from. The number after the decimal corresponds to which cutting within the line the cysts were collected from. For example, cyst population 2.5 was raised on plant line #2 and taken from plant cutting #5 within that line.

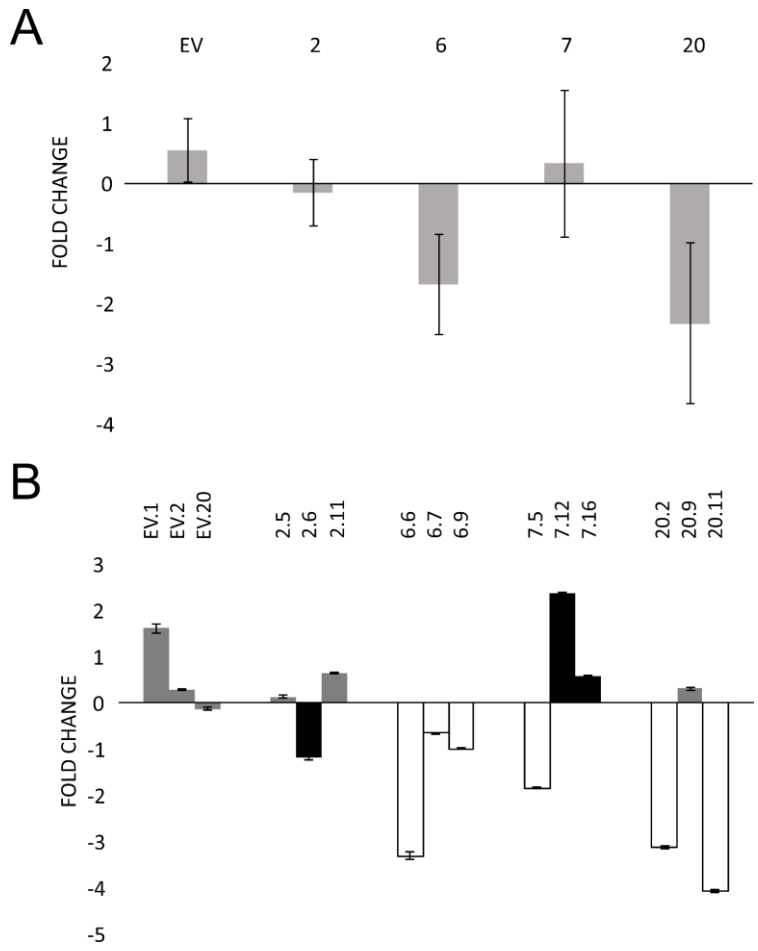


Figure 3.12 – qPCR analysis of annexin knock-down in PCN

A) Average level of knock-down per cyst population compared to control gene GAPDH. **B)** Knock-down of eggshell annexin within cyst populations raised on the same plant line. Bars in grey were used as no knock-down control, bars in white were used as RNAi test samples, bars in black were not used from this point forward. Error bars represent standard error.

Females present on the roots of transgenic plants were counted 7 weeks after planting. An ANOVA between samples suggests that there is a significant difference in cyst numbers between plant lines (P=0.0495) (Figure 3.13 A). However, this is caused by the reduced number of cysts counted on cuttings from plant line #7. This is likely due to the badly developed root systems in cuttings from this plant line and not as a result of annexin knock-down (7.6 Scoring root systems of transgenic Desiree cuttings).

Twenty cysts per population were imaged under a dissecting microscope. The cyst diameters were measured using ImageJ software. ANOVA testing was used to show that

there was no significant difference between sizes of cysts taken from the negative control plant lines (EV) (P=0.514). Similarly, there was no significant difference in cyst sizes in cysts from line 6.6, 20.11 and 20.2 when compared to the negative control populations of cysts (P=0.696 , P=0.367 , P=0.078 respectively) (**Figure 3.13 B**).

Dimensions of eggs removed from cysts were also measured using ImageJ software. Not all lines were tested due to limited availability of samples. However, measurements showed that eggshells from both RNAi lines 6.6 and 20.11 were significantly narrower than eggshells from the negative control line EV.20 (P=0.024 and P=0.007). Eggshells from test line 6.6 proved significantly longer than eggshells from the negative control line (P=0.029). However, eggshells from RNAi line 20.11 showed no significant difference in length compared to negative control eggs (P=0.737) (**Figure 3.13 C**). If assuming all eggshells are cylinders then the volume of the egg can also be calculated using $V=\pi r^2 h$. The volume of eggs from population 20.11 was not significantly different to line EV.20 (P=0.0618). However, the volume of eggs from population 6.6 was significantly less than that of the negative control line (P=0.0051) with eggs having on average 14.95% less volume (**data on accompanying CD**).

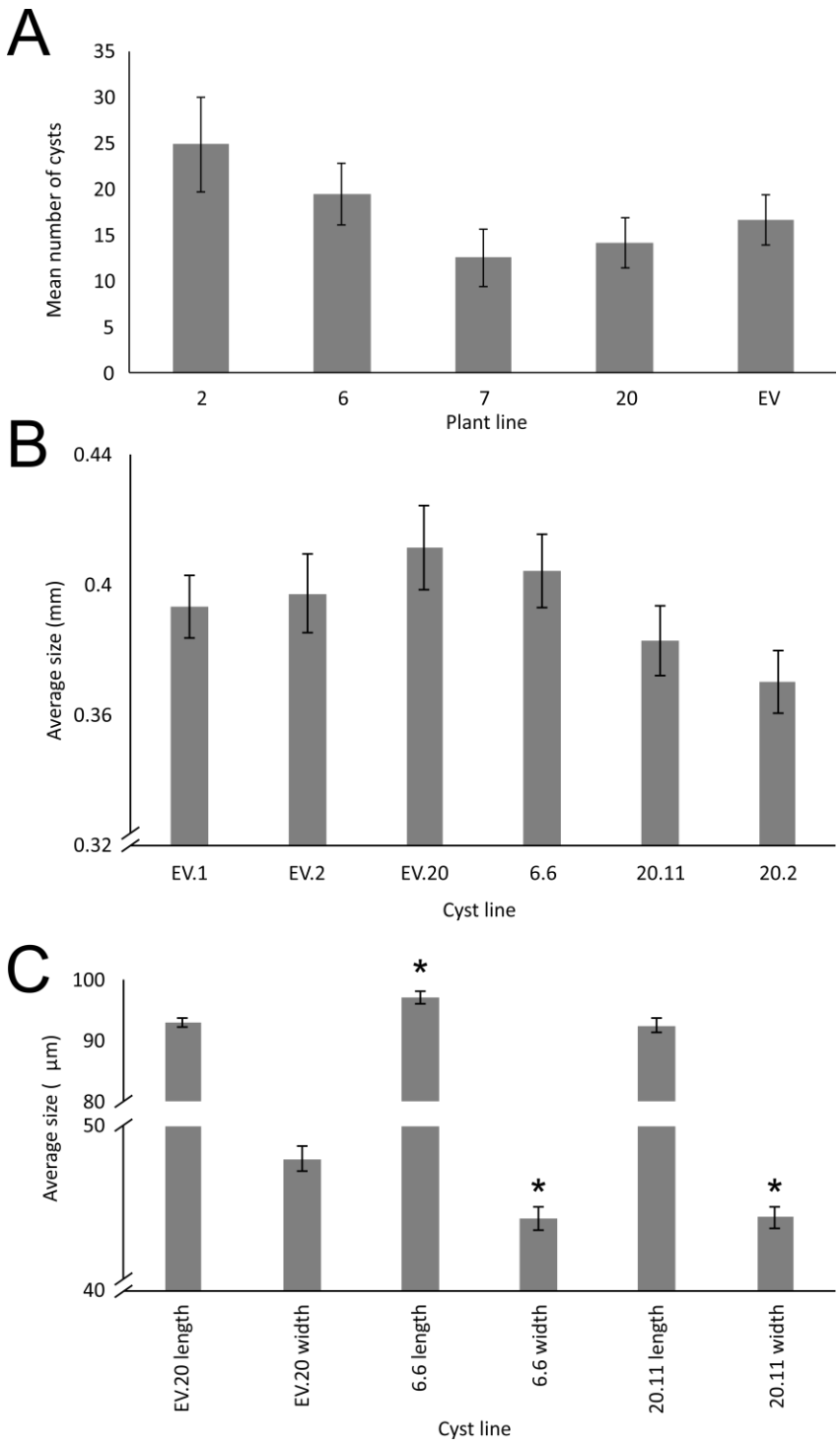


Figure 3.13 – RNAi and negative control cyst and eggshell dimensions

A) Number of females counted presenting on roots of transgenic Desiree 7 weeks after initial infection. B) Measurements of cyst diameter in populations taken from RNAi or negative control plant. No significant difference was seen between RNAi and negative control lines. N = 20 cysts C) Eggshell length and width measurments. Measured eggshells (n=32) showed that eggshells from populations with reduced availability of annexin were significantly narrower than those with no annexin knock-down. Where P<0.05 is marked with '*'. Error bars represent standard error.

When measuring eggshell sizes, differences in developing juveniles within the eggs from RNAi lines became apparent. Usually, developed PCN juveniles grow to fill the space within the eggshell. However, some juveniles from RNAi lines appeared smaller with increased space visible between the nematode cuticle and the eggshell. Similarly, a larger quantity of underdeveloped juveniles was noticeable in RNAi lines (**Figure 3.14 A & B**).

Cysts taken from either negative control, RNAi or the original population used to inoculate transgenic Desiree were cut open and the eggs were removed. Numbers of underdeveloped juveniles were counted before counting all eggs present within the cyst to give a percentage. Here, the term 'underdeveloped' was used to describe anything still within various stages of embryonic development. A t-test between cysts from the negative control line (EV) in addition to cyst populations that showed no annexin knock down (2.5 & 2.11) and cysts from the RNAi lines showed that there was a significant increase in the number of underdeveloped juveniles in cysts that showed annexin knock-down through qPCR (6.6, 20.2 & 20.11) ($P=7.03E^{-6}$). Individual t-tests between any cyst population and the negative control line (EV) showed a significantly increased number of underdeveloped juveniles (**Figure 3.14 C**).

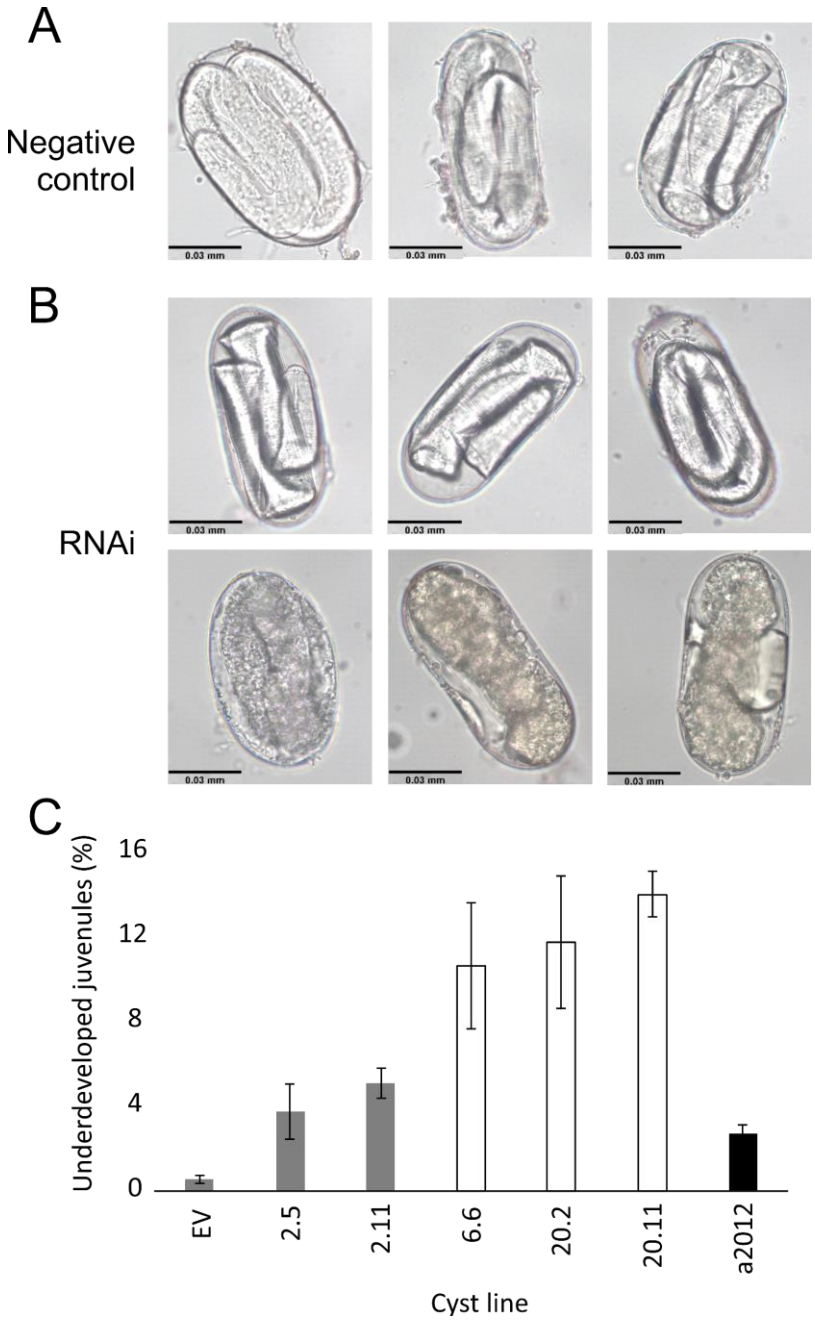


Figure 3.14 – Juvenile morphology within annexin knock-down eggs

A) Juveniles inside eggs from plant lines not expressing the RNAi hairpin. Juveniles are swollen and fill the entire eggshell. **B)** Top - juveniles inside eggs with reduced eggshell annexin availability. Juveniles appear shrunken, not filling the eggshell space. Bottom – underdeveloped juveniles found in RNAi lines. **C)** Percentage of underdeveloped juveniles found in negative control or RNAi cyst populations. Eggs from 5 cysts per population were used. There is a significant increase in underdeveloped juveniles in cysts populations that have reduced annexin availability. Cyst populations were also compared to a sample taken from the original population of cysts used to inoculate transgenic plants (a2012). All cyst populations show a significantly increased number of underdeveloped juveniles compared to the negative control line (EV). Error bars represent standard error.

The effect of the RNAi of the annexin on the PCN ability to hatch in response to TRD was tested using hatching assays. Hatching was monitored over a four-week period with cysts being exposed to fresh TRD each week. At this stage nematodes should be in diapause as they had not been exposed to a period of cold and would therefore not be expected to hatch. A significant increase in the cumulative number of juveniles hatching in annexin knock-down lines was visible after two weeks of hatching (P= 0.036, P= 0.014 and P= 0.003 in weeks 2, 3 and 4 respectively) (**Figure 3.15 A**). When not looking at the hatching data as averages it is noticeable that the three RNAi lines exhibiting the highest number of juveniles hatched were also the three lines of females showing increased annexin knock-down (**Figure 3.12 B, Figure 3.15 B**).

Combining the data from **Figure 3.14 C** and **Figure 3.15 B** allows the average percentage of juveniles hatched per cyst population to be calculated after removing the average number of underdeveloped juveniles from each replicate of the population (**Figure 3.15 C**). Correcting the data in this way presents a more representative result as the eggs that would never hatch within the experiment have been removed. Populations 6.6, 20.2 and 20.11 still show a significantly increased number of juveniles hatching after TRD exposure for 4 weeks compared to EV.2 (P=0.0495, 0.0064 and 0.0064 respectively).

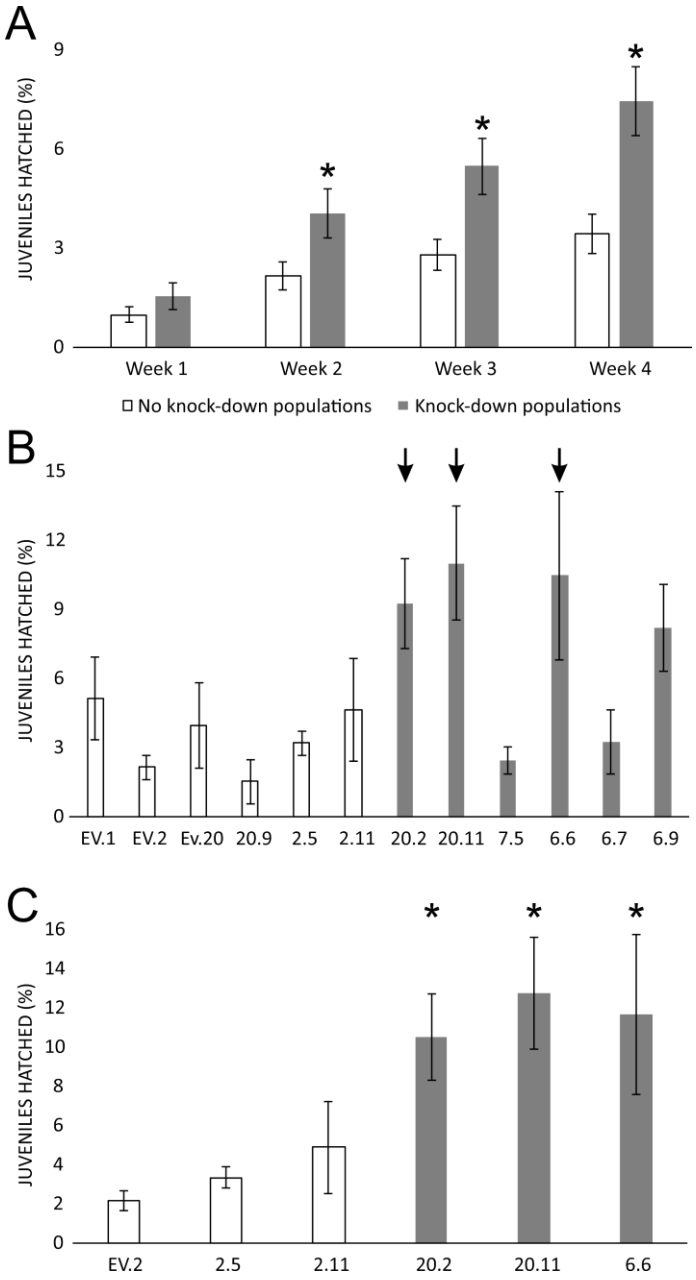


Figure 3.15 – Hatching of cyst populations after eggshell annexin knock-down.

A) Average percentage of juveniles hatched from RNAi lines or populations that showed no annexin knock down. Averages were calculated after 4 weeks of exposure to TRD. A significant increase in the percentage of juveniles hatched was visible after 2 weeks, marked with '*'. **B)** Percentage of juveniles hatched per cyst population. Bars in white showed no knock-down of the eggshell annexin whereas bars in grey showed annexin knock-down. Lines marked with an arrow showed the highest percentage of hatched juveniles, corresponding with these lines showing the largest amount of annexin knock-down as determined by qPCR. **C)** Percentage of juveniles hatched per cyst population after correcting for the number of underdeveloped juveniles per cyst population. All error bars represent standard error.

3.4 Discussion

The combination of methods used here has brought new light to our knowledge of the PCN eggshell. Technical developments have vastly increased mass spectrometer sensitivity allowing smaller sample sizes to be used for protein identification. Combining this with good quality genome resources allows identified peptides to be matched to predicted proteins. Furthermore, reliable RNAseq data providing gene expression values allows life-stage specific roles of identified proteins to be studied. However, although the application of these new techniques has allowed significant advances to be made, the methods used here are not perfect. Contamination from juvenile surface proteins is visible (GROS_g14202, Prior *et al.*, 2001) in the eggshell protein extraction data. Other identified proteins that may not have a role in eggshell function include the FAD binding domain proteins and neprylisin. Recent work shows both proteins were upregulated in juveniles following 8 hours exposure to root diffusates (Duceppe *et al.*, 2017). However, eggs used for the protein extractions were not exposed to RD in this study. Contamination could be reduced by washing the eggshells more but this risks losing sample through loss of the lipoprotein layers beneath the eggshell chitin.

Detection of chondroitin proteoglycans in PCN eggshell protein extractions demonstrates that structurally, PCN and *C. elegans* eggshells may be more similar than thought. As with *C. elegans*, CpG presence suggests that the outer lipid layer as described for general tylenchid eggs by Perry & Trett (1986) is potentially a CpG layer, as seen by Olson *et al.* (2012). Previous reports suggested that 59% of the eggshell is made up of protein but none of these proteins have been previously identified. The presence of a CpG layer can support this. Proteoglycans are a subclass of glycoproteins. Instead of a carbohydrate side chain, proteoglycans have side chains built from amino-sugars, containing nitrogen. Therefore, a protein content measurement based on nitrogen presence may be artificially inflated.

Due to the predicted place of calcium within the hatching cascade of PCN, the annexin GROS_g03104 was of interest. The unique peptides identified by mass spectrometry confirm that GROS_g03104 is the annexin from the protein sample and not another

similar member of the annexin family. Aligning GROS_g03104 to annexins in other nematode species shows an extremely high level of sequence similarity. However, one short glycine-rich domain stands out as being specific to the genus *Globodera.*

Expression data for the annexin in both *G. pallida* and *G. rostochiensis* shows a high level of expression at the egg producing stage. In *G. pallida,* the eggshell annexin is the 7th highest expressed gene in the cyst life stage, suggesting a key role of the protein within the nematode lifecycle. The lack of expression within the parasitic life stages rules out a role for this protein in parasitism. Comparing the RNA-seq data to the root-knot nematode *M. incognita* shows that there is no up-regulation in the egg producing stages for any of the top hits for this annexin. This reflects the difference in biology of PCN and RKN; as there is no described role for calcium in hatching of root-knot nematodes, meaning that there may be no need for an eggshell localised calcium binding protein.

Antibody production allowed immunolocalisation of GROS_g03104 to the PCN eggshell. It is thought that this is the first eggshell protein to be identified and localised to the eggshell of any parasitic nematode. Localisation was not carried out to a level of detail to identify which layer of the eggshell the protein resides due to time constraints. However, further localisation could be carried out with immunogold imaging under an electron microscope. Nevertheless, due to the calcium dependency and lipid binding properties of annexins (discussed in **chapter 4**) it is likely that GROS_g03104 sits in the lipoprotein layer beneath the chondroitin proteoglycan layer.

Structural prediction for GROS_g03104 was carried out based on homology searches against proteins with an already solved crystal structure. The predicted GROS_g03104 construct has 5 expected calcium binding sites. Two of these locations were omitted as there is no similar calcium binding site from annexin orthologues in other species. The first calcium binding site interacting with amino acids at position 191, 192, 193 and 233 represents a clear pocket where calcium can bind. This binding site can also be seen in other annexins with established structure. The second calcium binding site again matches homologous sites found on other annexins. However, the binding potential seems low due to only one amino acid (leu-77) being implicated and the coordination

number of calcium requiring six neighbouring ions or water molecules. The final calcium binding site is modified in GROS_g03104 due to the presence of the *Globodera* specific motif in this annexin (FFGIGNLGI). It is possible that this modified binding site has changed to allow docking of a much larger molecule. However, early docking modelling work for potential interactions between this binding site and three potential PCN hatching factors (solanoeclepin-A, α-chaconine and α-solanine) shows few interactions between this unique motif and hatching factors. This is possibly due to unknown compound bond angles for the hatching factors and unknown rotational values meaning that mostly rigid models were docked (data not included).

The potential for GROS_g03104 to act as a homo-multimer was also investigated through comparison to a *H. vulgaris* annexin XII homo-hexamer pore. Here, the *Globodera* specific motif within the protein is shown to be at the interacting surface between two GROS_g03104 monomers. Using a space filling model highlights the possibility for this motif to act as an embracing arm between the two monomers. However, without laboratory experiments such solving the protein crystal structure, gel filtration or native gel electrophoresis it is not possible to know whether the protein acts in a multimer or where the calcium binding sites truly are. Use of a GROS_g03104 homo-multimer to form a pore in the permeability barrier would make sense as this would allow tight regulation of molecule passage across the lipid bilayer. However, although the *H. vulgaris* pore would allow passage of water molecules, this annexin channel would be too narrow to allow movement of larger molecules such as trehalose. Without structural knowledge of a GROS_g03104 multimer it is impossible to say whether the diameter of this theoretical annexin pore would be larger.

After localising GROS_g03104 to the eggshell, annexin presence in *G. rostochiensis* eggs was consequently reduced via knock-downs using transgenic lines of Desiree expressing an annexin specific hairpin RNA. RNA interference in PCN is not an exact science. Insertion of recombinant DNA into host plants is unguided and so has the potential for off target effects through disruption of genes within the plant. This method also relies on the nematodes carrying out their lifecycle ordinarily and consuming the interfering

hairpin RNA. Finally, there is a chance that the hairpin RNA will not be processed correctly or simply will not interfere with the target nematode gene. Fortunately, here, qPCR revealed that expression of the eggshell annexin had been knocked-down in some cyst populations. Unfortunately, knock-down was not uniform across all populations taken from the same plant line. This means that annexin expression in populations not tested by qPCR may not have been altered.

Cyst sizes in eggshell annexin knock-down populations were unaltered. Eggshells formed within cysts and appeared regular, although in some cases they were of reduced size and volume. The first signs of abnormality were noticed in the increased number of under-developed juveniles present in annexin knock-down lines. Similarly, there was a visible size difference between some juveniles in RNAi populations and negative control populations. Combining these observations with the significant decrease in width of eggshells in RNAi populations suggests that annexin knock-down affected eggshell permeability. Increased permeability would allow increased movement of fluid out of the eggshell. Therefore, when the juvenile has developed, dehydration can occur at an increased rate, decreasing the juvenile to an unusually small volume and reducing the osmotic pressure within the eggshell, reducing the overall eggshell size. Increased under-development is a phenotype often associated with interfering with the permeability barrier of *C. elegans*. Interrupting the permeability barrier potentially allows small molecules signalling between embryonic cells to escape (Schierenberg & Junkersdorf, 1992; Olson *et al.*, 2012).

Significant increases of hatching in annexin RNAi populations were also seen with around double the percentage of nematodes hatched when compared to negative control lines after 4 weeks. A direct correlation between increased hatching and increased knock-down of the eggshell annexin was noticeable. Although the average percentage of hatching was less than 10% after 4 weeks, this becomes more significant when remembering that on average 10% of the juveniles from RNAi lines were underdeveloped. Furthermore, these hatching assays were carried out before the nematodes entered into an obligatory diapause.

There is still much to be learnt about PCN eggshells. The work in this chapter has allowed identification and localisation of the first PCN eggshell protein, annexin GROS_g03104. It is likely that reducing annexin availability reduces lipid binding, increasing the permeability of the eggshell. This allows increased dehydration of juveniles in the eggshells (reducing their size) and increases the ability to rehydrate after dormancy, increasing the number of juveniles hatching. Predictions of this annexin's structure highlights potential calcium binding sites, one of which has been modified by a small peptide motif specific to *Globodera spp.* Expansion of this binding site could suggest modification of this binding site to allow docking of a larger molecule/ion in place of calcium.

3.5 References

ATKINSON, H.J. & TAYLOR, J.D. 1983. A calcium-binding sialoglycoprotein associated with an apparent eggshell membrane of Globodera rostochiensis. *Annals of Applied Biology*, **102**, 345–354.

ATKINSON, H.J., ROBINSON, M.P. & PERRY, R.N. 1985. The Effect of Delayed Emergence On Infectivity of Juveniles of the Potato Cyst Nematode Globodera Rostochiensis. *Nematologica*, **31**, 171–178.

AVILA-SAKAR, A.J., CREUTZ, C.E. & KRETSINGER, R.H. 1998. Crystal structure of bovine annexin VI in a calcium-bound state. *Biochimica et biophysica acta*, **1387**, 103–116.

CARTAILLER, J.P., HAIGLER, H.T. & LUECKE, H. 2000. Annexin XII E105K Crystal Structure: Identification of a pH-Dependent Switch for Mutant Hexamerization †. *Biochemistry*, **39**, 2475–2483.

CLARKE, A.J. & HENNESSY, J. 1983. The role of calcium in the hatching of Globodera rostochiensis. *Revue de Nématologie*, **6**, 247–255.

CLARKE, A.J. & PERRY, R.N. 1980. Egg-shell permeability and hatching of Ascaris suum. *Parasitology*, **80**, 447–456.

CLARKE, A.J. & PERRY, R.N. 1985. Egg-shell calcium and the hatching of Globodera rostochiensis. *International Journal for Parasitology*, **15**, 511–516.

CLARKE, A.J. & PERRY, R.N. 1988. The induction of permeability in egg-shells of Ascaris suum prior to hatching. *International Journal for Parasitology*, **18**, 987–990.

CLARKE, A.J., COX, P.M. & SHEPHERD, A.M. 1967. The chemical composition of the egg shells of the potato cyst-nematode, Heterodera rostochiensis Woll. *The Biochemical journal*, **104**, 1056–1060.

COTTON, J.A., LILLEY, C.J., JONES, L.M., KIKUCHI, T., REID, A.J., THORPE, P., TSAI, I.J., … URWIN, P.E. 2014. The genome and life-stage specific transcriptomes of Globodera pallida elucidate key aspects of plant parasitism by a cyst nematode. *Genome Biology*, **15**, R43.

DEVINE, K. & RYAN, A. 2005. Comparison of the in-soil hatching responses of Globodera rostochiensis and G. pallida in the presence and absence of the host potato crop cv. British Queen. *Nematology*, **7**, 587–597.

DUCEPPE, M., LAFOND-LAPALME, J., PALOMARES-RIUS, J.E., SABEH, M., BLOK, V., MOFFETT, P. & MIMEE, B. 2017. Analysis of survival and hatching transcriptomes from potato cyst nematodes, Globodera rostochiensis and G. pallida. *Scientific Reports*, **7**, 3882.

EVES-VAN DEN AKKER, S., LAETSCH, D.R., THORPE, P., LILLEY, C.J., DANCHIN, E.G.J., DA ROCHA, M., RANCUREL, C., … JONES, J.T. 2016. The genome of the yellow potato cyst nematode, Globodera rostochiensis, reveals insights into the basis of parasitism and virulence. *Genome Biology*, **17**, 124.

FIORETTI, L., WARRY, A., PORTER, A., HAYDOCK, P. & CURTIS, R. 2001. Isolation and localisation of an annexin gene (gp-nex) from the potato cyst nematode, Globodera pallida. *Nematology*, **3**, 45–54.

FOOR, W.E. 1967. Ultrastructural aspects of oocyte development and shell formation in Ascaris lumbricoides. *The Journal of Parasitology*, **53**, 1245–1261.

GLEAVE, A.P. 1992. A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Molecular Biology*, **20**, 1203–1207.

KATZ, A., GLUSKER, J.P., BEEBE, S.A. & BOCK, C.W. 1996. Calcium Ion Coordination: A Comparison with That of Beryllium, Magnesium, and Zinc. *Journal of the American Chemical Society*, **118**, 5752–5763.

KELLEY, L.A., MEZULIS, S., YATES, C.M., WASS, M.N. & STERNBERG, M.J.E. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, **10**, 845–858.

KENNEDY, M.J., SCHOELZ, J.E., DONALD, P.A. & NIBLACK, T.L. 1997. Unique Immunogenic Proteins in Heterodera glycines Eggshells. *Journal of Nematology*, **29**, 276–281.

KERRY, B., DAVIES, K. & ESTEVES, I. 2009. Managing potato cyst nematode through maximising natural decline and population suppression. *Agriculture and Horticulture Development Board, Project REF R269*.

MASLER, E.P. & PERRY, R.N. 2018. Hatch, survival and sensory perception. *In* Perry, R.N., Moens, M. & Jones, J.T., *eds. Cyst Nematodes*. CABI, 44–73.

OLSON, S.K., GREENAN, G., DESAI, A., MÜLLER-REICHERT, T. & OEGEMA, K. 2012. Hierarchical assembly of the eggshell and permeability barrier in C. elegans. *Journal of Cell*

Biology, **198**, 731–748.

PERRY, R.N. 2002. Hatching. *In* Lee, D.L., *ed. The Biology of Nematodes*. Taylor & Francis, 147–169.

PERRY, R.N. & CLARKE, A.J. 1981. Hatching mechanisms of nematodes. *Parasitology*, **83**, 435–449.

PERRY, R.N. & TRETT, M.W. 1986. Ultrastructure of the eggshell of Heterodera schachtii and H. glycines (Nematoda: Tylenchida). *Revue de Nematologie*.

PRIOR, A., JONES, J.T., BLOK, V.C., BEAUCHAMP, J., MCDERMOTT, L., COOPER, A. & KENNEDY, M.W. 2001. A surface-associated retinol- and fatty acid-binding protein (Gp-FAR-1) from the potato cyst nematode Globodera pallida: lipid binding activities, structural analysis and expression pattern. *The Biochemical Journal*, **356**, 387–394.

SCHIERENBERG, E. & JUNKERSDORF, B. 1992. The role of eggshell and underlying vitelline membrane for normal pattern formation in the early C. elegans embryo. *Roux's Archives of Developmental Biology*, **202**, 10–16.

SCHNEIDER, C.A., RASBAND, W.S. & ELICEIRI, K.W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, **9**, 671–675.

STEIN, K.K. & GOLDEN, A. 2015. The C. elegans eggshell. *In WormBook: the online review of C. elegans biology*. 1–36. Accessed from: http://www.wormbook.org/chapters/www_eggshell/eggshell.html

TEFFT, P.M. & BONE, L.W. 1984. Zinc-mediated hatching of eggs of soybean cyst nematode,Heterodera glycines. *Journal of Chemical Ecology*, **10**, 361–372.

WASS, M.N., KELLEY, L.A. & STERNBERG, M.J.E. 2010. 3DLigandSite: predicting ligand-binding sites using similar structures. *Nucleic Acids Research*, **38**, W469-73.

WESLEY, S. V., HELLIWELL, C.A., SMITH, N.A., WANG, M.B., ROUSE, D.T., LIU, Q., GOODING, P.S., … WATERHOUSE, P.M. 2001. Construct design for efficient, effective and high-throughput gene silencing in plants. *The Plant journal : for cell and molecular biology*, **27**, 581–590.

4. Identification and characterisation of PCN eggshell lipids

4.1 Introduction

Electron microscopy shows that inner lipid layers are visible beneath the chitin layer of *G. rostochiensis* eggshells (Perry *et al.*, 1982). Before PCN juveniles hatch from their egg, there is a change in eggshell lipid permeability. This change may be due to binding of hatching factors or from protein interactions within the eggshell. However, ultimately it is the hydrophobic properties of the eggshell lipid layer that must be altered to allow the influx of water and rehydration of the quiescent nematodes before eclosion can occur.

4.1.1 Enzymes

Several studies have suggested that enzymes are used by some nematodes to begin permeabilisation. Considering that hatching factors are effective at extremely low concentrations, it is unlikely that they are solely responsible for physically enabling nematode hatching (Perry & Clarke 1981). In M*eloidogyne* spp. metabolic activation of the nematode occurs before the eggshell permeability change (Perry 2002). This would allow the unhatched juvenile to begin secreting eggshell specific enzymes only when the nematode is prepared to exit the protection of the eggshell. However, in PCN, a change in eggshell permeability occurs before metabolic activation, meaning either that any hatching enzymes must already exist within the eggshell structure or within the perivitelline fluid, or that degradation of the eggshell does not require enzyme activity.

Structural properties of the nematode eggshell can be attributed to the thick chitinous layer. Recently, the PCN transcriptome was investigated before and after initiation of the hatching cascade. This showed a significant upregulation of the *cht-2* gene, encoding a chitinase, following exposure to PRD (Duceppe *et al.* 2017). Similar genes have also

been identified in both *H. glycines* and *M. incognita* although expression of these genes has not yet been analysed in such detail as in PCN (Dautova *et al.* 2001). Previous work on *M. incognita* showed that the eggshell had become soft and flexible after hatching, probably as a result of enzymatic degradation (Bird 1968). However, unlike *M. incognita*, the PCN eggshell remains rigid after hatching (Perry *et al.* 1992). Additionally, commercially available chitinase has the ability to reduce *G. rostochiensis* hatching (Cronin *et al.* 1997), again suggesting the absence of a role for chitinase within the PCN hatching cascade.

Chitinase and lipase activity in hatching of *A. lumbricoides* has been noted previously (Rogers 1958, Fairbairn 1961). Hatching lipases were thought to be present in the egg fluid, repressed by an unknown internal inhibitor or possibly trehalose. However, if this was the case egg fluid associated hatching enzymes would only exhibit activity after release of egg fluid. In *Globodera spp.* release of this fluid does not happen until after lipid barrier permeability meaning that enzymes would still be inhibited (Perry & Clarke 1981). Perry *et al.* (1992) assayed potential lipases in *M. incognita* and *G. rostochiensis*. Clear lipase activity was present in hatching of *M. incognita*. However, no lipase activity was detected during the 3-week hatching period for *G. rostochiensis*.

4.1.2 Ascarosides

Ascarosides are seemingly nematode specific glycolipids. First identified in eggshells of *Ascaris spp.,* ascarosides have now been associated with numerous nematodes including *C. elegans*. This group of glycolipids obtains its name from the unique ascarylose sugar a 3,6-dideoxy-L-mannose (Bartley *et al.* 1996).

The ascarosides that were first identified in *Ascaris* eggshells were long chained (C_{27}-C_{35}) and saturated. Not all the hydrocarbon chains terminate with a carboxyl group; some can terminate with a branched methyl group or another ascarylose unit (**Figure 4.1**). True *Ascaris* eggshell ascarosides are thought to be a key component of the permeability barrier (Bartley *et al.* 1996). This is a function that is well supported by the structure of the identified eggshell ascarosides. Saturated hydrocarbon chains result in a higher melting temperature. Branching that would occur when a chain is unsaturated would

decrease the ability of the fatty acids to compact together. Likewise, the longer chain lengths exhibited by *Ascaris* ascarosides further increases the glycolipid melting temperature due to increased van der Waals interactions between neighbouring chains. A membrane consisting of these true ascarosides would form a strong permeability barrier. This is a feature that is necessary for the lifecycle of *Ascaris* as eggs of these nematodes must survive the harsh conditions found in animal digestive systems.

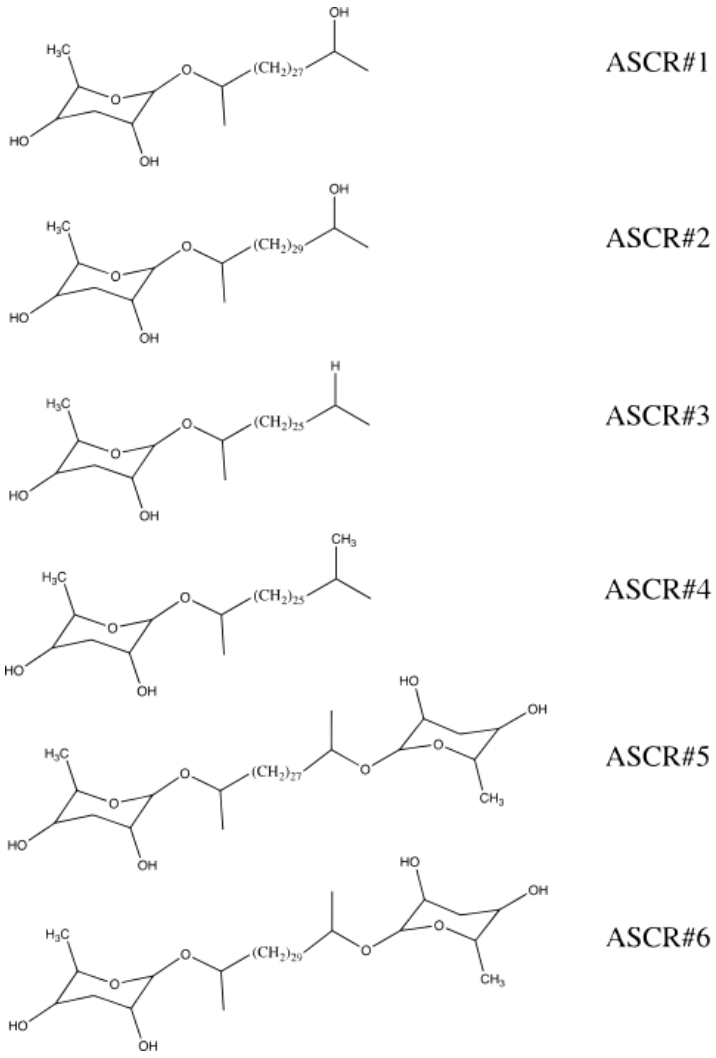


Figure 4.1 – True Ascaris eggshell ascarosides

Ascarosides originally identified in *Ascaris* as described by Bartley, Bennett, & Darben (1996). Unlike ascaroside later described by Jeong *et al.* (2005) in *C. elegans*, these true ascarosides exhibit much longer hydrocarbon chains and relatively unmodified ascarylose sugar units.

In *Ascaris spp.* the eggs become fertilised when passing through the oviduct and into the di-branched uterus. It is here that eggshell formation begins. Tarr & Fairbairn (1973)

showed that as the distance from the oviduct increases, the amount of esterified ascarosides decreases. This work also concluded that *Ascaris* eggshell ascarosides are not esterified. Therefore, final, detectable ascarosides will exhibit the same predicted 3,6- dideoxy sugar.

The presence of ascarosides has been inferred in *C. elegans* eggshells due to upregulation of carbohydrate and fatty acid synthesis and modification enzymes during egg deposition (Olson *et al.* 2012). However, the only two long chained ascarosides to be identified in *C. elegans* have been shown not to be associated with eggshell formation (Zagoriy *et al.* 2010). To date, no ascarosides have been located in eggshells of *C. elegans.*

A large range of non-eggshell associated ascaroside-like compounds have been discovered across life-stages of *C. elegans*. It is thought that this group of compounds act as signalling molecules between nematodes and are involved in inducing dauer stage formation (Jeong *et al.* 2005). These molecules do not resemble the true ascarosides found in *Ascaris* eggshells, exhibiting much shorter hydrocarbon chains and frequently with a modified ascarylose sugar (von Reuss *et al.* 2012). It is particularly important to differentiate between the dauer inducing ascarosides (daurmones) and true eggshell ascarosides as conventional naming terms often overlap between both sets of molecules. For example, ascaroside #2 found in eggshells of *Ascaris* has a long, branched, C32 side chain. Ascaroside #2 found in *C. elegans* is a much shorter variant with a C6 ketone side chain (Butcher *et al.* 2007). Recently it has been suggested that the ascarylose sugar is a basis for building a diverse range of more complicated signalling molecules (Panda *et al.* 2017).

Due to the theoretical ability of eggshell ascarosides to protect a developing juvenile from a harsh external environment they have been of interest to plant-parasitic nematologists. In *G. rostochiensis,* where eggs can lay dormant in the soil for up to 20 years, a strong permeability barrier is key to avoid premature eclosion. Clarke & Hennessy (1977) investigated the possibility of *G. rostochiensis* eggshell ascarosides but

were unable to identify any of these compounds. However, this study only looked for the true, long chained ascarosides as found in *Ascar*is.

4.1.3 Nematode eggshell phospholipids

Phospholipids are key components of cellular membranes. The hydrophobic properties of a phospholipid layer enable membranes to form a distinct separation between the two sides of the membrane. There has been surprisingly little research into phospholipids in PCN eggshells considering how important the eggshell permeability barrier is to a developing juvenile.

As with most cases, working with a model organism allows more straightforward analysis of the eggshell. However, even the eggshell of the model nematode *C. elegans* has not yet had its lipids fully characterised. Additionally, this free-living nematode's hatch is not dependent on a host, suggesting that *C. elegans* would not be a model for parasitic nematodes in this respect.

Eggshell lipid analysis of the ruminant parasite *Haemonchus contortus* highlighted four major lipid groups being identifiable within samples. These were, triglycerides, sterols, sterol esters and phospholipids. Of these phospholipids the major species were phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylinositol (PI) and phosphatidic acid (PA). No lipids specific to *H. contortus* eggshells were identified (Riou *et al.* 2007). The presence of ascarosides or glycolipids in these eggshells was not discussed. Sterols are only found in extremely low quantities in plant-parasitic nematodes (Chitwood & Lusby 1991). The presence of sterols in PCN eggshells is unlikely as cysts containing eggs only have minor amounts of sterols and steryl esters (Gibson *et al.* 1995).

In PPN, eggshell lipid composition appears to be similar to that of APN. *M. incognita* and *M. arenaria* both have eggs with an abundance of neutral lipids (>90%). Fatty acids are mostly longer chain, with low degrees of unsaturation as predicted for having increased function in forming strong membranes (Krusberg *et al.* 1973).

Although limited, there has been some research into PCN eggshell lipid species. Neutral lipids were found to be the major component of total lipid extracts making up 55-75%. Of these neutral lipids, triacylglycerols were the most predominant (Gibson *et al.* 1995), corresponding with findings in *H. contortus* and *Meloidogyne spp.*. To our knowledge, no specific PCN eggshell glycerophospholipids have been identified.

Currently no research has looked at lipids extracted solely from eggshells, instead looking at lipids extracted from cysts with eggshells inside. This is a far easier way of collecting lipid samples that may contain eggshell lipids and with enough controls this type of analysis could allow differences between samples with and without eggshells to be identified. However, this indirect method of analysing eggshell lipids is susceptible to contamination and bias due to unknown numbers of juveniles present inside the cysts. Developing PCN are packed full of phospholipids to aid survival of the juvenile through diapause and to provide the hatched J2 with the energy required to locate and invade a host. Therefore, the subtle differences between a lipid sample taken from juveniles and a lipid sample taken from cysts containing eggshells are likely to be far too small to allow definitive conclusions about the lipids present in eggshells to be drawn.

4.1.4 Chapter 4 aims

The aim of this chapter is to identify eggshell lipids separated from purified PCN eggshell samples. Lipids bound by the eggshell annexin GROS_g03104 will also be identified. Total lipid extractions from purified PCN eggshells will be carried out to identify any eggshell associated lipids. It is hoped that extracting lipids directly from eggshells, instead of using cysts containing eggshells, will highlight subtleties that could otherwise be overlooked. Previous knowledge of nematode eggshell ascarosides will be used to search for presence of PCN eggshell associated ascarosides.

4.2 Methods

4.2.1 Collecting eggshells

Eggshells were collected using sonication and floatation methods described in general methods section **2.1.3**. Due to the possibility that ultrasonication could fracture eggshells and damage the inner lipid layer, eggshells for lipid extractions were also fractured by placing samples between two glass sheets and applying pressure to the top sheet. The sheets were then washed into a glass beaker using dH_2O, transferred to centrifuge tubes and centrifuged (2500 rpm, 10 minutes) to collect the nematode samples. Empty eggshells were separated from this mixture using the sodium-potassium tartrate gradient method described in **2.1.3**.

4.2.2 Lipid dot blots

Recombinant GROS_g03104 (annexin) was obtained using methods described in general methods section **2.6.5**. Membrane lipid strips were purchased from Echelon Biosciences. These membranes are dotted with 15 different lipids that are biologically important in cell membranes.

Lipid strips were blocked in 3% fatty acid-free BSA in protein binding buffer TTBS (50 mM Tris pH 8, 150 mM NaCl, 0.1% Tween20 containing either 1 mM $CaCl_2$ or EGTA) for 1 hour at room temperature in the dark. The membrane was then incubated at 4 °C overnight in the same buffer, this time containing recombinant 10 µg/mL GROS_g03104. Membranes were washed 3 times in PBS with 0.1% Tween20 before incubating in anti-GROS_g03104 primary antibodies in 0.1% fatty acid-free BSA (**3.2.4**) for 1 hour at room temperature. Membranes were washed 3 times in PBS to remove unbound primary antibodies before exposing membranes to anti-rabbit secondary antibodies for 1 hour at room temperature. Membranes were washed in PBS to remove secondary antibody solution and imaged using a LI-COR Odyssey CLx imaging system.

4.2.3 Eggshell lipid extraction

Any solvents that were used in the extraction and preparation of lipid samples were analytical HPLC grade. Where possible, glassware was used after solvents were introduced to samples to avoid stripping chemicals from plastic containers.

Samples for lipid extraction were pelleted and resuspended in 200 µL of PBS before being transferred to a glass vial. 750 µL of a 2:1 methanol:chloroform mix was added to the sample which was left shaking overnight at 4 °C. A biphasic solution was formed by adding 250 µL H_2O and 250 µL chloroform, the mixture was mixed and left to separate. The bottom lipid-containing chloroform layer was moved to a new glass vial and dried under N_2 gas. Lipid samples were stored in the cold room at 4 °C.

Due to the small sample sizes, lipid extraction using methyl tert-butyl ether (MTBE) was also used for some samples. This work did not look to compare the two extraction methods and both extraction methods yielded similar results.

Eggshells were concentrated to 40 µL samples and transferred to a glass vial. To this, 300 µL methanol was added and the vial was briefly vortexed. 1 mL MTBE was added and the sample was left shaking on a vibrating platform at room temperature for 2 hours. A biphasic solution was created by adding 250 µL of dH_2O, to ensure the two phases were completely formed, vials were centrifuged for 10 minutes at 1000 x g. The upper phase was extracted and dried under a stream of N_2 gas. To aid evaporation of the MTBE, methanol was added to the sample when needed. Dry lipid samples were stored at 4 °C until analysis.

4.2.4 Electrospray Tandem Mass Spectrometry

Dried lipid samples were resuspended in 15 µL 2:1 methanol:chloroform and 15 µL 6:7:2 acetylnitrile:isopropanol:water. Samples were analysed using electrospray tandem mass spectrometry carried out on an AB-Sciex Qtrap 4000 triple quadrupole mass spectrometer. An associated Advion TriVersa NanoMate connected to the nano-electrospray ionisation source was used to load samples.

Negative ion mode produced survey scans of lipid species containing PE, PI or PS head groups. Positive ion mode detected PC head groups. Survey spectra were acquired within the range of 120-1600 m/z. Tandem mass spectrometry allowed daughter ions to be fragmented from peaks of interest using nitrogen as a collision gas. Each spectrum consisted of a minimum of 20 scans. Lipid species were interpreted using theoretical values found within the Lipid MAPS (Metabolites and Pathways Strategy) database (http://www.lipidmaps.org/). After analysis lipids were returned to the glass vial, dried under N_2 gas and stored at 4 °C.

4.2.5 Analysis of Fatty Acid Methyl-esters (FAMES)

Fatty acids required methylation before analysis. This was achieved by adding 1.5 mL methanol, 0.2 mL toluene and 0.3 mL of 8% HCl in methanol:water (85:15) to the dried lipid sample. Samples were incubated at 65 °C overnight. Methylated samples were dried under N_2 and resuspended in 0.5 mL hexane and 0.5 mL dH_2O. The top hexane layer was removed and transferred to a new glass vial then dried once more under N_2. Methylated fatty acids were then resuspended in 50 µL DCM, 20 µL of this solution was moved to a fresh autosampler vial with a glass insert for GC-MS analysis.

Analysis was carried out on an Agilent Technologies GC-6890N gas chromatograph together with an MS detector-5973. Separation was performed on a Phenomenex ZB-5 column with a program of 10 minutes at 70°C followed by a gradient increase to 220°C increasing at a rate of 5 °C per minute. Samples were held at 220 °C for a further 5 minutes. Spectra were acquired from 50-550 amu. Fatty acids were identified based on retention time and fragmentation patterns. Comparisons were also drawn against bacterial FAME standards.

Lipid Home (http://www.lipidhome.co.uk/) was used as a reference to compare samples against libraries of GC-MS data and fragmentation patterns.

4.2.6 Ascarosides

A synthetic version of ascaroside #18 was created in house (Fraser, 2017) and used to identify how an ascaroside would act under ESI-MS-MS conditions. After visualisation of

splitting patterns under MS conditions had been observed, the mass spectrometer could be tuned to attempt to heighten appearance of ascarosides compared to other lipid species present in the sample.

4.2.7 Carbohydrate extraction and analysis

Fifty *G. rostochiensis* cysts were crushed using a Potter Elvehjem homogeniser before being transferred to a glass vial. Juveniles (20,000) of the two species were hatched using TRD (**2.1.2**) before being sucrose floated and transferred to an Eppendorf tube. Nematodes were frozen in liquid nitrogen, crushed with a micro pestle and transferred to a glass vial. 10 µL of 1 mM monosaccharide standards were transferred to glass vials. The same synthetic ascaroside that was created in house (**4.2.5**) was used as an ascarylose standard following acid hydrolysis. All samples were frozen to -80 °C before being vacuum freeze dried.

After samples were dry, 200 µL of 6M HCl was added. The vials were heated to 110 °C overnight to hydrolyse polymeric units down to their monomeric components. Acid hydrolysed samples were dried in a speedvac vacuum concentrator. Once dry, samples were washed 3 times in 50 µL water to remove remaining HCl. Samples were then washed and dried twice more in 50 µL methanol; after this the sample could be stored at -20 °C.

Derivatization of samples was required to increase volatility of monosaccharides before GC-MS. Samples were reduced by shaking in 10 µL of a 20 mg/mL methoxyamine hydrochloride in pyridine solution for 1.5 hours. Free hydroxyls were silyated by adding 91 µL of MSTFA (N-Methyl-N-(trimethylsilyl) trifluoroacetamide) to all samples. Vials were placed on a shaking platform at room temperature for a further 30 minutes before being placed in an auto-sampler ready for GC-MS analysis. Samples were derivatized no more than 30 minutes before the sample was to be analysed.

GC-MS analysis was carried out on an Agilent Technologies GC-6890N gas chromatograph together with an MS detector-5973. Separation was performed by a Supelco SE-54 fused silica column with a program of 10 minutes at 70 °C followed by a

gradient increase to 220 °C increasing at a rate of 5 °C per minute. Monosaccharides were identified based upon retention time and fragmentation patterns. Comparisons between samples and monosaccharide standards were conducted.

4.2.8 Ascaroside analysis from *G. rostochiensis* juveniles

G. rostochiensis juveniles were hatched in TRD (**2.1.2**). Controls of water pre-exposure to tomato roots and TRD were taken, 200 mL each. Juveniles were left to hatch for 12 days before being removed from the TRD by centrifugation. 1 M HCl was added to the post hatching TRD (10% v/v) to acidify the solution. 10 mL butan-1-ol was then added per 100 mL of post hatch TRD:HCl mixture and the solution was inverted several times. The organic layer was removed from the top before another butan-1-ol extraction was carried out. The two organic phases were combined and dried in a speed vac vacuum concentrator.

Samples were taken for ascaroside analysis by ESI-MS-MS. Additional testing was carried out to detect ascarylose monosaccharides after acid hydrolysis and TMS derivatisation. Monosaccharides were identified using GC-MS as described in section **4.2.7**.

4.3 Results

4.3.1 GROS_g03104 lipid binding

The previously identified eggshell annexin (GROS_g03104, chapter 3) was predicted to exhibit calcium dependent lipid binding abilities. The lipid binding capability of the annexin was tested against common membrane lipids using a lipid blot (**Figure 4.2**). Lipid binding was tested in the presence or absence of $CaCl_2$. In both cases the protein showed clear binding to phosphatidylinositol 3,4,5 trisphosphate ($PI(3,4,5)P_3$). However, in the presence of calcium faint, but clear binding was seen to other PI species.

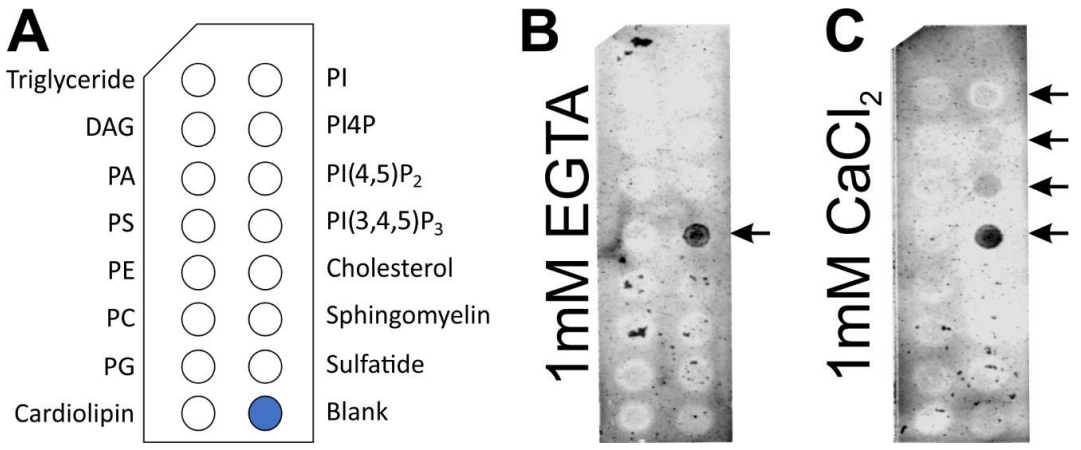


Figure 4.2 – Eggshell annexin lipid binding

A) Schematic map of lipids spotted onto the hydrophobic membrane. **B)** Binding of annexin to $PI(3,4,5)P_3$ in an absence of calcium. **C)** Binding of annexin in the presence of calcium. Here, increased binding to other PI lipid species is visible as indicated with arrows.

4.3.2 PCN eggshell lipids

As of yet, no data exist identifying PCN eggshell lipids from lipids extracted solely from eggshells. Eggshells were collected following pressing between two glass slides (**4.2.1**) to avoid shearing any lipid layers through sonication. Lipids were then extracted either using conventional chloroform:methanol extraction techniques or using MTBE before being dried under a stream of N_2 (**4.2.3**). Total lipid samples were submitted to Caroline

Horsburgh (School of Chemistry, University of St. Andrews) for high resolution electrospray mass spectrometry using a Thermo Exactive Orbitrap (Figure 4.3).

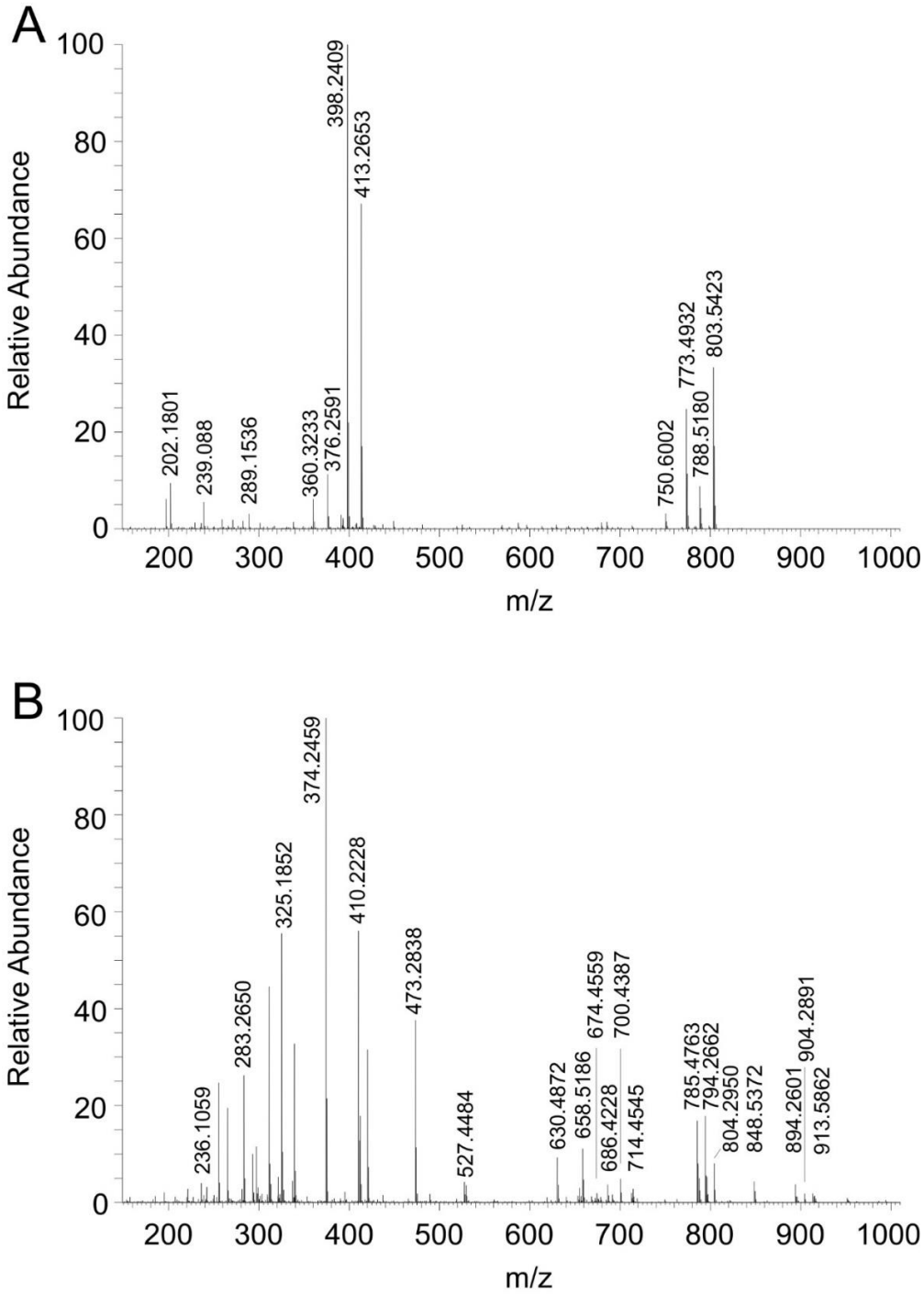


Figure 4.3 – High resolution electrospray mass spectrometry surveys of eggshell lipid extractions.

Survey scans in both positive **(A)** and negative **(B)** ion modes were carried out to identify parent ions. Data shown here was collected from MTBE extracted eggshell lipids. No differences were observed between MTBE and MeOH:$CHCl_3$ extraction techniques.

Parent ions identified by high resolution electrospray were compared between samples and reoccurring ions were further examined using tandem electrospray mass spectrometry for analysis of daughter ions. This allowed for identification of head groups associated with the lipid species. Neutral loss of m/z 87 is indicative of phosphatidylserine headgroups whereas precursors of m/z 184 are indicative of phosphatidylcholines. Under negative ionisation, a loss of m/z 241 is used to identify phosphatidylinositol species. Similarly, in negative ionisation mode, a loss of m/z 196 identifies potential phosphatidylethanolamine lipids. All glycerophospholipids show a loss of m/z 153 due to the loss of a cyclo-phosphoglycerol anion, therefore, a loss of m/z 153 scan can be used to differentiate between glycerophospholipids and other lipid species such as triacylglycerols.

MS-MS analysis was often complicated by the overwhelming presence of an unidentifiable peak with m/z 374 (m/z 398 in positive polarity). The accurate mass of ions obtained from high resolution electrospray mass spectrometry were searched in the Lipid Maps® (https://www.lipidmaps.org/) database to confirm the identified ion could match the suggested head group. Only masses greater than m/z 600 were analysed.

Finally, GC-MS was used to identify and quantify fatty acid methyl esters (FAMEs) confirming that the fatty acids that would be needed to create the lipid/phospholipid parent can be found within eggshell lipid extractions (**Figure 4.4**). The chromatogram here shows extraction of ions with m/z 74. This not only reduces background noise from artefacts related to column bleeding but also highlights methylated fatty acids. This ion is a signature of the McLafferty rearrangement ion and is a signature for FAME analysis (Christie & Han 2010). Further GC-MS spectra can be found in **7.8 GC-MS spectra**.

Table 4.1 summarises the lipids found within PCN eggshells.

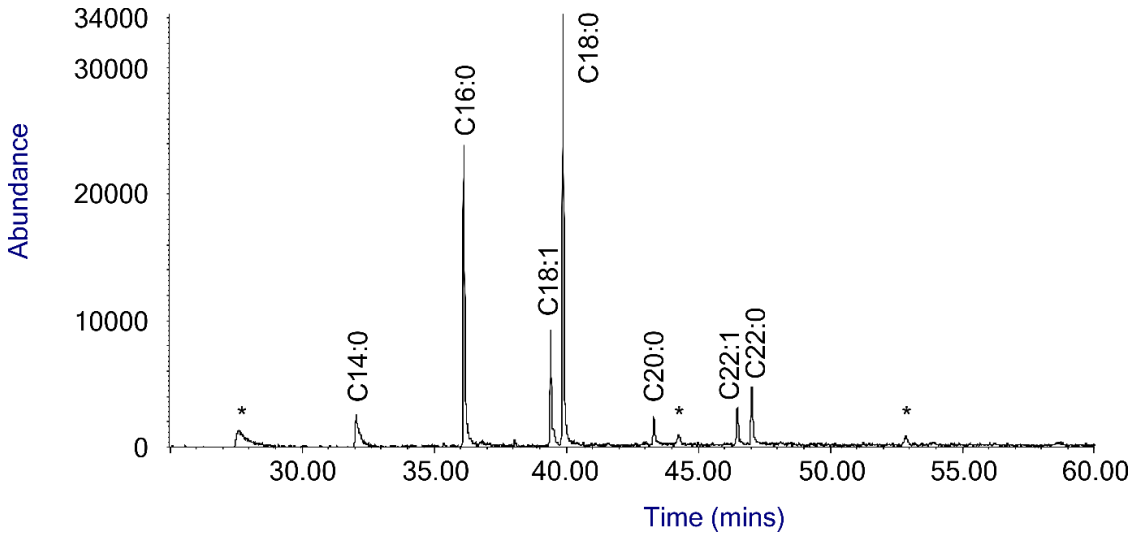


Figure 4.4 – Fatty acid methyl ester analysis of PCN eggshell lipids

Chromatogram showing ion extract of m/z 74. GC-MS analysis of FAMEs identified an abundance of C18 and C16 length chains in PCN eggshells. Contaminants, marked with * were also noticed. These likely result from bleed from the column.

Table 4.1 – PCN eggshell lipids summary

Summary of lipids identified from PCN eggshells. Peak number reflects the mass identified by the mass spectrometer with polarity noting whether this peak was visible after positive or negative ionisation. This data is visible in figure 4.3. Head group scans were used to identify commonly lost masses under MS-MS conditions.

Peak number	Polarity (+/-)	Lipid Maps database suggestion from accurate mass	Head group scan	Lipid
803.5	+	PT36:1 PC37:0 PC38:7	Not on P153 Not on P184	FAMEs would support PT36:1(18:1,18:0). PCN ability to synthesise phosphothreonine is not known
788.5	+	TAG47:2 PG37:2	Not on P153	Glycosyldiacylglycerol – 1,2-dioctadecanoyl-3-O-(6-deoxy-6 amino-a-D-glucosyl)-sn-glycerol (18:0, 18:0, supported by FAMEs)
773.5	+	PC35:1 PE38:1	Not on P153 Not on P184	Contamination from machine
750.6	+	TAG44:0 PG34:0	Not on P153	TAG44:0 (16:0,14:0,14:0)

685.4	+	PE32:3 DAG40:0	Not on P153	DG40:0 (22:0,18:0)
913.6	-	TAG56:2 PI40:4	On P153 On P241	PI40:4 FAMEs do not support. Likely present but in trace amounts
904.3	-	PI39:2 TAG56:7	903 on P241 905 on P153	FAMEs do not support PI39:2
894.3	-	PI 38:0 TAG56:12	Not on P241, m/z 241 in daughter scan suggesting PI	PI 38:0 (18:0, 20:0)
848.5	-	PG 41:0 TAG51:0	Not on P153	FAMEs do not support suggested TAG
804.3	-	PG38:1 TAG48:1	Not on P153	TAG48:1 (18:1,16:0,14:0)
794.3	-	TAG48:6 PA43:4	Not on P153	FAMEs do not support suggested TAG
785.5	-	PE39:2 PS36:3	Not on P153 Not on P196 Not on NL87	Glycosyldiacylglycerol – 1,2-dioctadecanoyl-3-O-(6-deoxy-6 amino-a-D-glucosyl)-sn-glycerol (18:0, 18:0, supported by FAMEs)
714.5	-	PG32:4 DG43:4	Not on P153	m/z 785.5 – NH_2, – C4:0. 16:0, 16:0 supported by FAMEs
700.4	-	PA36:2 DG42:4	On P153 On NL87	PA36:2 (18:1,18:1) PS30:3 unlikely from FAMEs
686.4	-	DG41:4 PG30:4	Not on P153	714.5 – C2:0 16:0, 14:0 supported by FAMEs
674.5	-	PA34:1	On P153	PA34:1 (18:1,16:0)
658.5	-	DG39:4 PA33:2	On P153, not DG	Degradation of 674
630.5	-	DG37:4 PA31:2	On P153, not DG	Degradation of 674

4.3.2.1 PIP_3 in PCN eggshells

As the annexin, GROS_g03104, has already been localised to the eggshell and has been shown to have a strong affinity for PIP_3 (**Figure 4.2**), it is assumed that PIP_3 species will also be found in PCN eggshells. Electrospray tandem mass spectrometry was used to identify PCN eggshell lipids. Mass spectrometry of this sort can identify loss of ions from larger compounds and so can also be used to identify loss of the PIP_3 head group with a mass of m/z 481. This identified two peaks, one of which (m/z 1246.8) matches the mass of PIP_3 46:0 on the lipid maps database (**Figure 4.5**). Although FAME analysis can support the presence of a C22:0 chain there was no sign of C24:0 to give a total of C46:0 (**Figure 4.4**). Similarly to before, sample contamination with an unidentifiable mass of 374 complicated MS-MS analysis of the potential PIP_3 with m/z 1246.8.

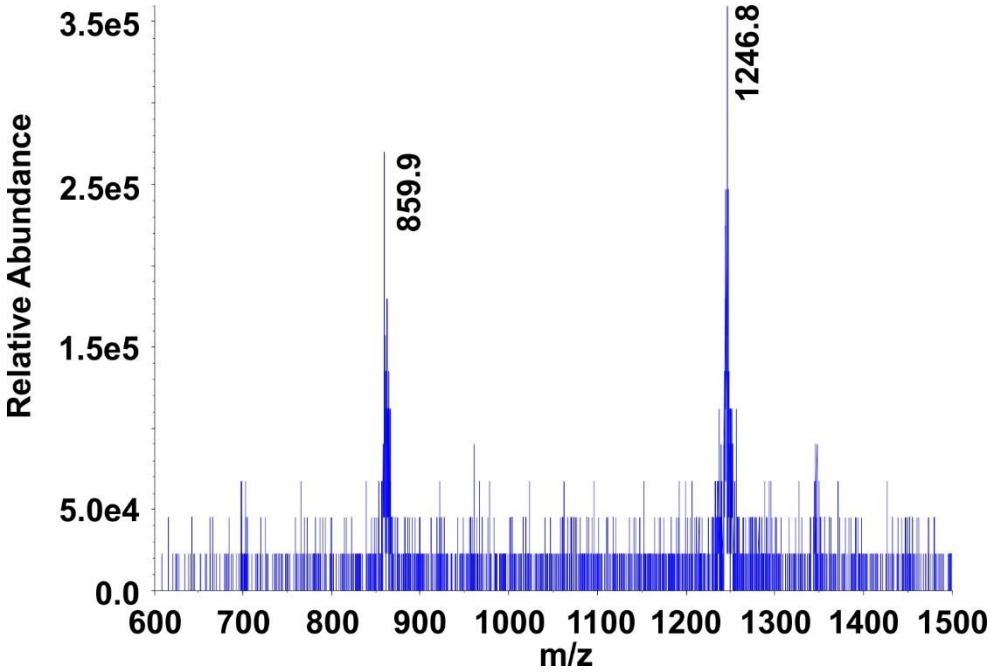


Figure 4.5 – Electrospray survey scan for PCN eggshell PIP_3
Precursor of m/z 481 was used to identify potential PIP_3 species in PCN eggshell lipid extractions. An ion of m/z 1246.8 matches the predicted mass for PIP_3 46:0 corresponding with Lipid Maps.

4.3.3 PCN eggshell ascarosides

Another group of potential PCN eggshell lipids are the ascarosides. Before being able to detect ascarosides in eggshell lipid extractions, an ascaroside control needed to be analysed to see how it would behave under ESI-MS-MS conditions. The same ascaroside

splitting pattern as seen in Figure 4.6 matches results seen by Bartley *et al.* (1996). Neutral loss of m/z 148 was also seen by Zagoriy *et al*. (2010). Using neutral loss of m/z 130 is further supported as eggshell ascarosides are predicted not to be esterified meaning the mass of ascarylose will not change from the expected m/z 130 (Tarr & Fairbairn 1973).

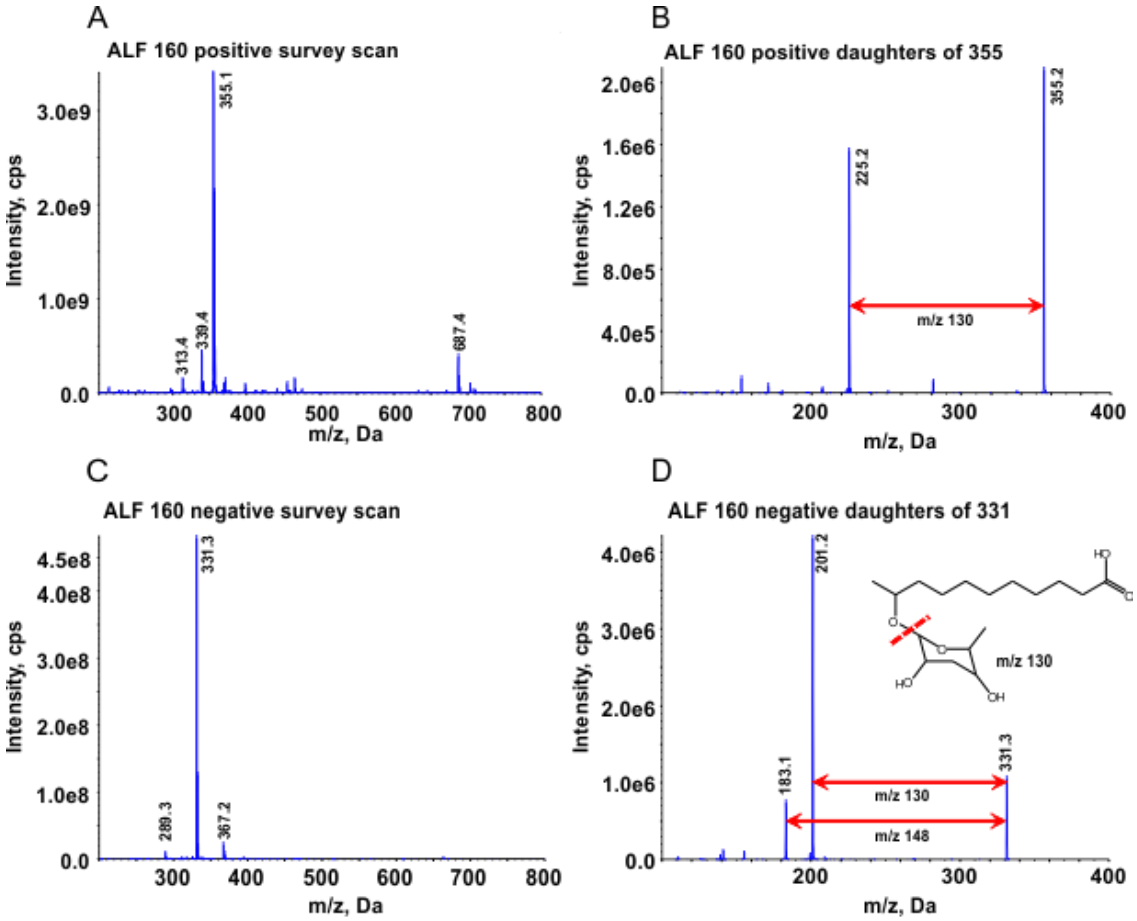


Figure 4.6 – ESI-MS/MS of synthetic ascaroside #18 (ALF160) control

A) Positive survey scan for ascr#18 showing sodium adduct of ascr#18 at m/z 355, [H^+] m/z 332. **B)** Fragmentation of ascr#18 sodium adduct m/z 355 shows loss of m/z 130. **C)** Negative survey scan for ascr#18 showing m/z of 331. **D)** Fragmentation of ascr#18 shows loss of m/z 130 and m/z 148. Peak at 201 corresponds with loss of ascarylose resulting in oxygenated undecylic acid.

After confirming ascaroside splitting patterns under ESI-MS-MS conditions the mass spectrometer can be tuned to detect lipid species that readily lose m/z 130. Using a quadrupole MS system allows parent ions to be scanned as usual before being fragmented by collision energy. Fragmented ions are then measured, ions that have now

lost the desired mass (m/z 130) are then plotted with the lost mass added back to the total mass (**Figure 4.7**). As with any lipid sample analysis, survey scans were carried out before MS-MS.

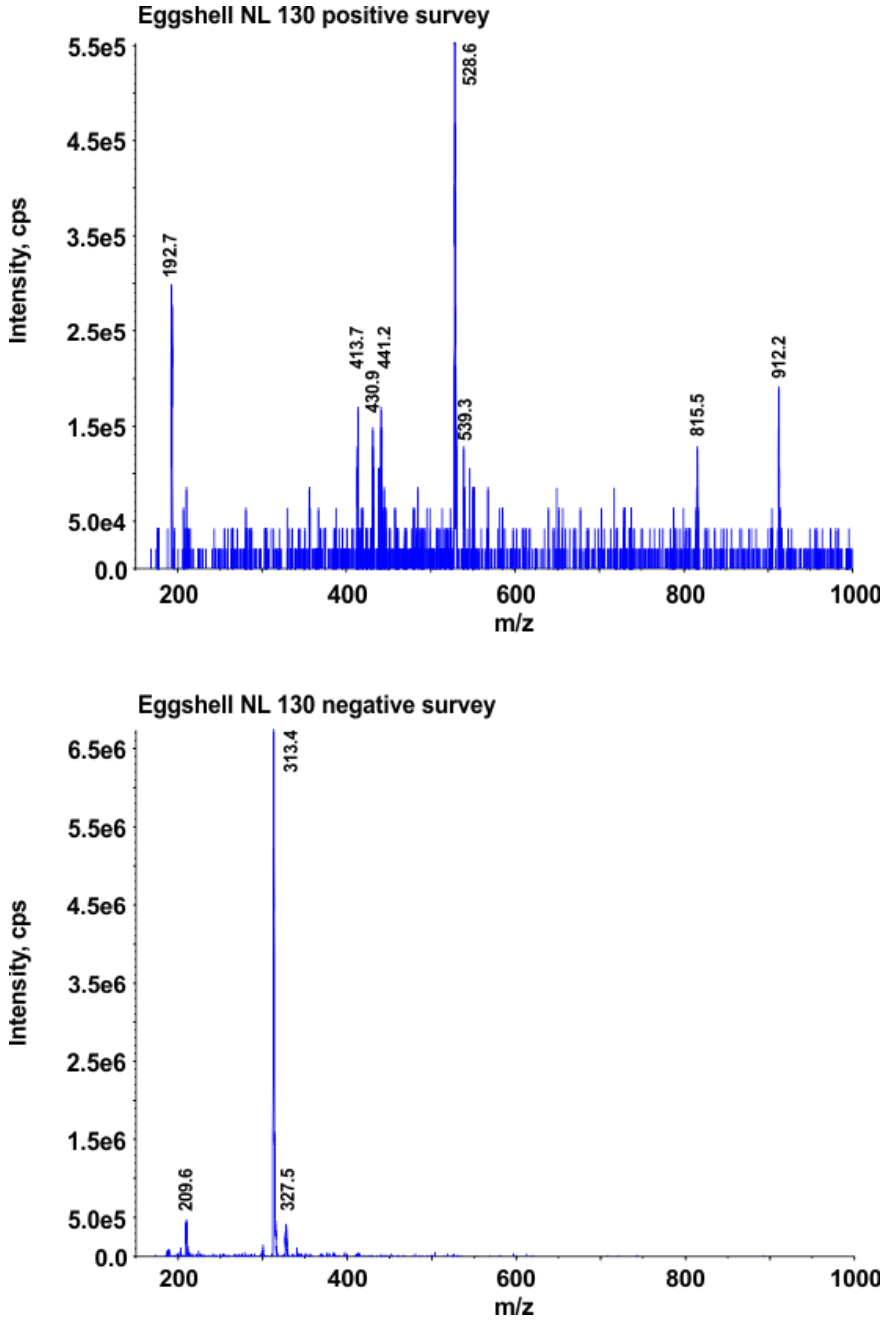


Figure 4.7 – Neutral loss of m/z 130 surveys from eggshell lipid extracts
Top) NL130 survey scan, positive mode. **Bottom)** NL130 survey scan, negative mode. Surveys scans comprised from a minimum of 138 MCA scans. Masses were detected between 150-1000 amu as this is the expected range for any eggshell associated ascarosides.

The NL m/z 130 surveys appear to show that the eggshell lipid extractions are full of ascarosides or ascaroside like compounds all capable of losing m/z 130. However, a pitfall of this type of mass spectrometry is that any compounds that are capable of losing

the desired m/z will be plotted on the spectra. This can be seen in **Figure 4.8** which shows MS-MS daughter scans for some of the masses detect by the NL 130 surveys.

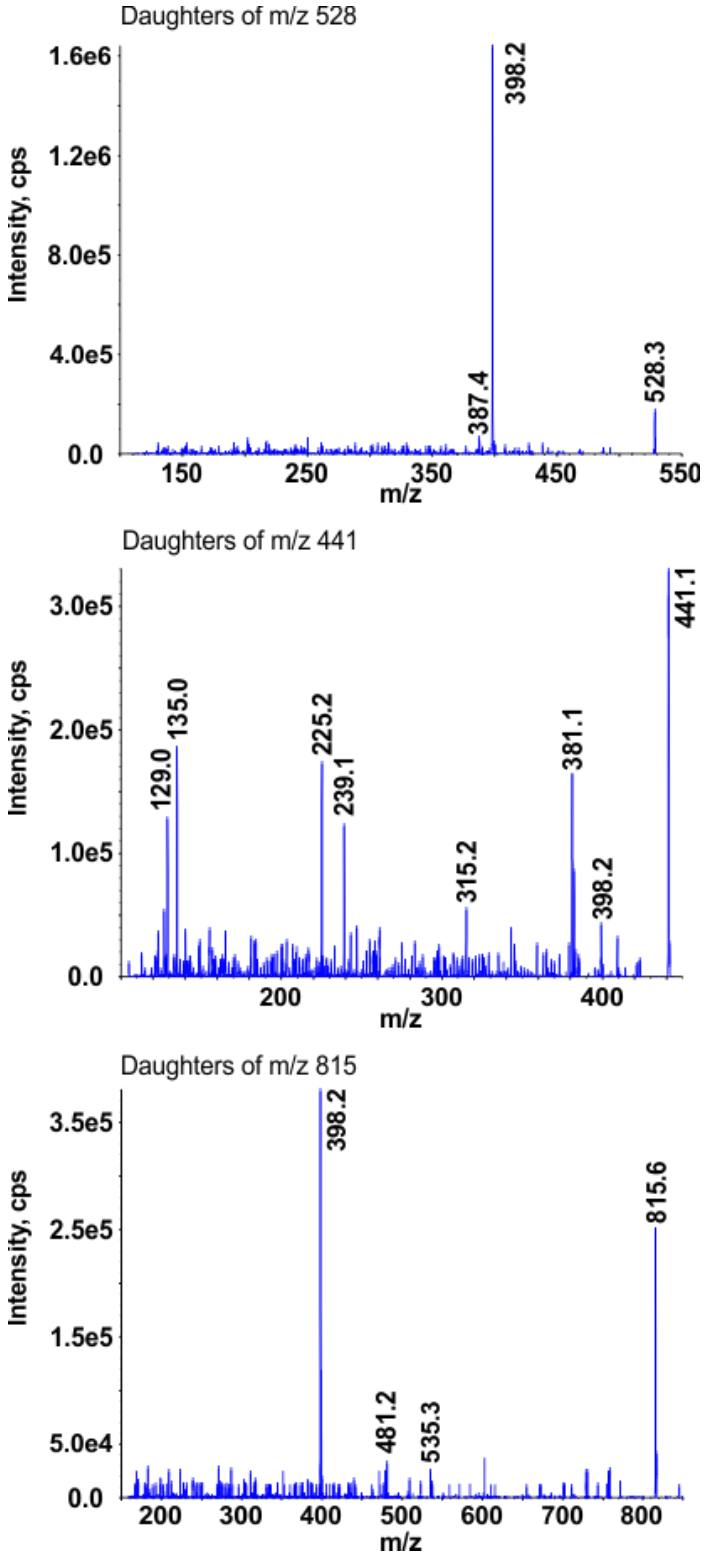


Figure 4.8 – Daughter scans for eggshell lipid extraction peaks showing NL 130

Daughter scans taken over a minimum of 16 MCA scans. Clear loss of m/z 130 is seen in d528 (top). Neither daughter scans 441 (middle) or 815 (bottom) show clear loss of m/z 130. All spectra show unknown contaminating peak with m/z 398.

An unknown contaminating peak with m/z 398 (also seen in **Figure 4.8** and discussed previously) causes appearance of an ascaroside in eggshell lipid extraction with m/z 528. However, as seen in the other daughter scans an ion of m/z 398 is also present. The 'ascaroside' with m/z 528 only seems enriched because of the high abundance of the m/z 398 contamination in this sample. The mass spectrometer shows an abundance of m/z 598 as it easily dissociates from m/z 130 to produce m/z 398. The other daughter scans from these ions do not resemble a clear loss of m/z 130 and do not resemble ascarosides as seen in the positive control.

Similarly to ions found in the positive mode NL130 scan, the ions in the negative mode NL130 scans (not shown) do not resemble ascarosides. Instead these daughter scans show a variety of peaks all increasing in mass around a base peak of m/z 183. For the same reason as above, there is clear enrichment of m/z 313 in this spectrum.

To confirm absence of ascarosides in PCN eggshell lipid samples carbohydrate analysis was carried out to observe the ascarylose sugar unit. Due to the natural rarity of 3,6 dideoxy sugars, ascarylose can be easily picked out from other monosaccharides in a mixture. However, it first must be removed from any hydrocarbon chains or alternative modifications by acid hydrolysis. Acid hydrolysis of whole nematode samples whether J2 or cyst keeps the experiment unbiased as there is no purification or enrichment of any groups of compounds. After hydrolysis remaining hydroxyl groups must be derivatized to increase volatility of monosaccharide units for GC-MS (**Figure 4.9**).

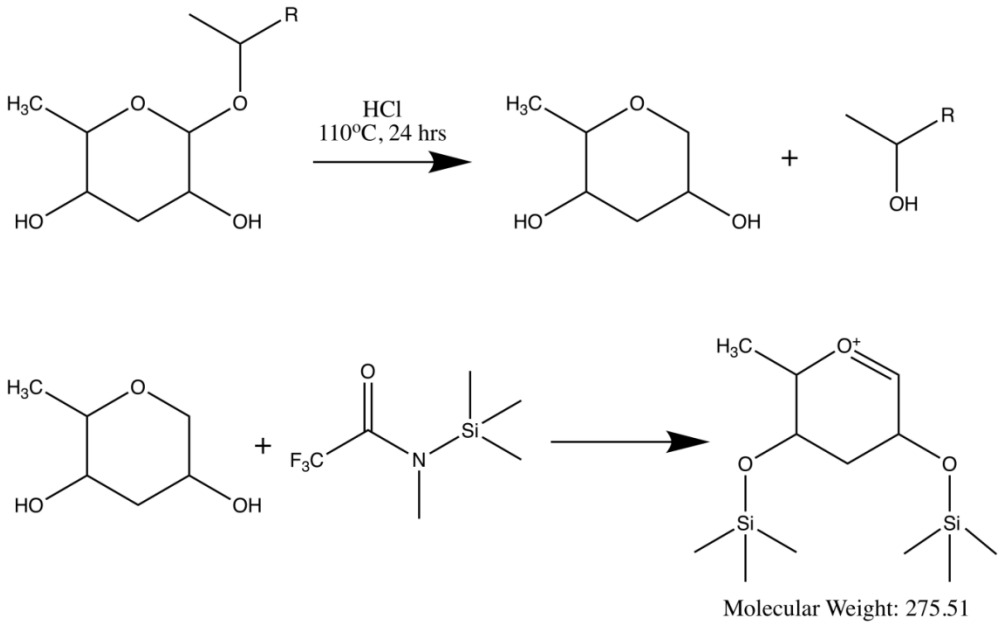


Figure 4.9 – Acid hydrolysis and TMS derivatization of ascarylose
Before GC-MS can be used to identify the presence or absence of ascarylose in nematode samples the sugar unit must be derivatized to increase volatility. Derivatization with MSTFA follows acid hydrolysis which breaks all compounds down to their base units.

The unique molecular weight of m/z 275 resulting from hydrolysis and TMS derivatization of ascarylose allows ion extractions to select any peaks containing a fragment of m/z 275. An additional ion of m/z 185 can be seen resulting from loss of one silyated hydroxyl group. (von Reuss *et al.,* 2017). These ions are not visible in hydrolysis and TMS derivatization of other monosaccharides unless they are also 3,6-dideoxy sugars. Similarly, most pentose and hexose sugars with hydroxyl groups on all their carbons can be selected due to presence of m/z 204 or m/z 217 diagnostic ions (Chen *et al.,* 2002). On the rare occurrence that there are other monosaccharides exhibiting the same number and arrangement of hydroxyl groups as ascarylose the retention times of these sugars will be used to discern the ascarylose. This retention time effect can be seen when separating various pentose or hexose sugars which all exhibit full hydroxylation (**Figure 4.10**). Due to the dideoxy nature of ascarylose, this ion at 204 cannot form under TMS derivatization conditions. Instead ascarylose forms a unique ion at m/z 275, which equally is not present in fully hydroxylated monosaccharides. **Figure 4.10 B** shows that TMS derivatised ascarylose elutes the column after 10.6 minutes.

Fragmentation of TMS derivatised ascarylose highlights characteristic peaks for TMS derivatisation are seen at m/z 73, 75, 147. Diagnostic ascarylose peaks are found at m/z 185 & 275 (**Figure 4.10 C**).

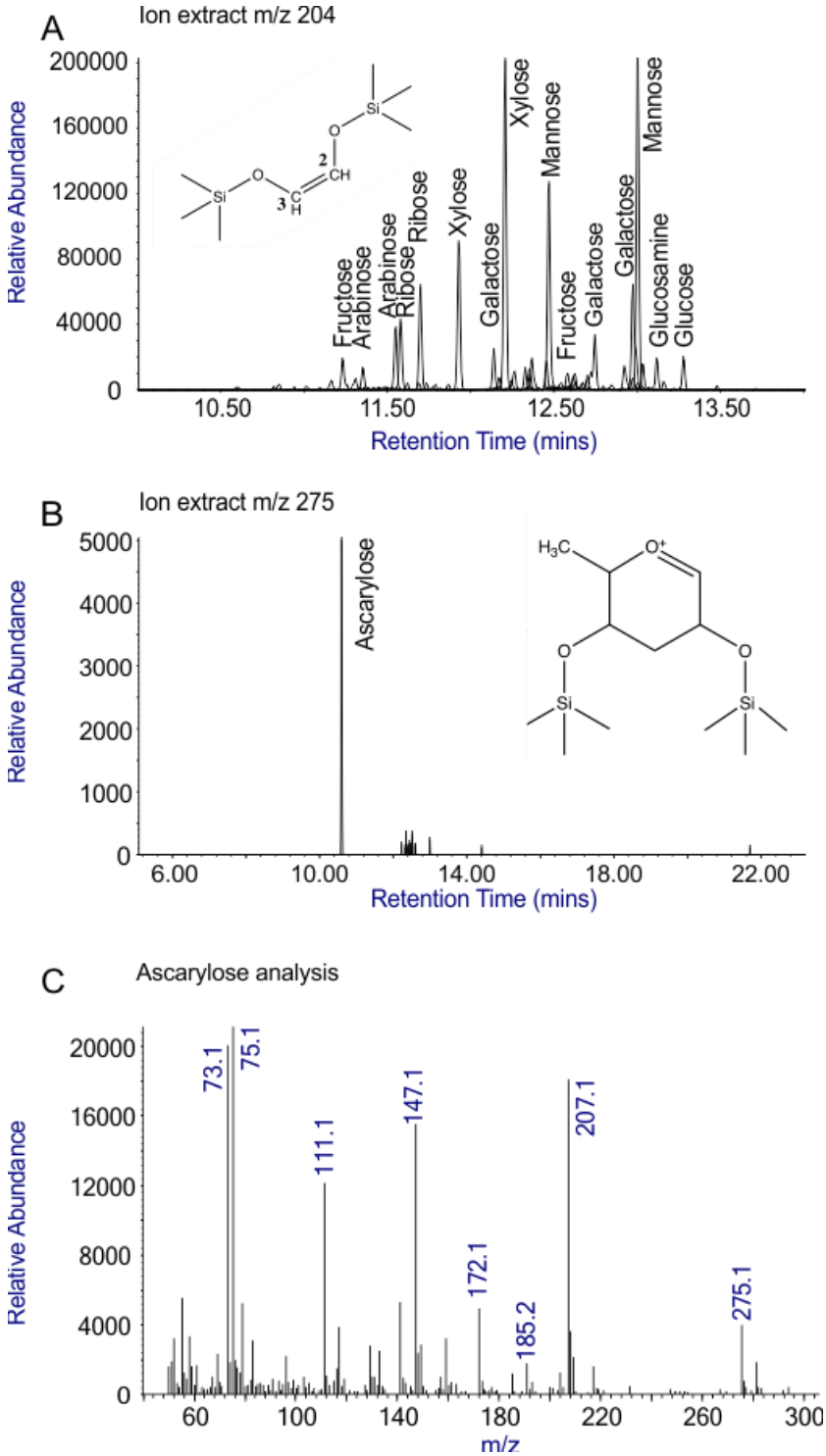


Figure 4.10 – Ascarylose analysis by GC-MS

A) Ion extract m/z 204 showing specificity of this ion in fully hydroxylated sugar standards. **B**) Ion extract m/z 275 from sugar standards showing specificity of this ion in ascarylose only. **C**) Fragmentation of ascarylose standard peak from elution at 10.6 minutes showing characteristic peaks of ascarylose TMS derivitisation at m/z 275 and m/z 185.

Acid hydrolysis of PCN juveniles and cysts will show whether ascarylose is present (**Figure 4.11, A**). It is expected that far more ascarylose will be present in cyst samples as these contain the eggshells where the ascaroside layers are hypothesised to be. Using whole, crushed cysts for this experiment is as viable as using eggshells alone as the acid hydrolysis removes any potential bias from extracting the eggshells. Acid hydrolysis of eggshell lipid extractions, although potentially biased, will aid identification of any ascarosides that have been concentrated due to extraction techniques, making them more apparent (**Figure 4.11 B**).

Ascarosides have been well documented in the model nematode *C. elegans*. However, *C. elegans* cannot be used as an *in vivo* test for identification of ascarylose following acid hydrolysis and TMS derivatization. All current research on *C. elegans* ascarosides identifies ascarosides in the liquid culture that the nematodes are grown in. From unpresented work with the model nematode for this chapter it became apparent that analysis of *C. elegans* cultured on plates is not sufficient for identifying ascarosides. This is likely due to the nematode excreting ascarosides into their surroundings and only producing ascarosides as and when they are required and in small quantities. However, as PCN juveniles are hatched in liquid hatching factor (TRD) the remaining hatching factor, after removal of the nematodes, can be analysed for ascarosides to confirm the ability of PCN to produce ascarosides (**Figure 4.11 C**).

No ascarylose was detectable in *G. rostochiensis* cysts, juveniles, eggshell lipid extractions or liquid culture used to hatch juveniles.

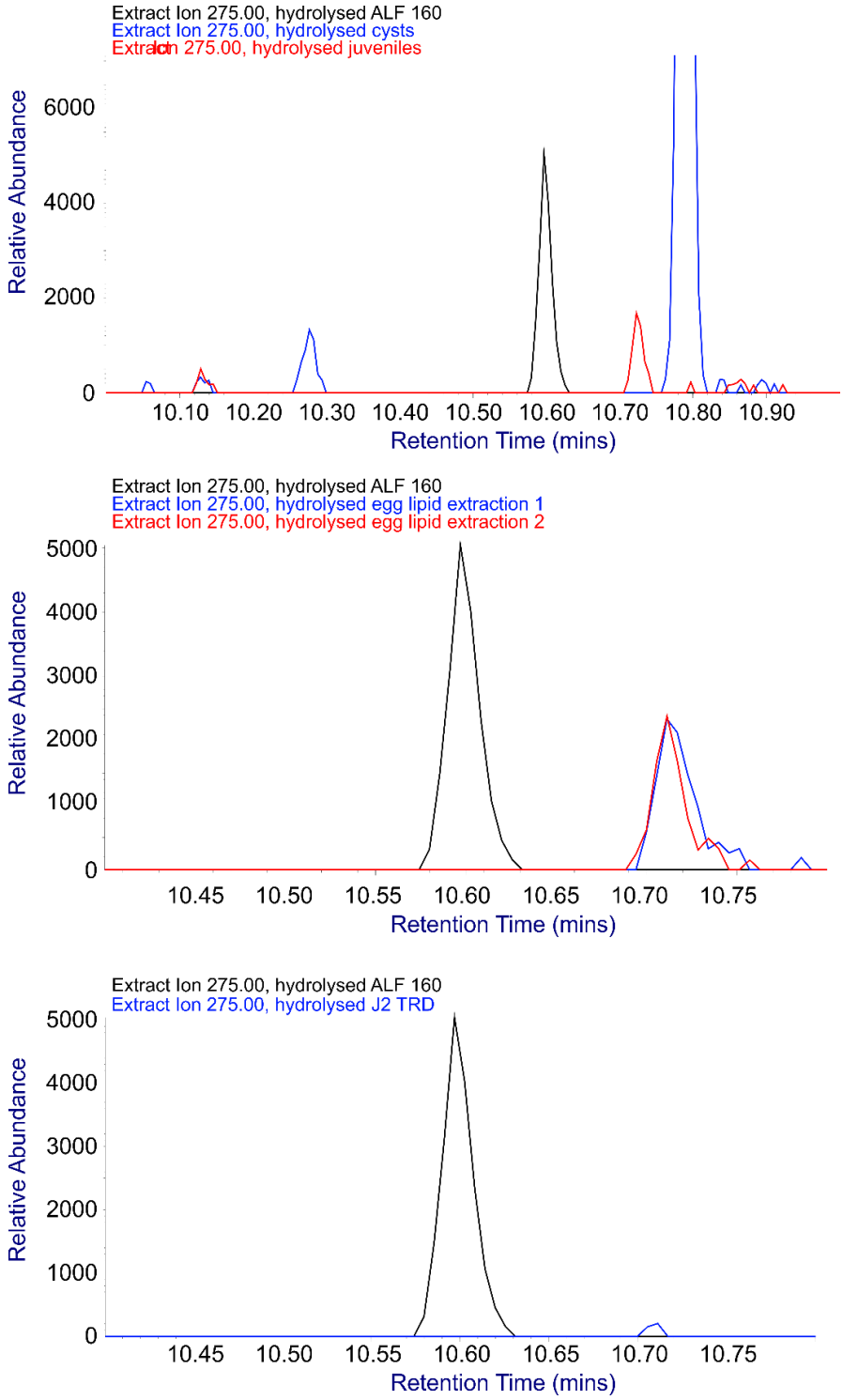


Figure 4.11– GC-MS of PCN carbohydrate samples

A) Extract ion m/z 275 comparing acid hydrolysed ascaroside control and acid hydrolysed cyst and juvenile samples. B) Extract ion m/z 275 comparing acid hydrolysed ascaroside control and acid hydrolysed eggshell lipid extractions. C) Extract ion m/z 275 comparing acid hydrolysed ascaroside control and acid hydrolysed TRD taken after use for PCN hatching.

4.4 Discussion

The previously identified and localised eggshell protein GROS_g03104 (**chapter 3**) showed affinity for PIP_3. Binding was shown with and without the presence of calcium. However, when calcium was present, faint binding could be seen to other phosphatidylinositol species. This reflects the annexin function as a calcium dependent lipid binding protein. Although it was not fully confirmed, there is a strong likelihood that PCN eggshells contain PIP_3 46:0. Phosphatidylinositol 3,4,5-trisphosphates are noted for their downstream signalling capabilities, in terms of IP4 and DAG, a role which here could be used to pass signals from or to the dormant juvenile after the eggshell has been exposed to hatching factors. Annexins have a strong affinity for acidic lipid species, usually binding to PA and PS. It is possible that GROS_g03104 was attracted by the abundant highly acidic PI and phosphate bound PI species. Therefore, this blot may not fully reflect the native binding of this annexin.

Aside from PI species, analysis of eggshell lipid extractions identified various tri- and di-acylglycerols. This reinforces previous work that acylglycerols make up 55-75% of neutral lipids identified in PCN eggshells (Gibson *et al.* 1995).

Phosphatidic acid species were also identified in lipid extractions. PA is maintained in low levels within the cell as it is the base for many other glycerophospholipids and so applied elsewhere quickly. Therefore, the presence of PA in eggshell lipid samples would suggest a role for the lipid within the eggshell itself. PA species have been previously shown to be aggregated by calcium (Blackwood *et al.* 1997, Koter *et al.* 1978). This aggregation allows for changing membrane shape to form PA microdomains, altering membrane structure (Stillwell 2016).

Glycolipid presence in *C. elegans* eggshells has been strongly inferred in previous work. These glycolipids were assumed to be ascarosides due to their presence in other nematode eggshells and the increase in expression of fatty acid and sugar synthesis and modifying enzymes during eggshell production (Olson *et al.,* 2012). However, no ascarosides have yet been identified in any *C. elegans* eggshell lipid extractions. The identification of glycolipids here (m/z 714.5, m/z 686.4 and amine containing glycolipid

at m/z 785.5) would explain the need for fatty acid and sugar modifying enzymes during eggshell formation and would clarify a lack of detectable eggshell ascarosides. These lipids are also neutral lipids, backing up that neutral lipids are a major component of many species of nematode eggshells.

Interestingly, analysis of FAMEs highlighted the presence of C16:0 and C18:0 with relatively little C20:0. However, in PCN juveniles C18:1, C20:4 and C20:1 make up the majority of fatty acids (**7.9 PCN juvenile FAMEs**), similar fatty acid traces were also recorded by Holz *et al.* (1998). Furthermore, longer chained C22 is barely visible in lipid extractions from juveniles whereas it was more prominent than C20 in eggshell FAMEs. FAMEs identified here support previous work showing that nematode eggshell lipids are mostly longer chained and saturated. The increased strength in intermolecular forces offered by increased larger, saturated fatty acids in PCN eggshells possibly reflects the impermeable properties of the eggshells.

Supporting previous work from Bartley *et al*., (1996) and Zagoriy *et al*., (2010), the in house synthesised ascaroside #18 showed distinctive neutral loss of m/z 130. Using detection methods set up to accommodate for this neutral loss of m/z 130 did not identify any ascarosides in eggshell lipid extractions, agreeing with Clarke & Hennessy (1977). Unlike the work carried out by Clarke and Hennessy investigations into the presence of the characteristic ascarylose sugar took place.

Acid hydrolysis of whole juveniles or cysts removes the need for specific lipid extractions or purification, therefore removing chances of bias. Comparing ascarylose from the synthetic ascaroside with acid hydrolysed juveniles and cysts showed no ascarylose to be present in any PCN samples, confirming tandem mass spectrometry work. Similarly, ascarylose was not found in samples of hydrolysed eggshell lipid extractions.

Up to this point all work on ascarosides in *C. elegans* has not identified ascarosides within the nematode itself, instead extracting ascarosides from the liquid culture that the nematodes were grown in. A similar protocol was used to extract any ascarosides

from the TRD that PCN were hatched in. However, once again, no ascarylose or ascarosides were discovered.

The lack of ascarosides in PCN eggshells is somewhat unexpected. *Ascaris* spp. must survive the harsh conditions of the intestine to carry out their lifecycle. The use of ascarosides to protect the developing juvenile therefore makes sense. Although PCN eggshells can also be subjected to harsh conditions, eggshell ascarosides are not the only difference between *Ascaris* and other nematode species. For example, sterols are an abundant group of molecules in *Ascaris* lipids (Fairbairn 1957). However, within cyst nematodes sterols are almost non-existent (Chitwood *et al.,* 1985, Orcutt *et al.,* 1978). The lack of ascarylose in samples extracted from juveniles comes as more of a surprise due to the apparent abundance of the glycolipid within the kingdom Nematoda. Work on *C. elegans* highlights the possibility that ascarosides act as hermaphrodite attracters when produced by males (Srinivasan *et al.* 2012). In PCN it is understandable that this trait would be lost due to the entirely different mode of reproduction. For example, in other cyst nematodes vanillic acid has been identified as a sex pheromone (Jaffe *et al.*, 1989). Nematode species that do not contain ascarosides are not unheard of. Previous work identifying conservation of ascarosides across nematode species showed that the only tested PPN species (*Pratylenchus penetrans*) did not contain any ascarosides (Choe *et al.* 2012). However, other research shows ascarosides have been associated with the excreted metabolome of *Meloidogyne* spp., *H. glycines* and *Pratylenchus brachyurus*. Of the identified ascarosides, ascr#18 in particular was found to act as a nematode-associated molecular pattern, detectable by *Arabidopsis*. Upon detection of the ascaroside, the plant induces a conserved immune response. Solanaceous species were shown to be extremely sensitive to ascr#18 detecting the molecule at 1-10 nM concentrations compared to micromolar quantities required to induce the same response in *Arabidopsis* (Manosalva *et al.*, 2015). It is possible that PCN therefore lost the ability to produce ascarosides to reduce the likelihood of being detected by their host, avoiding the host immune response.

A major difficulty with these analyses was availability of nematode material. Ideally the use of more nematode eggshells would allow sufficient lipid extraction and consequently greater signal detection under various MS conditions. Although full, comprehensive analysis of all PCN eggshell lipids was not likely achieved here, the work does identify a range of eggshell lipid species that are present. The techniques used here can form the basis for similar work in the future.

4.5 References

BARTLEY, J.P., BENNETT, E.A. & DARBEN, P.A. 1996. Structure of the ascarosides from Ascaris suum. *Journal of natural products*, **59**, 921–926.

BIRD, A.F. 1968. Changes Associated with Parasitism in Nematodes. III. Ultrastructure of the Egg Shell, Larval Cuticle, and Contents of the Subventral Esophageal Glands in Meloidogyne javanica, with Some Observations on Hatching. *Journal of Parasitology*, **54**, 475.

BLACKWOOD, R.A., SMOLEN, J.E., TRANSUE, A., HESSLER, R.J., HARSH, D.M., BROWER, R.C. & FRENCH, S. 1997. Phospholipase D activity facilitates Ca2+-induced aggregation and fusion of complex liposomes. *American journal of physiology*, **272**, C1279-85.

BUTCHER, R.A., FUJITA, M., SCHROEDER, F.C. & CLARDY, J. 2007. Small-molecule pheromones that control dauer development in Caenorhabditis elegans. *Nature Chemical Biology*, **3**, 420–422.

CHEN, J., YAGER, C., REYNOLDS, R., PALMIERI, M. & SEGAL, S. 2002. Erythrocyte galactose 1-phosphate quantified by isotope-dilution gas chromatography-mass spectrometry. *Clinical chemistry*, **48**, 604–612.

CHITWOOD, D.J., HUTZELL, P.A. & LUSBY, W.R. 1985. Sterol Composition of the Corn Cyst Nematode, Heterodera zeae, and Corn Roots. *Journal of Nematology,* **17**, 64-68.

CHITWOOD, D.J. & LUSBY, W.R. 1991. Metabolism of plant sterols by nematodes. *Lipids*, **26**, 619–627.

CHOE, A., VON REUSS, S.H., KOGAN, D., GASSER, R.B., PLATZER, E.G., SCHROEDER, F.C. & STERNBERG, P.W. 2012. Ascaroside Signalling Is Widely Conserved among Nematodes. *Current Biology*, **22**, 772–780.

CHRISTIE, W.W. & HAN, X. 2010. *Lipid analysis : isolation, separation, identification and lipidomic analysis*. Oily Press, an imprint of PJ Barnes & Associates, 428 pp.

CLARKE, A.J. & HENNESSY, J. 1977. Lipids of Globodera rostochiensis. *Rothamsted Experimental Station Report for 1976 Part 1*, **1**, 205 Available at: http://www.era.rothamsted.ac.uk/eradoc/article/ResReport1976p1-199-221.

CRONIN, D., MOËNNE-LOCCOZ, Y., DUNNE, C. & O 'GARA, F. 1997. Inhibition of egg hatch of the potato cyst nematode Globodera rostochiensis by chitinase-producing

bacteria. *European Journal of Plant Pathology*, **103**, 433–440.

DAUTOVA, M., GOMMERS, F., ABAD, P., SMANT, G., BAKKER, J. & ROSSO, M. 2001. Single pass cDNA sequencing - a powerful tool to analyse gene expression in preparasitic juveniles of the southern root-knot nematode Meloidogyne incognita. *Nematology*, **3**, 129–139.

DUCEPPE, M., LAFOND-LAPALME, J., PALOMARES-RIUS, J.E., SABEH, M., BLOK, V., MOFFETT, P. & MIMEE, B. 2017. Analysis of survival and hatching transcriptomes from potato cyst nematodes, Globodera rostochiensis and G. pallida. *Scientific Reports*, **7**, 3882.

FAIRBAIRN, D. 1957. The biochemistry of Ascaris. *Experimental Parasitology*, **6**, 491–554.

FAIRBAIRN, D. 1961. The in vitro hatching of Ascaris lumbricoides eggs. *Canadian Journal of Zoology*, **39**, 153–162.

FRASER, A.L. 2017. Target Elucidation of Novel Trypanosomatid Inhibitors. *Target Elucidation of Novel Trypanosomatid Inhibitors*. University of St. Andrews. PhD Thesis.

GIBSON, D.M., MOREAU, R.A., MCNEIL, G.P. & BRODIE, B.B. 1995. Lipid Composition of Cyst Stages of Globodera rostochiensis. *Journal of Nematology*, **27**, 304–311.

HOLZ, R.A., WRIGHT, D.J. & PERRY, R.N. 1998. Changes in the lipid content and fatty acid composition of 2nd-stage juveniles of Globodera rostochiensis after rehydration, exposure to the hatching stimulus and hatch. *Parasitology*, **116**, 183–190.

JAFFE, H., HUETTEL, R.N., DEMILO, A.B., HAYES, D.K. & REBOIS, R. V. 1989. Isolation and identification of a compound from soybean cyst nematode,Heterodera glycines, with sex pheromone activity. *Journal of Chemical Ecology*, **15**, 2031–2043.

JEONG, P., JUNG, M., YIM, Y., KIM, H., PARK, M., HONG, E., LEE, W., KIM, Y.H., KIM, K. & PAIK, Y. 2005. Chemical structure and biological activity of the Caenorhabditis elegans dauer-inducing pheromone. *Nature*, **433**, 541–545.

KOTER, M., DE KRUIJFF, B. & VAN DEENEN, L.L.M. 1978. Calcium-induced aggregation and fusion of mixed phosphatidylcholine-phosphatidic acid vesicles as studied by 31P NMR. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, **514**, 255–263.

KRUSBERG, L.R., HUSSEY, R.S. & FLETCHER, C.L. 1973. Lipid and fatty acid composition of females and eggs of Meloidogyne incognita and M. arenaria. Pergamon Press, 335-

341 pp.

MANOSALVA, P., MANOHAR, M., VON REUSS, S.H., CHEN, S., KOCH, A., KAPLAN, F., CHOE, A., … KLESSIG, D.F. 2015. Conserved nematode signalling molecules elicit plant defenses and pathogen resistance. *Nature Communications*, **6**, 7795.

OLSON, S.K., GREENAN, G., DESAI, A., MÜLLER-REICHERT, T. & OEGEMA, K. 2012. Hierarchical assembly of the eggshell and permeability barrier in C. Elegans. *Journal of Cell Biology*, **198**, 731–748.

ORCUTT, D.M., FOX, J.A. & JAKE, C.A. 1978. The Sterol, Fatty Acid, and Hydrocarbon Composition of Globodera solanacearum, 264-268 pp.

PANDA, O., AKAGI, A.E., ARTYUKHIN, A.B., JUDKINS, J.C., LE, H.H., MAHANTI, P., COHEN, S.M., STERNBERG, P.W. & SCHROEDER, F.C. 2017. Biosynthesis of Modular Ascarosides in C. elegans. *Angewandte Chemie International Edition*, **56**, 4729–4733.

PERRY, R.N. 2002. Hatching. *In* Lee, D.L., *ed. The Biology of Nematodes*. Taylor & Francis, 147–165.

PERRY, R.N. & CLARKE, A.J. 1981. Hatching mechanisms of nematodes. *Parasitology*, **83**, 435–449.

PERRY, R.N., WHARTON, D.A. & CLARKE, A.J. 1982. The structure of the egg-shell of Globodera rostochiensis (Nematoda: Tylenchida). *International Journal for Parasitology*, **12**, 481–485.

PERRY, R.N., KNOX, D.P. & BEANE, J. 1992. Enzymes released during hatching of Globodera rostochiensis and Meloidogyne incognita. *Fundam. appl. NemalOl*, **15**, 283–288.

RIOU, M., GRASSEAU, I., BLESBOIS, E. & KERBOEUF, D. 2007. Relationships between sterol/phospholipid composition and xenobiotic transport in nematodes. *Parasotol. Res*, **100**, 1125–1134.

ROGERS, W.P. 1958. Physiology of the Hatching of Eggs of Ascaris lumbricoides. *Nature*, **181**, 1410–1411.

SRINIVASAN, J., VON REUSS, S.H., BOSE, N., ZASLAVER, A., MAHANTI, P., HO, M.C., O'DOHERTY, O.G., EDISON, A.S., STERNBERG, P.W. & SCHROEDER, F.C. 2012. A Modular Library of Small Molecule Signals Regulates Social Behaviors in Caenorhabditis elegans Sengupta, P., *ed. PLoS Biology*, **10**.

STILLWELL, W. 2016. *Membrane Polar Lipids*. Elsevier, 63-87 pp.

TARR, G.E. & FAIRBAIRN, D. 1973. Conversion of ascaroside esters to free ascarosides in fertilized eggs of Ascaris suum (nematoda). *The Journal of parasitology*, **59**, 428–433.

VON REUSS, S.H., BOSE, N., SRINIVASAN, J., YIM, J.J., JUDKINS, J.C., STERNBERG, P.W. & SCHROEDER, F.C. 2012. Comparative Metabolomics Reveals Biogenesis of Ascarosides, a Modular Library of Small-Molecule Signals in *C. elegans*. *Journal of the American Chemical Society*, **134**, 1817–1824.

VON REUSS, S.H., DOLKE, F. & DONG, C. 2017. Ascaroside Profiling of Caenorhabditis elegans Using Gas Chromatography–Electron Ionization Mass Spectrometry. *Analytical Chemistry*, **89**, 10570–10577.

ZAGORIY, V., MATYASH, V. & KURZCHALIA, T. 2010. Long-Chain O-Ascarosyl-alkanediols Are Constitutive Components of Caenorhabditis elegans but Do Not Induce Dauer Larva Formation. *Chemistry & Biodiversity*, **7**, 2016–2022.

5. Identification and characterisation of juvenile surface proteins

5.1 Introduction

After eclosion, the nematode no longer has the protective eggshell to shield the vulnerable juvenile from environmental extremes or predation. The cuticle is now the main contact point between the parasite and the host, or the external environment, and therefore must be the new, adaptable, barrier able to protect the nematode from the outside world. In addition to the physical protection provided by the cuticle itself, the hypodermis also produces secreted defensive proteins onto the nematode's cuticle surface, or from here into the immediate surrounding environment (Davies & Curtis, 2011). This chapter will only focus on protection given by proteins released via the cuticle.

5.1.1 The surface coat

The surface coat is a collection of secretions found on the surface of the nematode cuticle with non-structural and often defensive properties (Davies & Curtis, 2011). Proteins found here differ from those released from the widely examined oesophageal glands as they are released into the apoplastic space instead of being secreted directly into host cells. There is strong evidence that proteins shuttled to the surface coat originate in the hypodermis (Sharon *et al.*, 2002). However, it is also possible that surface coat proteins are secreted from the gland cells surrounding the main sense organs of the nematode, the amphids, forming a surface coat as the nematode moves through the secretions (Spiegel & McClure, 1995; Hu *et al.*, 2000)

Although proteins found in the surface coat have a different function to those secreted from the oesophagus both groups of proteins share some similar properties. Both groups of proteins would be expected to be expressed at parasitic or pre-parasitic life

stages. As these are secreted proteins they are unlikely to have a predicted transmembrane domain and would most likely have a predicted signal peptide for secretion, although leaderless secretion is also widely seen across the animal kingdom (Andrei *et al.*, 1999; Bendtsen *et al.*, 2004).

5.1.2 Survival outside a host

Although *C. elegans* cannot be a model for surface coat changes in a parasitic lifestyle, it has potential utility as a model for nematode survival in the environment outside a host, including protection from attack by other pathogens. For example, surface associated proteins in *C. elegans* play a major role in the innate immune defence of the nematode. Mutants in the nucleotide sugar transporter protein (srf-3) become resistant to bacterial pathogens such as *Microbacterium nematophilum* and *Yersinia pseudotuberculosis*. Changes to this protein inhibit bacteria from adhering to the cuticle of the nematode due to altered surface glycosylation. Nematodes carrying this mutation can therefore feed and grow much quicker as they are not repressed by bacteria. However, the cuticles of nematodes carrying the mutation appear to be much weaker than those of wild type nematodes (Höflich *et al.*, 2004).

When outside a host, parasitic nematodes are vulnerable to predation. Shedding material from the surface coat offers protection by distracting potential predators away from the body of the nematode (Riddle *et al.*, 1997). The ability to rapidly turn over the surface coat also allows nematodes to evade pathogens that would bind to their cuticle such as *Pasteuria penetrans. Pasteuria* spores bind to the nematode surface, penetrate the cuticle and grow and reproduce inside the nematode (Tian *et al.*, 2007). However, if the nematode can produce a layer above this epicuticle, this could prevent cuticular attachment by the bacterial spores. This rapid surface coat turnover has been demonstrated in PCN. Antibodies raised against excretory/secretory products impaired movement of *G. rostochiensis*, but after a period of 1-2 hours, regular locomotion resumed (Fioretti *et al.*, 2002). Maintenance of motility is key for plant-parasitic nematodes to give the nematode as much time as possible to reach the host root before lipid reserves are used (Robinson *et al.*, 1985).

5.1.3 Survival in a host

Once at an appropriate feeding site, the nematode can be offered some protection from pathogens and environmental extremes by the host. This is reflected by a reduced complement of genes associated with immune function in the genomes of several unrelated plant-parasitic nematodes (Kikuchi *et al.*, 2017). However, to obtain these benefits and maintain feeding within the host, the nematode must now avoid detection. In animal-parasitic nematodes there are two main strategies used to hide from immune responses. The first uses surface associated molecules that mimic the host, allowing the nematode to camouflage itself from detection. For example, the canine roundworm *Toxicara canis*, presents a surface glycan which resembles host blood group antigens (Maizels, 2013). These surface glycoproteins are termed mucins and have been identified on the surface of a variety of nematodes, including parasitic and non-parasitic species. However, no function other than evading host adaptive immune responses has been attributed to mucins.

The second strategy relies upon release of defensive surface proteins that can directly interact with the hosts defensive immune response. *Brugia pahangi* produces a secreted Cu-Zn superoxide dismutase which acts to detoxify superoxide produced by host defensive leukocytes (Tang *et al.*, 1994). Reduction in superoxides in turn can act as an anti-inflammatory, which ultimately would allow the invasive nematode to propagate further (Serena *et al.*, 2015). Other nematode surface proteins that detoxify host induced oxidative bursts, include the *Onchocerca volvulus* surface associated glutathione-S-transferase (Ov-GST-1) (Liebau *et al.*, 1994; Sommer *et al.*, 2003).

However, not all defensive nematode surface proteins act by reducing damage from reactive oxygen species. A *Haemonchus contortus* cystatin was recently localised to the body surface of the nematode. It is thought this protein can inhibit host cathepsins, altering processing of antigens (Wang *et al.*, 2017). Other parasitic nematode cystatins have been shown to inhibit proteases ultimately reducing T-cell response to the parasite (Hartmann & Lucius, 2003).

Shedding surface coat material can also be used within a host to aid infection (Blaxter & Maizels, 1992). Both the entomopathogenic nematode *Romanomermis culicivorax* and the trichinosis inducing, *Trichinella spiralis,* shed their surface coats to remove immune defence products that have recognised and bound to the nematode (Modha *et al.*, 1999; Shamseldean *et al.*, 2007). This allows the nematode to continue its lifecycle without being detected by the host.

Although plant-parasitic nematodes do not have to avoid an adaptive immune system, they must still avoid detection by their host plant, and activation of innate immune responses, to successfully set up and maintain a feeding site. Antibodies raised against *Meloidogyne incognita* surface coat proteins cross react with material in plant cells (Gravato-Nobre *et al.*, 1999). This would suggest that like animal-parasitic nematodes, *M. incognita* exhibits some sort of host mimicry on its surface. Additionally, antibodies raised against *M. javanica* surface coats were used to show that, as with other nematode species, surface coat shedding is also used by PPN inside the host (Sharon *et al.*, 2002).

5.1.4 PCN surface proteins

In PCN, some key surface proteins have been described. A peroxiredoxin has been identified on the surface of *G. rostochiensis* that specifically metabolises hydrogen peroxide, a key element of plant defence responses (Robertson *et al.*, 2000). This is the only surface localised protein from plant-parasitic nematodes known to lack an N-terminal signal peptide, usually used for protein secretion. Classically, a signal peptide is located at the N-terminus of a protein and is an indicator for the protein to be passed through the endoplasmic reticulum and Golgi into vesicles that ultimately fuse with the cell membrane, leading to secretion of the proteins (Coleman *et al.*, 1985). Non-classical or leaderless secretion occurs across the animal kingdom and is a feature of many inflammatory cytokines and is a suggested feature for parasite surface proteins capable of aiding host infection (Nickel, 2003). However, detailed pathways for non-classical secretion are yet to be clarified.

A fatty acid and retinol binding protein (FAR-1) has also been identified on the surface of *G. pallida*. Gp-FAR-1 is capable of binding a range of fatty acids, including linolenic and linoleic acids, both of which are precursor molecules within the jasmonic acid pathway (Prior *et al.*, 2001). Binding of these lipids is thought to allow the nematode to obstruct the plant defence pathway. Surface associated FAR proteins have now been described in numerous other parasitic nematodes (Iberkleid *et al.*, 2013; Le *et al.*, 2016; Phani *et al.*, 2017). While it has never been specifically identified, it is assumed that this protein would also be present on the surface of *G. rostochiensis*.

A glutathione peroxidase (GpX) has also been identified on the *G. rostochiensis* surface. This protein has a function in metabolism of hydrogen peroxide, as well as a range of host fatty acid hydroperoxides, which are used within the plant defence signalling pathway (Jones *et al.*, 2004).

Although all of these proteins except the peroxiredoxin have a signal peptide, they all lack transmembrane domains and are expressed at nematode life cycles reflecting parasitism. Expression of all these proteins was localised to the hypodermis and antibodies to peroxiredoxin and FAR-1 were used to confirm that the protein is present on the nematode surface. These proteins function to protect the nematode from plant defences, typically by aiming to reduce the availability of potentially damaging reactive oxygen species.

5.1.5 Identifying novel proteins

Bacterial plant pathogens secrete effector proteins to aid infection of their host. These proteins can be easily characterised as they are secreted using the type 3 secretion system and have a type 3 signal sequence at the N-terminus. In the absence of the type 3 secretion system, the bacteria are unable to infect the plant (Alfano & Collmer, 2004). In oomycetes, many effectors can be identified on the basis of the presence of RxLR-EER motifs immediately downstream of a signal peptide (Whisson *et al.*, 2007). More recently, upstream promoter motifs were discovered associated with *G. rostochiensis* dorsal gland effectors, termed the DOG box motif (Eves-van den Akker *et al.*, 2016). Similarly, a motif (STATAWAARS) was found to be associated with genes, including

effectors, expressed in the gland cells in *B. xylophilus* (Espada *et al.*, 2018). In both cases, knowledge of the promoter element facilitated further prediction of novel effectors in these nematode species.

The STATAWAARS and DOG box motifs, in addition to the previously identified *C. elegans* muscle tissue promoter (GuhaThakurta *et al.*, 2004), provide evidence that nematodes commonly use tissue specific promoters. Therefore, it is possible that similar motifs responsible for hypodermal expression of proteins could also exist. However, not enough genes known to be expressed in the hypodermis have been discovered to be able to reliably search for upstream sequences. This is possibly due to the difficulties with collecting proteins that are released to the cuticle surface of the nematode. If such an upstream motif was to be found, more surface coat proteins, including proteins acting to counteract host defence responses, could be identified, further developing the understanding of parasite-host interactions.

5.1.6 Chapter 5 Aims

The aim of this chapter was to identify proteins present on the surface of PCN juveniles, including those that suppress host defence responses. This required a novel method for extracting nematode surface proteins to be developed.

5.2 Methods

5.2.1 Biological materials

Globodera rostochiensis juveniles were collected using the techniques described in section **2.1.2**. Collected nematodes were used immediately and were never stored for longer than 24 hours.

5.2.2 Surface protein extraction methods

A variety of extraction methods were used to check for consistency between protein extractions. All surface protein extractions were carried out on a minimum of approximately 10,000 nematodes to give strong peptide reads/identification by mass spectrometer.

5.2.2.1 Biotin Pull-downs

Nematodes were collected as described in section **2.1.2** and washed twice in PBS. After washing, nematodes were resuspended in 0.9 mL ice cold PBS with 100 µL 1mg/mL biotin N-hydroxy succinimide (Sigma). The reaction was kept on ice, turning the tube at regular intervals for 20 minutes. Samples were pelleted out (1020 x g, 10 minutes) of the biotin solution and resuspended in 2 mg/mL glycine to bind to any unbound biotin. Biotinylated nematodes were removed from the glycine solution and fixed in 4% paraformaldehyde overnight at 4 °C. Surface biotinylation was tagged by adding 5 µg/mL Dylight594-streptavidin conjugate to nematode samples. Reactions were kept on ice in the dark for 1 hour. Unbound streptavidin conjugates were removed by washing in PBS. Surface biotinylation was then visualised using a confocal microscope reading emission at 618 nm.

5.2.2.1.1 Extraction of Biotinylated Proteins

Surface biotinylated nematodes were homogenised using a bead beater and 0.5 mm beads. 0.25 µL of magnetic streptavidin beads (Dynabeads MyOne Streptavidin, ThermoFisher) were added to the nematode homogenates and left rotating for 3 hours at 4 °C. Magnetic beads were separated from the homogenate using a magnetic rack and the beads were washed several times in PBS. The beads were resuspended in

20 mM ammonium bicarbonate and sent for mass spectrometry analysis as described in Section **2.2.5**.

5.2.2.2 Sodium Dodecyl-Sulphate Surface Coat Removals

This method was first described by Spiegel *et al.* (1996) for the removal of nematode surface coats. However, the method was adapted here to increase specificity between secreted and excreted proteins by first surface labelling the samples using the techniques described in **5.2.2.1.** After biotinylating, samples were incubated for 1 hour at 25 °C in a solution of 1% sodium dodecyl-sulphate (SDS) in PBS with 1% protease inhibitor cocktail (cOmplete^{Tm}, Roche). Samples were centrifuged at 1020 x g for 10 minutes. Biotinylated proteins were pulled down from the supernatant as described in **5.2.2.1.1.** However, the streptavidin beads were eluted, concentrated and run into a polyacrylamide gel to remove any SDS (**2.4.3**). The gel was stopped and stained with Coomassie blue before separation of the proteins could occur. Proteins were excised from the gel using a sterile blade and sent for mass spectrometry analysis.

5.2.2.3 Solvent Surface Coat Removals

Solvents were used to extract the surface coat in a similar way to the SDS surface coat removal technique. Solvents were of analytical HPLC grade and glassware was used where possible. Nematode samples were incubated in 80:20 MeOH:H_2O at 4 °C on a shaking platform for 24 hours. Nematodes were removed from the solvents by running the solutions through a 20-micron mesh sieve. Nematodes on the sieve were then checked under the microscope for any signs of cuticle lysis. Increased specificity for surface proteins was achieved by surface labelling proteins using techniques described in section **5.2.2.1**. Samples were dried down under a nitrogen gas line and resuspended in 20 mM ammonium bicarbonate for analysis.

5.2.3 *in situ* hybridisation

Localisation of gene expression was carried out using *in situ* hybridisation. Polymerase chain reactions to amplify gene fragments for *in situ* hybridisation used GoTaq polymerase and followed the conditions described in section **2.3.4**. Once DNA fragments had been amplified, they were cloned into the pGEM-T easy vector following steps

described in section **2.3.7**. Primers synthesised for cloning of *in situ* probes can be found in **7.1 Primer table, use and Tm**.

5.2.3.1 DIG-labelled Probes

Single stranded DNA antisense probes were used for *in situ* hybridisation and incorporated digoxigenin labelled UTP in place of TTP. Probes were generated using an asymmetric PCR reaction. These reactions contained 4 µL of 5 x GoTaq buffer (Promega), 1.2 µL of 25 mM $MgCl_2$, 1.5 µL DIG-labelled dNTPs, 4 µL of 10µM sequence specific reverse primer, 1 µL Taq polymerase and 50 ng of PCR product template. The reaction mixture was then brought up to a final volume of 20 µL with dH_2O. The same PCR cycle conditions described in section **2.3.4** were used (primer specific annealing temperatures found in **7.1 Primer table, use and Tm**. Gel electrophoresis was used to assess incorporation of DIG-dNTPs as seen by a smear with some distinct larger bands on the gel. Negative control probes were created by using 4 µL of sequence specific forward primer in place of the reverse primer. Probes were stored at -20 ˚C until use.

5.2.3.2 Cutting of Nematodes

Juveniles that were collected for *in situ* hybridisation were fixed in 4% paraformaldehyde (in PBS) for 24 hours before the cutting process. Nematodes were concentrated in a micro centrifuge tube, resuspended in 0.4% paraformaldehyde and transferred to a glass slide. Cutting the juveniles into segments was achieved by scraping a single edge razor blade attached to a vibrating aquarium air pump across the surface of the glass slide. Once sufficiently cut, as judged by examination under a binocular microscope, the nematode segments were transferred to a fresh micro centrifuge tube.

5.2.3.3 Permeabilisation, Hybridisation & Staining

Nematode fragments were washed twice in M9 buffer pelleting the fragments by centrifuging at 5900 x g for 2 minutes between washes. After the final wash the segments were resuspended in 0.5 mL proteinase-K solution (0.5 mg/mL) and incubated at room temperature for 30 minutes on a rotator. Proteinase-K was removed by washing once with M9 buffer after which the nematode pellet was placed on deep frozen ice for 15 minutes. The nematode fragments were then resuspended for 30 seconds in 1 mL of

methanol stored at -20 ˚C. Nematode sections were pelleted by centrifuging at 15700 x g for 1 minute and were resuspended for 1 minute in 1 mL acetone which was stored at -20 ˚C. Once again fragments were centrifuged to form a pellet and acetone was removed until 100 µL of supernatant remained. Nematodes were then rehydrated by gradually adding DEPC-treated dH_2O until the volume in the tube was 200 µL. The supernatant was removed after further centrifugation and 500 µL of hybridisation buffer (**7.12 Buffer recipes**) was added to remove any remaining acetone. Nematode fragments were pelleted out of the hybridisation buffer and 150 µL of fresh hybridisation buffer was added per hybridisation reaction. Cut nematode fragments were evenly distributed between new 0.5 mL micro centrifuge tubes.

Nematode segments were pre-hybridised by rotating the sample tubes at 50 ℃ for 15 minutes. Whilst the nematode segments were undergoing pre-hybridisation, the DIG-labelled DNA probes (2.4.1) were heat denatured for 10 minutes at 100 ˚C before being cooled directly on ice. Denatured probes were transferred to the corresponding nematode suspension in the micro centrifuge tubes. Samples were left rotating at 50 ℃ overnight to allow the probes and nematode mRNA to hybridise.

After overnight incubation samples were washed three times with 4x SSC (**7.12 Buffer recipes**) at 50 ˚C, each wash lasted 15 minutes. This was followed by three further 20 minute washes with 0.1 x SSC/ 0.1% SDS, also at 50˚C. Any remaining SDS was removed by a wash with maleic acid buffer (**7.12 Buffer recipes**). Any non-specific binding was blocked by incubating the nematode fragments in 1 x blocking reagent for 30 minutes. Nematode fragments were then incubated in a 1:1000 dilution of anti-digoxigenin-Alkaline phosphatase conjugated antibody for 2 hours at room temperature. Three 15 minute washes in maleic acid buffer were used to remove any unbound antibodies and avoid unspecific background staining.

Nematodes were stained by incubating in a mixture of Nitro Blue tetrazolium and X-phosphate staining solution (3.4 µL NBT and 3.5 µL X-phosphate per 1 mL alkaline phosphatase detection buffer) overnight. Staining was halted by washing the nematodes twice in 0.01% Tween-20 in dH_2O.

5.2.4 Immunolocalisation

Immunolocalisation was carried out as described in section **2.5**. Anti Gp-FAR-1 was obtained from the nematology lab in the James Hutton Institute. This antiserum could also be used for detection of *Globodera rostochiensis* FAR-1 due to the strong similarity (89% at amino acid level) between the proteins from *G. pallida* and *G. rostochiensis*.

5.2.5 Expressing surface proteins for functional studies

Due to the large numbers of rare codons in the genes encoding candidate surface proteins, the range of bacterial cells that can be used for expressing these proteins is limited. Furthermore, repeating rare codons of the same type mean that even codon-optimised cells struggle to express candidate proteins. For these reasons a new vector was created allowing the use of plants for expression of the proteins.

ApSiPR is a gateway vector that carries a gene providing spectinomycin resistance. The vector fuses an N-terminal apoplastic targeting signal peptide to the protein of interest with further C-terminal fusion to RFP for imaging or purification. Further information and sequence details can be found in **7.2 ApSiPR vector**. The vector was created from a pH7RWG2 backbone with a signal peptide taken from a pK7FWG2 vector that was modified to contain an apoplast targeting signal peptide used by Karimi *et al.* (2002).

5.2.6 Plant based studies

Nicotiana benthamiana was used to express identified nematode proteins. Fluorescently tagging proteins allowed identification of protein localisation within the plant. *N. benthamiana* was used in place of the host plant as it is a distant relative of PCN host plants, but has several practical advantages compared to tomato and potato; most notably it can be used for transient expression of material in leaves using *Agrobacterium tumefaciens*.

Genes encoding identified surface proteins that were tested in plants were amplified using KOD polymerase reactions as described in section **2.3.4**. Gene specific forward primers lacked the signal peptide and included the Kozak sequence (ACCATG). Two reverse primers were synthesised for each gene, one including a stop codon and one

without the stop codon. This allowed for addition or omission of the C-terminal fluorescent protein tag for localisation when expressed *in planta*. Genes were then cloned into the pCR8/GW/TOPO vector using the process outlined in section **2.3.8**. Genes were transferred to the ApSiPR vector using LR recombination as described in section **2.3.10**.

5.2.6.1 *Agrobacterium* Transformation

Agrobacterium tumefaciens strain GV3101 was used for *N. benthamiana* infiltrations. This strain contains an additional helper vector (pBBR1MCS5-VIRG0N54D) that expresses the *Agrobacterium* VirG virulence gene. The protein encoded by this gene aids efficiency of plant cell transformation by increasing the likelihood of t-DNA transfer events due to a VirG mutation in GV3101at position 54 as described by Van Der Fits *et al.* (2000).

Electroporation was used to transfer the recombinant Gateway vectors (**2.3.10**) into *Agrobacterium* cells. Transformed cells were incubated for 2 hours at 28 **°**C before being spread onto LB-RGS (rifampicin, gentamycin, spectinomycin) plates. The plates were left to dry and were incubated at 28 **°**C for 48 hours.

Colony PCR was used to confirm gene presence and vector accuracy in *Agrobacterium* cells. Colonies identified as containing the recombinant Gateway vector were transferred to liquid cultures (5 mL LB, containing 5 µL spectinomycin, 1.25 µL gentamycin, 1.25 µL rifampicin, concentrations in section **2.3.13**, **Table 2.2**) and were left shaking at 28 ˚C for 24 hours.

5.2.6.2 *Nicotiana benthamiana* Infiltration

Agrobacterium cells were spun down at 2164 x g for 6 minutes before being washed twice and resuspended in infiltration buffer (100 mL dH_2O, 1 ml of 1M MES, 1 mL of 1M $MgCl_2$, 250 µL of 0.1M acetosyringone). The optical density (OD_{600}) was read using a spectrophotometer. Remaining bacterial suspensions were left shaking in the dark at room temperature **for 2-3 hours.** Suspensions were diluted to a final optical density of 0.2.

A needle was used to make a small scratch on the underside of selected leaves. Dilute bacterial suspensions were then injected into the scratch site, while supporting the tissue on the topside of the leaf. Plants were left at room temperature for two days to allow the bacteria to infect and for genes encoded on plasmids to be expressed.

5.2.6.3 Apoplastic fluid extraction & purification of recombinant proteins

Apoplastic fluid extraction was carried out as described by O'Leary *et al.* (2014). Briefly, leaves were removed from the plant by cutting through the petiole with a razor blade. Leaves were rinsed in SDW to remove surface contaminants before being placed in the chamber of a 50 mL syringe. The syringe was filled with water and the plunger was used to apply positive and negative pressure to the sealed chamber, this infiltrated the water into the leaf. The leaves were then patted dry with paper towel before being wrapped around a 5 mL pipette tip. The tip and leaf were then wrapped in parafilm before being placed into a 20 mL syringe chamber. The syringe was placed into a 50 mL centrifuge tube and centrifuged at 4 °C for 10 minutes at 1000 x g. Apoplastic fluid was transferred from the centrifuge tube to 2 mL Eppendorf tubes. Samples were centrifuged at 15000 x g for a further 5 minutes to remove cell debris.

Recombinant proteins were pulled down from the apoplastic fluid using the RFP tag. Apoplastic fluid was diluted 1:1 with binding solution (10 mM Tris.HCl pH 7.5, 150mM NaCl, 0.5 mM EDTA and protease inhibitors. Apoplastic fluid was incubated at 4 °C with RFP trap_MA (Chromotek) for 2 hours. Beads were magnetically separated, washed and eluted using 50 µL of 0.2 M glycine at pH 2.5 for 30 seconds. The supernatant was then neutralised with Tris base pH 10.4 (5 µL of 1 M stock). Purified proteins were run on a poly-acrylamide gel to check purity (**2.4.3**).

5.2.7 Confocal Imaging of Transformed Plants

Imaging of transformed tissues was carried out using a Zeiss LSM710 laser scanning microscope mounted on an AxioImager z2 motorised upright microscope. *N. benthamiana* strain CB157 was used for these experiments as this contains a nuclear marker allowing easier localisation. Plants were infiltrated with *Agrobacterium* at OD_{600}

0.2. expressing either the ApSiPR vector, for apoplastic expression, or the pK7FWG2 vector for subcellular expression.

5.2.8 ROS assays

ROS burst assays were used to test changes to plant defence responses generated in response to bacterial flg-22 in the presence of recombinant nematode surface proteins. WT *N. benthamiana* was infiltrated with *Agrobacterium* at OD_{600} 0.3. Leaf discs of plants infiltrated with empty *Agrobacterium* GV3101, GFP and RFP controls, pK7FWG2-GENE::STOP and ApSiPR-GENE::STOP were cut 24 hours after infiltration. The leaf discs were loaded into a 96 well plate and left to rest, protected from light, overnight in 150 µL dH_2O.

The following morning, the water was removed from the wells and replaced with ROS assay mixture (0.5 mM L-012, 20 µg/mL horseradish peroxidase and 100 nM flg-22). Controls containing no flg-22 were used to test that no other factors were initiating a ROS burst.

The 96 well plate was placed immediately into a Varioskan LUX, luminescence plate reader (Thermofisher). Readings at 450 nm were taken every 500 ms for about 1 hour. All plates were repeated in triplicate with 16 replicates per plate.

5.2.8.1 Inducing ROS bursts in N. benthamiana with oligogalacturonides

Oligogalacturonides (GAT114, Elicityl) were purchased to mimic cell wall breakdown products. *Arabidopsis* recognises these oligogalacturonides as damage associated molecular patterns (DAMPs) and initiates a ROS burst in response (Shah *et al.*, 2017).
Wild type *N. benthamiana* was infiltrated with *Agrobacterium* at OD_{600} 0.3. Leaf discs of plants infiltrated with empty *Agrobacterium*, GFP and RFP controls, pK7FWG2-1153::STOP and ApSiPR-1153::STOP were cut 24 hours after infiltration. The leaf discs were loaded into a 96 well plate and left to rest, protected from light overnight in 150 µL dH_2O.

The water was removed from the wells and replaced with ROS assay mixture (0.5 mM L-012, 20 µg/mL horseradish peroxidase and 500 µg/mL oligogalacturonides). ROS assays were carried out as above (**5.2.8**).

5.2.9 Ellman's assay for the detection of free thiols

An Ellman's assay was carried out to show whether conserved cysteines in GROS_g02583 form disulphide bridges within the protein or whether thiol groups are left free.

GROS_g02583 fused to RFP in the ApSiPR vector was expressed in *N. benthamiana* using *Agrobacterium* mediated transformations as described in section **5.2.6.2**. Apoplastic fluid containing recombinant GROS_g02583 was then extracted using the protocol described in section **5.2.6.3**. The RFP tag does not contain cysteine and so will not affect the binding of Ellman's reagent.

A set of cysteine hydrochloride monohydrate standards were created (0-1.5 mM) in 0.1 M sodium phosphate (pH 8.0) with 1 mM EDTA. Ellman's reagent was diluted to 4 mg/mL. Each test reaction contained 50 µL Ellman's reagent in 2.5 mL sodium phosphate buffer. 250 µL of standard or test sample was added, mixed, and incubated at room temperature for 15 minutes. Absorbance was measured at 412 nm. Two test samples of GROS_g02583 were tested at two different concentrations.

To check that the assay had worked, GROS_g02583 was denatured with DTT and heat. This denatures cysteine-cysteine interactions and allows the Ellman's reagent to bind.

5.2.10 Cell death assay

Leaves of 4-week old *N. benthamiana* were infiltrated as described in section **5.2.6.2** using concentrations of OD_{600} 0.1, 0.2 or 0.3. Leaves were infiltrated with test genes in addition to GFP, RFP and unmodified *Agrobacterium* controls. After infiltration, plants were left on the lab bench for 2 weeks. Leaves were removed from the plant and imaged in both white light and under ultra violet lamps.

5.2.11 Upstream motif analysis

Sequences upstream of hypodermal genes were analysed for conserved motifs. Upstream regions were extracted using get_upstream_regions.py (Peter Thorpe, https://github.com/peterthorpe5 GNU GENERAL PUBLIC LICENSE). Motif enrichment was calculated using HOMER comparing the hypodermal gene list to the DOGbox gene list as a control as these genes are known not to be expressed in the hypodermis.

5.2.12 Phylogenetic analysis

A phylogenetic tree was created to analyse relationships between homologous proteins in various nematode species. Sequences were obtained using BLASTP or tBLASTN against genomes or transcriptome data from published sources downloaded either through https://parasite.wormbase.org or from the James Hutton Institute Intranet and aligned using the MUSCLE client with default settings.

Gaps in the alignment were trimmed using TrimAL (-gappyout) (Capella-Gutiérrez *et al.*, 2009).Trimmed, aligned sequences were subjected to model selection Whelan and Goldman (WAG) with GAMMA and invariable sites (+G,+I). The phylogenetic tree was created with Baysien inference (Mr. Bayes) with 500,000 generations and standard sample frequency and burn in rate of 10% and 25% respectively using TOPALi software (v2.5) (Milne *et al.*, 2009).

5.3 Results

The small size of PCN juveniles makes obtaining enough proteinaceous surface coat material for mass spectrometry a challenge. Normally this can be overcome by simply using more nematodes. However, as discussed in section **5.1.3**, PCN gland cells are very active at this life stage. This makes it easy to contaminate surface protein extractions with oesophageal gland proteins. Therefore, a new method was developed for increasing the likelihood of labelling the surface proteins of the nematode with biotin, therefore increasing the chances of extracting surface proteins over gland cell proteins. Surface proteins could then be pulled down and concentrated using streptavidin. After extraction, proteins were analysed by electrospray ionisation time of flight mass spectrometry, carried out by the mass spectrometry service at the University of St. Andrews.

Surface protein extractions were carried out 5 times using techniques described in section **5.2.2.** Controls containing no biotin were used to identify the possibility of any naturally biotinylated proteins that would be pulled down by streptavidin. Protein data was collated and screened for candidate proteins. It was assumed that all proteins secreted to the surface of the nematode would have a signal peptide. This was therefore used as a filter to reduce the number of protein candidates identified. Peptides identified by mass spectrometry have a confidence interval of 95% and total protein confidence interval of 99% as decided by the default settings on Scaffold 4 proteome software. Candidate proteins are shown in **Table 5.1**.

Table 5.1 – Proteins identified from juvenile surface protein extractions.
Candidate surface coat proteins appeared in two or more protein extractions. The peptide match confidence limit was set to 95% and protein match confidence limits were set to 99%. The average number of unique peptides from the two or more analyses has been rounded to the nearest whole number.

Gene ID	Protein function	Molecular weight	Average number of unique peptides identified	Average Percentage coverage
GROS_g13102	Glycoside hydrolase family 31	106 kDa	11	13.2

GROS_g09144	Peptidase S8 subtilisin related protein	108 kDa	6	8.7
GROS_g01970	Amidase	69 kDa	3	6.9
GROS_g02910	Tyrosinase-like	163 kDa	4	5.0
GROS_g12305	Heat shock protein	95 kDa	4	7.6
GROS_g02873	Late embryogenesis abundant protein	23 kDa	4	22.1
GROS_g02968	EGF- like calcium binding protein	211 kDa	5	3.5
GROS_g09940	Cyclophilin-type peptidyl-prolyl cis trans isomerase	22 kDa	3	15.5
GROS_g09658	Transthyretin	16 kDa	2	9.9
GROS_g06693	Metalloproteinase Inhibitor	20 kDa	2	7.7
GROS_g01765	Hypothetical oesophageal gland protein	28 kDa	13	42.5
GROS_g08190	Chorismate mutase	32 kDa	6	28.0
GROS_g04677	Endoglucanase	50 kDa	3	10.7
GROS_g07949	Glycoside hydrolase family 5	37 kDa	4	23.6
GROS_g02583	Hypothetical protein	23 kDa	3	15.7
GROS_g04366	Pectate lyase 2	27 kDa	3	12.5
GROS_g14202	Fatty acid and retinol binding protein	22 kDa	4	18.9
GROS_g02490	Glutathione peroxidase	28 kDa	3	16.3
GROS_g07968	Pectate lyase 1	28 kDa	2	17.4
GROS_g10505	Glycoside hydrolase family 5	62 kDa	4	11.3
GROS_g07793	Integrin beta subunit	92 kDa	8	12.3

GROS_g01882	Histidine kinase-like	34 kDa	7	25.1
GROS_g05049	Fibronectin	443 kDa	32	11.7
GROS_g01153	Galectin-5	37 kDa	4	12.7
GROS_g06000	ABC transporter substrate-binding protein	47 kDa	3	7.7
GROS_g00761	Integrin alpha	135 kDa	6	8.3
GROS_g01391	Heat shock protein	82 kDa	10	16.7
GROS_g10494	Fatty acid oxidation complex	89 kDa	8	13.2

Identification of proteins such as the fatty acid and retinol binding protein (FAR-1, GROS_g14202) and glutathione peroxidase (GROS_g02490) act as positive controls for the surface protein extraction technique, as both proteins have previously been identified and localised to PCN surfaces (Prior *et al.*, 2001; Jones *et al.*, 2004). Proteins associated with nematode esophageal glands were also identified such as glycoside hydrolases, pectate lyases and chorismate mutase. These proteins were not studied further.

FAR-1 is reported to be present on the surface of many other nematode species (Le *et al.*, 2016; Phani *et al.*, 2017). However, it has never been localised to the surface of *G. rostochiensis*. Due to the sequence similarity between Gr-FAR-1 and Gp-FAR-1, antibodies produced for *G. pallida* could be used for both species (**Figure 5.1**). This showed that Gr-FAR-1 is localised on the surface of *G. rostochiensis*. A pre-immune serum negative control was used to confirm that the secondary antibody was not natively binding to the surface of *G. rostochiensis*. These experiments confirm that the techniques used here are capable of identifying proteins present on the surface of *G. rostochiensis*. *In situ* hybridisation was used to localise gene expression. Gr-FAR-1 mRNA localised to the hypodermis in both juvenile and parasitic life-stages (**Figure 5.1**). In parasitic stages, hypodermis staining develops from being segmented along the length of the nematode to becoming one continuous structure resembling three stripes from the head to tail of the nematode. This staining pattern is seen as the syncytial

hypodermis begins to swell throughout allowing staining throughout the cell, not only around the nuclei as seen in juvenile hypodermal staining (**Figure 5.2**).

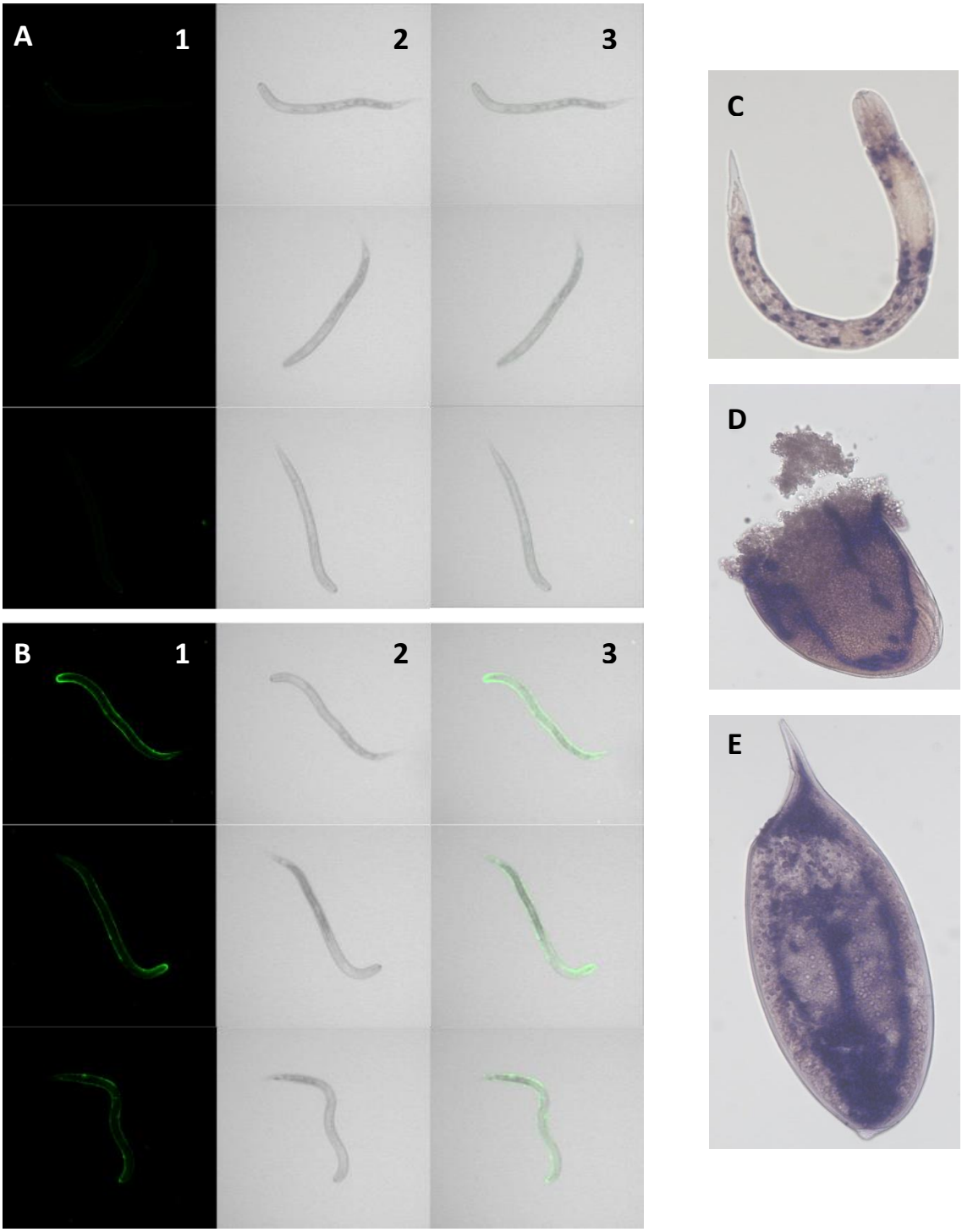


Figure 5.1 – Immunofluorescence and in situ hybridisation of Gr-FAR-1.
Antibodies raised against Gp-FAR-1 were used to localise the fatty acid and retinol binding protein. FAR-1 is present on the surface of *G. rostochiensis* juveniles. A – pre-immune serum. B – anti-serum. 1 – Presence of anti-Gp-FAR-1 detected by anti-rabbit-cy3 (green); fluorescence signal can be seen around the surface of the juvenile. 2 – DIC image of juvenile. 3 – merge of left and centre. C, D, E – hypodermis localisation of Gr-FAR-1 mRNA using in situ hybridisation in juvenile (C) or parasitic stage (D, E) *G. rostochiensis* Images have been brightened 20% for printing

Creating antibodies for all proteins identified in **table 5.1** was not possible as it would be too costly. Therefore, *in situ* hybridisation was used to localise gene expression for these proteins. Proteins released to the surface of the nematode are normally created in and secreted from the hypodermis. Glutathione peroxidase was used as a positive control for hypodermal *in situ* hybridisation staining. Punctate patches of purple staining along the sides of the nematode are indicative of hypodermal staining. At the J2 life stage the syncytial hypodermis is very thin but swells around the nuclei, it is around these areas that the anti-digoxigenin and NBT stain are most visible (Endo & Wyss, 1992) (**Figure 5.2**).

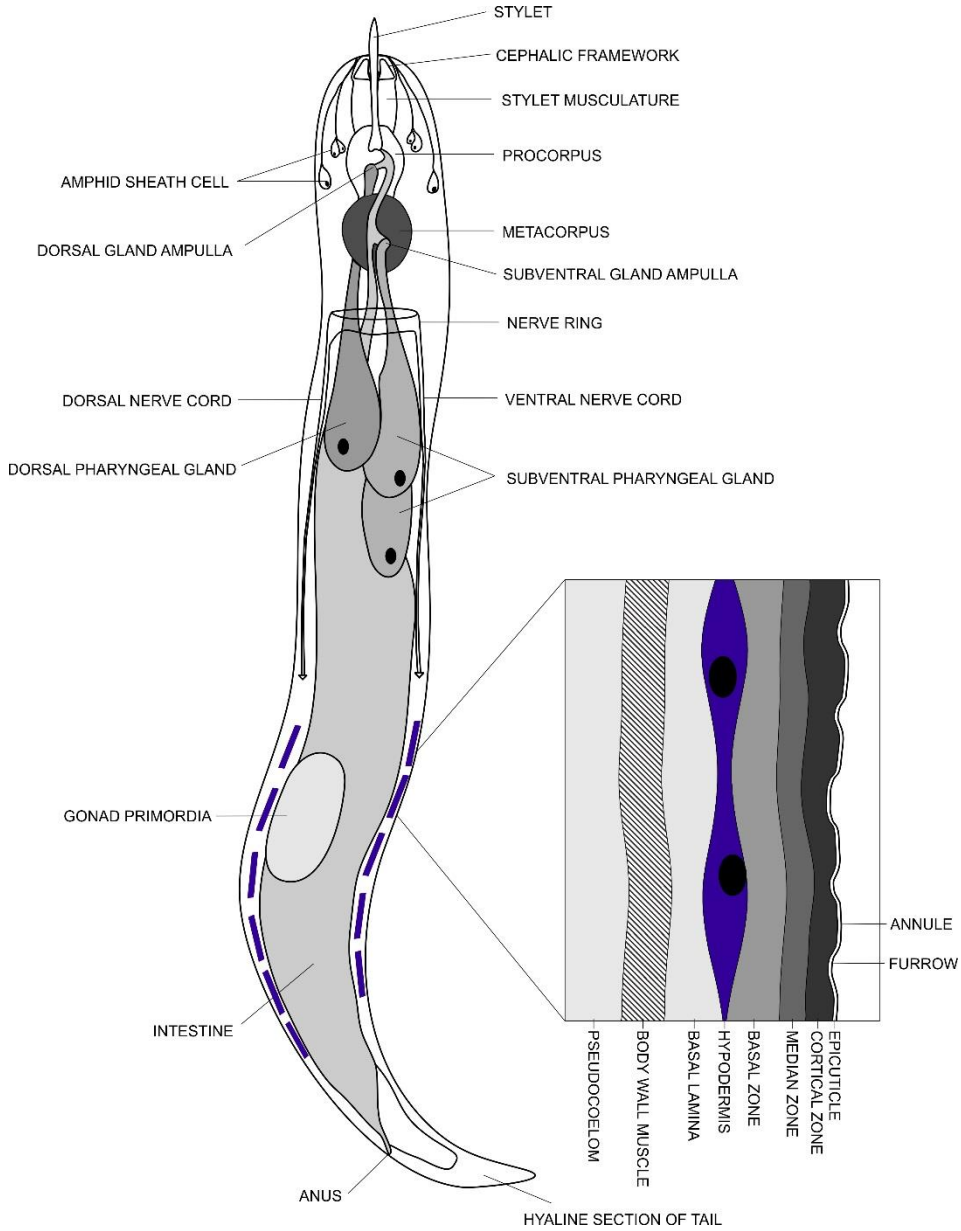


Figure 5.2 – Schematic of PCN juvenile with hypodermis staining following in situ hybridisation

Staining is found throughout the syncytial hypodermis but is more apparent around areas where nuclei are present due to swelling of the cell accommodating for these structures.

In situ hybridisation using a positive control, GROS_g02490, glutathione peroxidase showed staining in the hypodermis as expected (Figure 5.3). No staining was visible in negative controls (not shown). In addition, analysis of RNAseq data available for *G. rostochiensis* showed upregulation of GROS_g02490 at the juvenile lifestage as expected (Figure 5.3).

In situ hybridisation was carried out on several candidates from **Table 5.1.** Candidates were selected based on reliability of mass spectrometry data and comparisons to similar proteins identified in literature searches. Results from *in situ* hybridisations are summarised in **Table 5.2**, and shown in **Figure 5.4**.

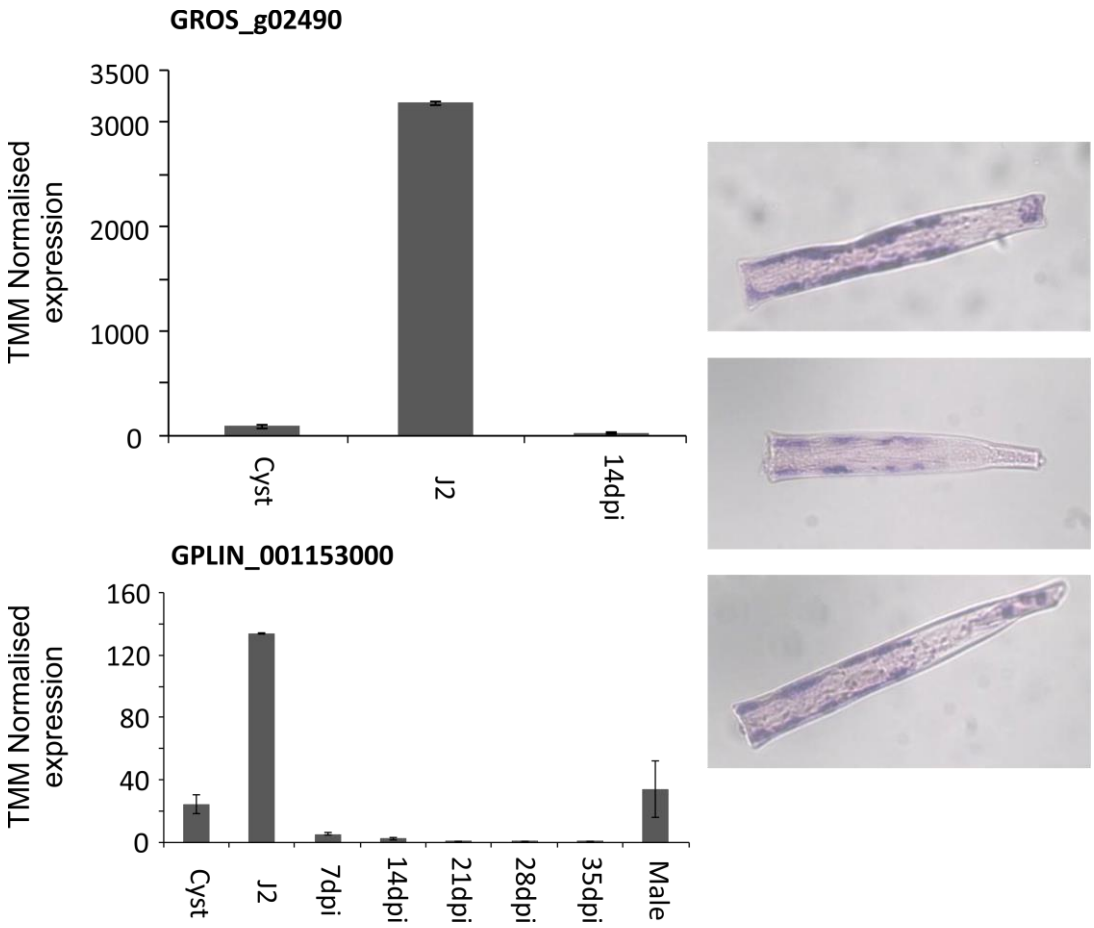


Figure 5.3 – Expression data and in situ hybridisation staining of GROS_g02490.
Heightened expression for this gene can be seen in juvenile stages of both *G. rostochiensis* (GROS_g02490) and the *G. pallida* orthologue (GPLIN_001153000), reflecting the need for the protein during early infection of the host. Staining of the probe (purple) is localised to the hypodermis as shown by the segmented patches along the sides of the nematode.

Table 5.2 – Location of expression for target genes identified by mass spectrometry.
In situ hybridisation was used to localise expression of mRNA. Proteins being expressed onto the nematodes' surface are expected to be produced in the hypodermis. **Hyp** = hypodermis. **GP** = genital primordia. **NR** = nerve ring. **NC** = nerve chord. **SvG** = subventral gland **ND** = not detected.

Gene Name	Predicted/Known Function	*In situ* localisation
GROS_g02490	Glutathione peroxidase	Hyp
GROS_g01153	Galectin-5	Hyp
GROS_g06693	Metalloproteinase Inhibitor	Hyp
GROS_g10494	Fatty acid oxidation complex	GP/Hyp
GROS_g01391	Heat shock protein	GP/Hyp
GROS_g01882	Histidine kinase-like	SvG
GROS_g06000	ABC transporter substrate-binding protein	ND
GROS_g05049	Fibronectin	Hyp
GROS_g01970	Amidase	NR
GROS_g02910	Tyrosinase-like	ND
GROS_g09658	Transthyretin	NC
GROS_g02583	Hypothetical protein	Hyp

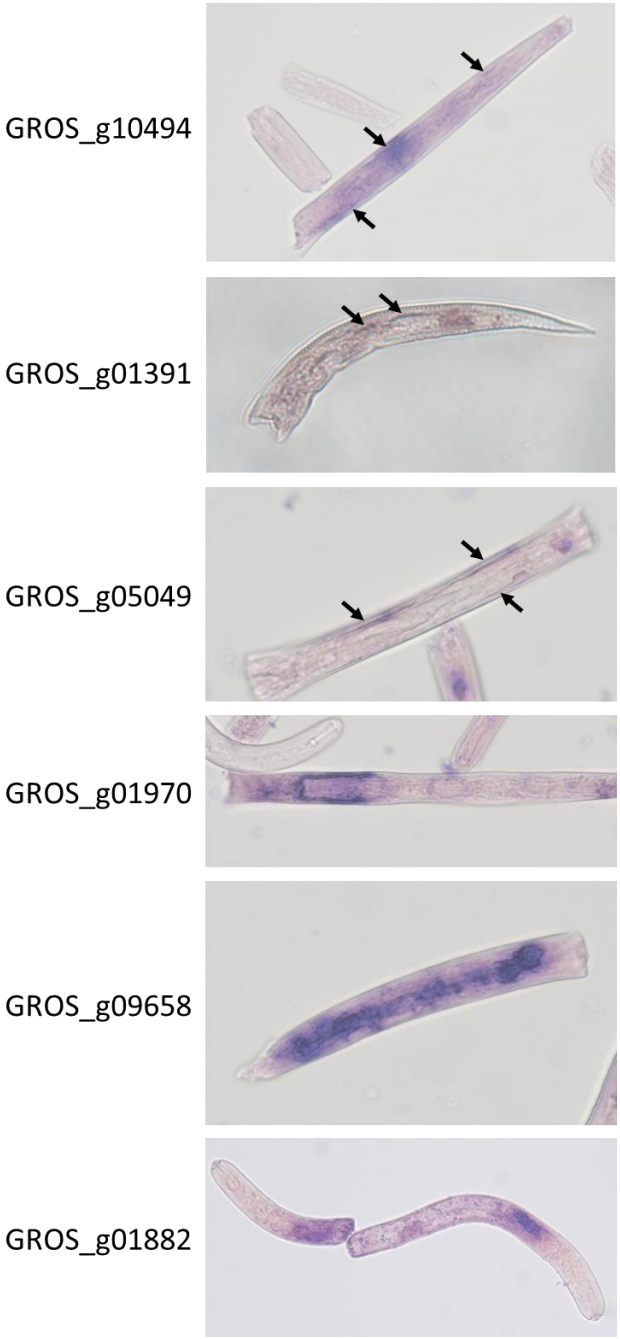


Figure 5.4 – Localisation of gene expression using in situ hybridisation.

In situ hybridisation was used to localise gene expression of various surface protein candidates identified in surface extraction mass spectrometry data. Anti-sense probe staining for GROS_g10494, GROS_g01391 and GROS_g05049 can be seen in the hypodermis as indicated by black arrows. GROS_g01970 was localised to the nerve ring. Speckled staining along the central length of the nematode suggests that GROS_g09658 is expressed in the nerve chord. Elongated ovular staining towards the anterior suggests GROS_g01882 is expressed in the sub ventral gland.

Following *in situ* hybridisation results, three candidates (GROS_g06693, GROS_g01153 & GROS_g02583) were taken for further research based on their potential roles on the surface of the nematode. These proteins may disrupt host defences through protein-protein interactions, protein-carbohydrate interactions or direct protein-ROS interactions.

5.3.1 GROS_g06693, metalloproteinase inhibitor

GROS_g06693 was identified several times in mass spectrometry results, has expression data showing increased expression in the J2 life stage and is clearly expressed in the hypodermis (**Figure 5.5**). This protein was previously identified by Wang *et al*. (2001), and identified on the basis of similarity as a conserved regulator of innate immunity, CRI-2. In this work the protein was said to be localised in the oesophageal gland. However, in this study, the mRNA was never detected by *in situ* hybridisation within the gland cells. This protein has a signal peptide and no transmembrane domains.

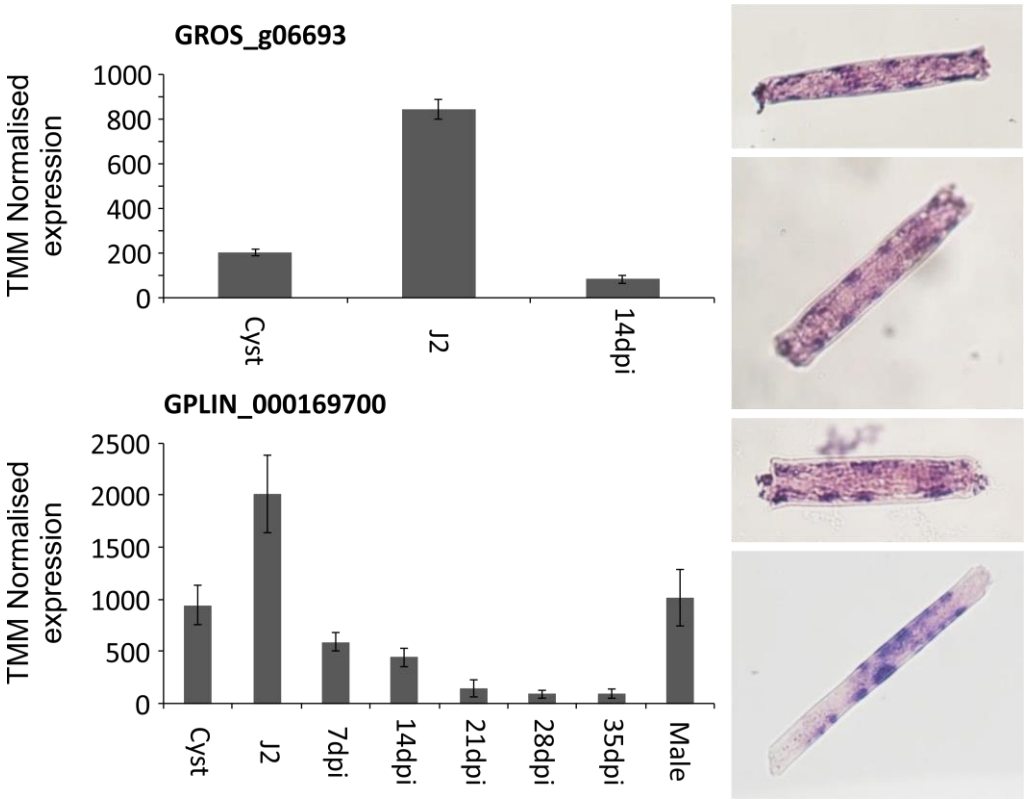


Figure 5.5 – Expression data and in situ hybridisation for GROS_g06693

Expression of this gene is heightened in juvenile stages of both *G. rostochiensis* (GROS_g06693) and the *G. pallida* orthologue (GPLIN_000169700), reflecting the possible need for the protein during early infection of the host. *In situ* hybridisation localises this mRNA to the hypodermis, no staining was seen in negative controls.

Due to being previously identified as a nematode excretory/secretory protein, expression data matching that of a protein that could aid infection of a host and *in situ* hybridisation results confirming gene expression within the hypodermis, GROS_g06693 was further tested for its ability to interfere within a plant defence pathway. A reactive oxygen species induction assay was used on leaf discs of *N. benthamiana* expressing GROS_g06693 inside cells or GROS_g06693 that was trafficked to the apoplast. ROS defences in the plant were induced with a bacterial flagellin protein (flg-22). No flg-22 controls, 'water', were used to ensure there was no other source of ROS initiation. ROS produced in response to flg-22 ultimately leads to production of luminescence when in the presence of horseradish peroxidase and L-012 (Mei *et al.*, 2018). If recombinant GROS_g06693 interacts with the plant defences, there will be a detectable change in availability of ROS for light production. Results for the ability of GROS_g06693 to alter ROS availability were inconclusive therefore; all data collected from all

spectrophotometer reads were averaged and graphed (Figure 5.6). There were no significant differences between ROS production in the samples. The delayed reaction to flg-22 in the leaf discs expressing GROS_g06693 in the ApSiPR vector is due to these samples being the last to be introduced to the PAMP on the 96-well plate.

Changes in ROS were also measured after inducing ROS with cell wall breakdown products (**5.2.8.1**). However, no significant differences in ROS production under these conditions were noted (not shown).

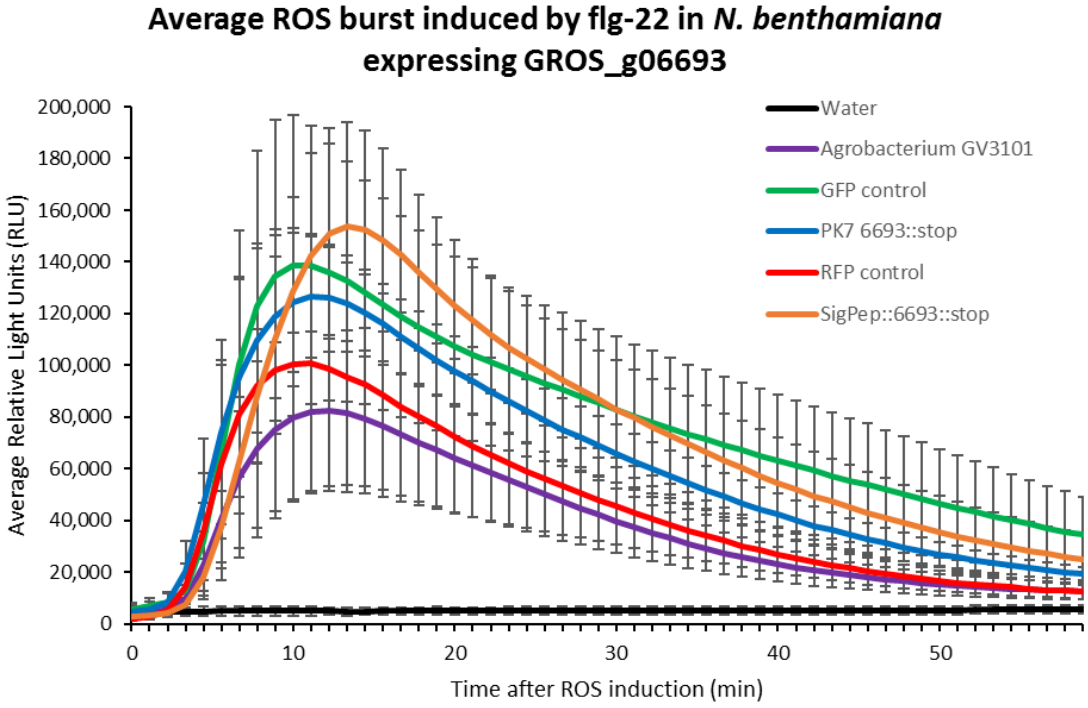


Figure 5.6 – ROS burst induced by flg-22 in *N. benthamiana* leaf discs expressing GROS_g06693

Due to inconclusive results between plates of ROS assays in leaf discs expressing GROS_g06693, results from all assays were averaged. This showed no significant difference in ROS availability with or without expression of GROS_g06693. Error bars represent standard error.

GROS_g06693 is similar to inhibitors of matrix metallopeptidases (MMPs). Attempts were made to identify the MMPs that could be inhibited within the plant. GROS_g06693 was expressed in *N. benthamiana* through *Agrobacterium* infiltration and sent to the apoplast using the ApSiPR vector. Plants were left to grow for 3 days before apoplastic fluid was extracted and GROS_g06693 was pulled down using the bound RFP tag. Any MMPs bound by GROS_g06693 were expected to also be purified using this process.

However, mass spectrometry on protein samples was unable to identify any MMPs. Recombinant GROS_g06693 was detected (not shown).

5.3.2 GROS_g01153, galectin-5

GROS_g01153 has a signal peptide, no transmembrane domains, was identified in several surface protein extractions and is localised to the hypodermis. GROS_g01153 has domains suggesting a galectin function. Galectins are proteins that specifically bind to β-galactoside sugars. This galectin is upregulated in all life-stages actively migrating through the host plant, there is no increase in expression during life-stages associated with nematode feeding (**Figure 5.7**). This could suggest that the protein differs from other galectins and is not required for digestion.

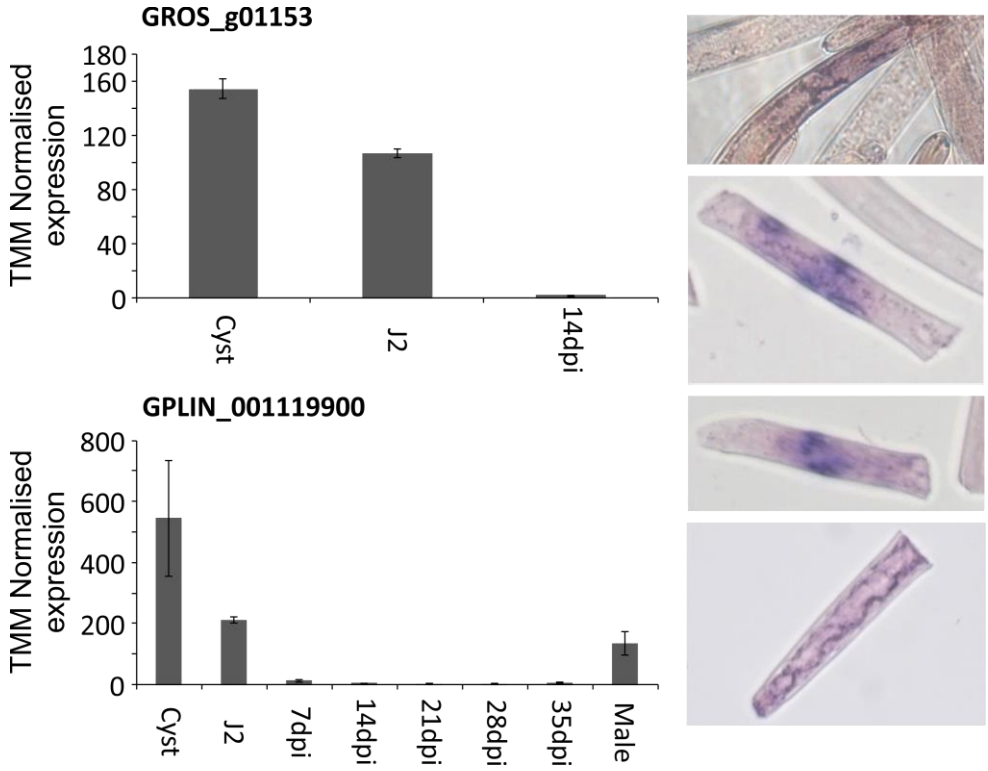


Figure 5.7 – Expression data and in situ hybridisation for GROS_g01153
Although there is no clear upregulation of this gene at the juvenile lifestage there is also little expression throughout the infective stages of both *G. rostochiensis* (GROS_g01153) and the *G. pallida* orthologue (GPLIN_001119900). This suggests the protein is not being used for digestion within the intestine. *In situ* hybridisation localises the mRNA to the hypodermis, no staining was visible in negative controls.

Plants respond to damage through the detection of damage associated molecular patterns (DAMPs). One example of a DAMP could be cell wall breakdown products produced as a result of the action of cell wall degrading enzymes produced by the pathogen. Upon detecting damage, the plant produces a defence-like response (Souza *et al.*, 2017; Shah *et al.*, 2017). GROS_g01153 could bind to cell wall breakdown products, potentially soaking up host DAMPs created during the nematodes destructive passage through the root. Binding these DAMPs could offer the nematode protection from the plant defences that would respond to detection of oligogalacturonides. Cell wall breakdown products were mimicked by 10-15mer oligogalacturonides (GAT114, Elicityl) as previously used by Shah et al., (2017). No oligogalacturonide controls, 'H_2O', were used to ensure there was no other source of ROS initiation Recombinant GROS_g01153 was expressed in *N. benthamiana* and kept within the cell or sent to the apoplast using the ApSiPR vector. When expressed within the cell, the recombinant galectin significantly reduced the ROS burst created in response to detection of DAMPs (**Figure 5.8**). Although there was some reduction in detectable ROS when the protein was present in the apoplast, this was less than what was seen when the protein was present in the cytoplasm and was not significantly different from the control.

Attempts to recombinantly express this protein were carried out. However, protein translation appeared to terminate shortly into the sequence. The plasmid was checked to confirm no mutations resulting in a premature stop codon had occurred. Stunted translation was possibly an effect of the high rare codon usage in this protein, using tRNA optimised expression cell lines (BL21) did not support expression. Due to time constraints and the ability to produce the protein in planta for ROS assays, attempts to recombinantly express this protein were stopped.

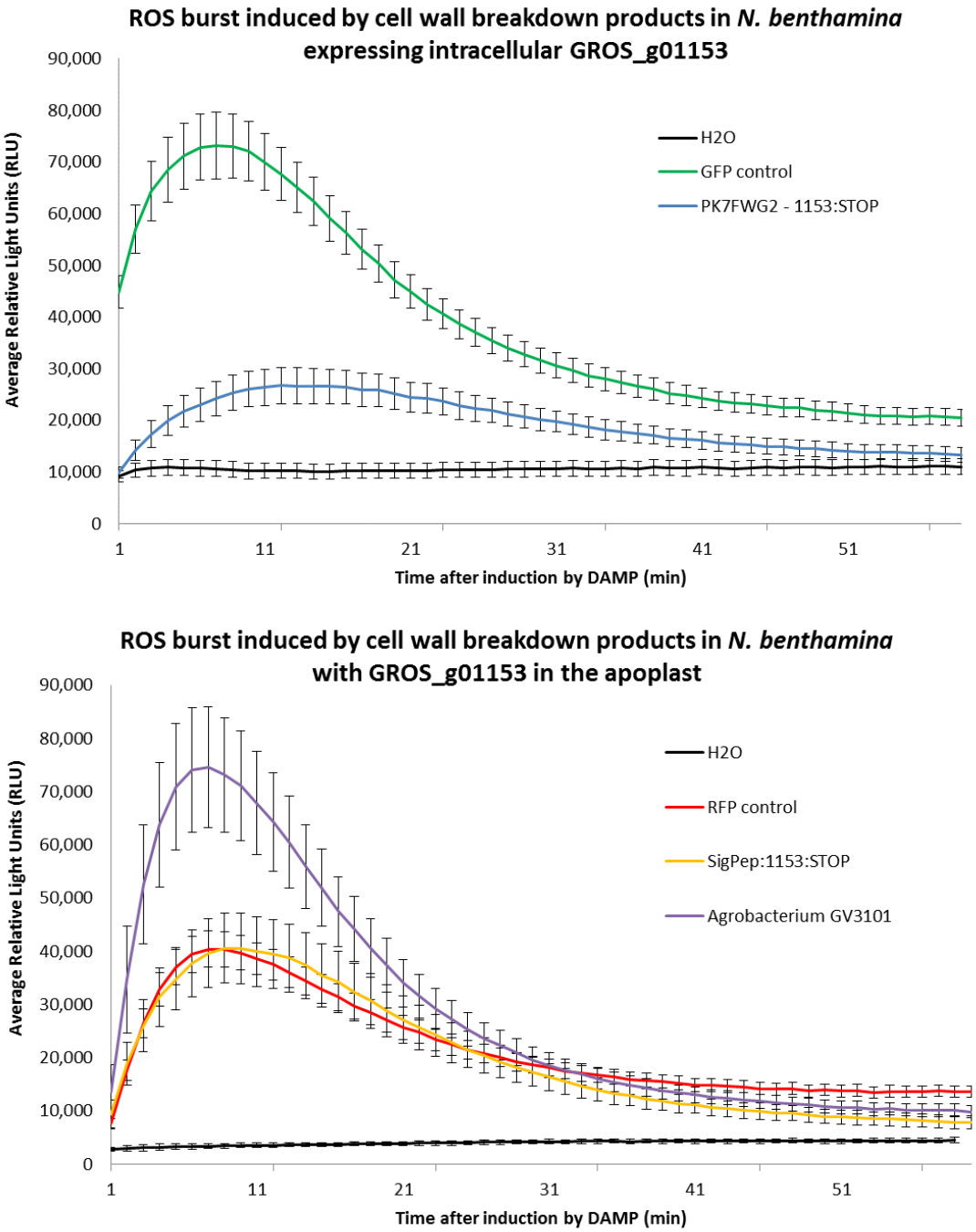


Figure 5.8 – ROS burst induced by cell wall breakdown products (oligogalacturonides) in *N. benthamiana* leaves expressing recombinant GROS_g01153.

Oligogalacturonides (GAT114, Elicityl) were used to induce a ROS burst in leaves of *N. benthamiana*. When GROS_g01153 was recombinantly expressed in the leaves using the pK7FWG2 vector there was a decrease in the plant response, reflecting the ability of this protein to detoxify these DAMPs (**top**). The same effect was not seen when the recombinant protein was sent to the apoplast (**bottom**). It is possible no effect is seen here as the protein has bound to the cell walls and so is no longer present to bind to the polysaccharides. Error bars represent standard error.

5.3.3 GROS_g02583, hypothetical protein

GROS_g02583 has no known domains when searched against the Pfam database. This protein was detected multiple times in mass spectrometry data and the gene was successfully cloned from *G. rostochiensis* cDNA. GROS_g02583 has a predicted signal peptide and no transmembrane domains were detected. Expression data shows upregulation of the gene after invasion of the roots and little expression at J2 (**Figure 5.9**). This perhaps reflects a role in responding to stress from the plant environment and not one experienced from soil. Expression of this gene was localised to the hypodermis (**Figure 5.9**). GROS_g02583 did not match any hits against *G. pallida* genome assembly, therefore the gene had to be mapped against RNAseq transcriptomic data collected for *G. pallida* (Cotton *et al.*, 2014). This was achieved by trimming raw unmapped transcriptome data (ERP001236) using the trimmomatic program (v0.32) (Bolger *et al.*, 2014). Next, the gene of interest, GROS_g02583, was matched to the trimmed reads using Bowtie (Bowtie2) (Langmead & Salzberg, 2012). After sorting, indexed files were entered into Bedtools (2.27.0) (Quinlan & Hall, 2010) which outputs gene lengths and counts. Expression data could then be obtained by entering this data into edgeR (Robinson *et al.*, 2010; McCarthy *et al.*, 2012). GROS_g02583 was found to match *G. pallida* comp36_c0_seq1.

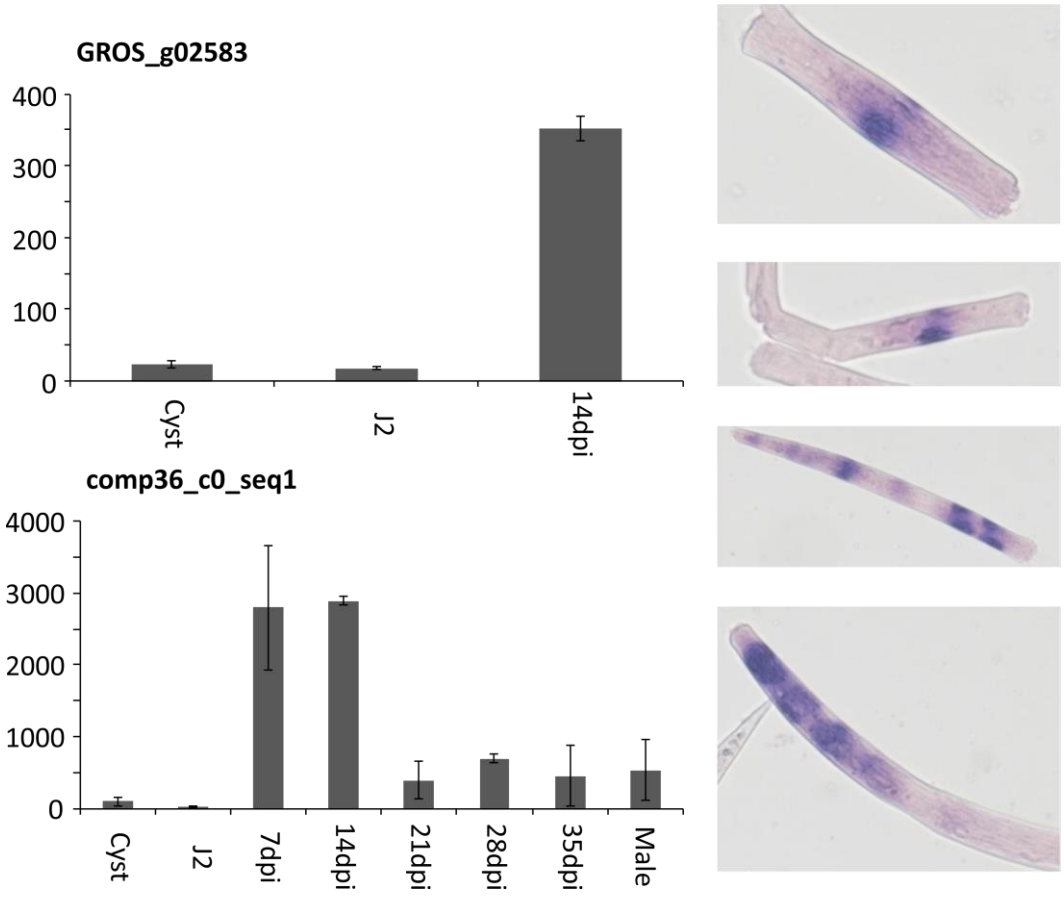


Figure 5.9 – Expression data and in situ hybridisation for GROS_g02583.
Expression data in both species of PCN show large upregulation of GROS_g02583 after infection of the host roots in both *G. rostochiensis* (GROS_g02583) and the *G. pallida* orthologue (comp36_c0_seq1). This suggests that the protein is not required upon immediate invasion or whilst the nematode is mobilising to the host. *G. pallida* expression values were calculated after mapping GROS_g02583 against *G. pallida* transcriptome data. In situ hybridisation localises the mRNA to the hypodermis.

Subcellular localisation was carried out to identify any apparent localisation of GROS_g02583 within plant cells. Expressing GROS_g02583 within cells of *N. benthamiana* using the gateway vector pk7FWG2 showed that the protein was aggregating within the cytoplasm and not localising to any obvious structure. However, when the protein was expressed using the ApSiPR vector (**5.2.6.1**) which fuses the protein to an apoplastic specific signal peptide, recombinant protein aggregation was no longer visible (Figure 5.10). Primarily, this highlights that the expression vectors used here direct recombinant proteins within the cell or to the apoplast.

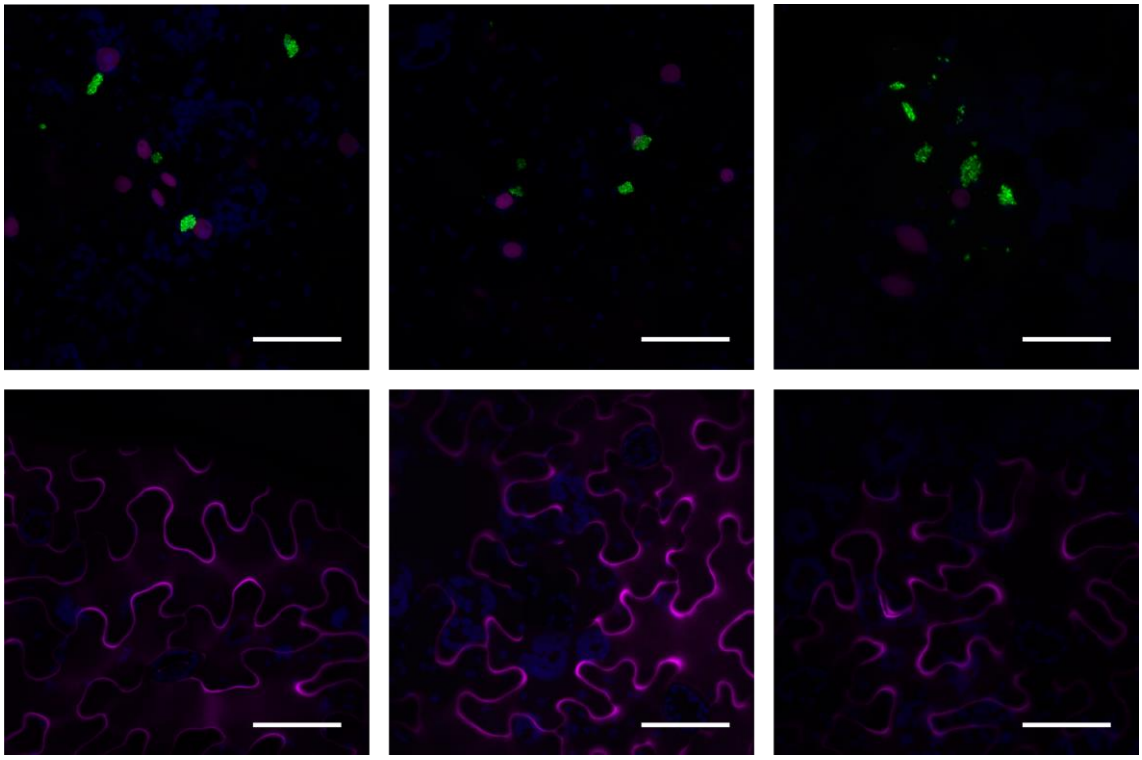


Figure 5.10 – Subcellular localisation of GROS_g02583.
Top – GROS_g02583 expressed by pk7FWG2 vector, C-terminal GFP tag (green). Protein is aggregated suggesting possible misfolding or denaturing when within the cell. Bottom – GROS_g02583 expressed by ApSiPR vector, C-terminal RFP tag (magenta). No aggregation of recombinant protein is seen when the protein is forwarded to the apoplast, possibly resulting from a more favourable pH. Scale bars represent 50 µm. Images have been brightened by 20% for printing.

Further information about this protein was gathered by aligning protein homologues across various nematode species. Across 21 tested species, all protein sequences showed 16 highly conserved cysteines (Figure 5.11). Sequences from other species were obtained using BLASTP or tBLASTN against genomes or transcriptome data from published sources downloaded either through https://parasite.wormbase.org or from the James Hutton Institute Intranet (C. elegans Sequencing Consortium, 1998; Ghedin *et al.*, 2007; Kikuchi *et al.*, 2011; Godel *et al.*, 2012; Laing *et al.*, 2013; Cotton *et al.*, 2014; Tang *et al.*, 2014; Bauters *et al.*, 2014; Schwarz *et al.*, 2015; Zhu *et al.*, 2015; Vieira *et al.*, 2015; Kulkarni *et al.*, 2016; McNulty *et al.*, 2016; Zheng *et al.*, 2016; Eves-Van Den Akker *et al.*, 2016; Eves-van den Akker *et al.*, 2016; Cotton *et al.*, 2017; Phillips *et al.*, 2017; Blanc-Mathieu *et al.*, 2017; Eves-Van Den Akker, 2018, personal communication; Pokhare, 2018, personal communication)

The conserved cysteines follow a clear pattern of $CXCX_5CX_{10}CX_3CX_{16}CX_{13}CX_8CX_{49-51}CX_3CX_{7-8}CX_{3-5}CX_{16}CX_{14}CXCX_{12-13}$ across all species of tested nematodes. Spaces between cysteines are equal across all species towards the beginning of the sequence. Additionally, the first half of the protein shows higher conservation than the C-terminal end of the protein suggesting that this may be an active region for this protein, or a region of structural significance. Interestingly, all sequences have a predicted signal peptide except for the *P. penetrans* protein. However, this is likely to be due to misprediction of the protein as many other sequences also have a methionine at this point.

Sequences aligned here were also used to create a tree based on sequence similarities. However, this did not suggest anything unusual and the tree resembled a normal nematode phylogenetic tree (**Figure 5.12**).

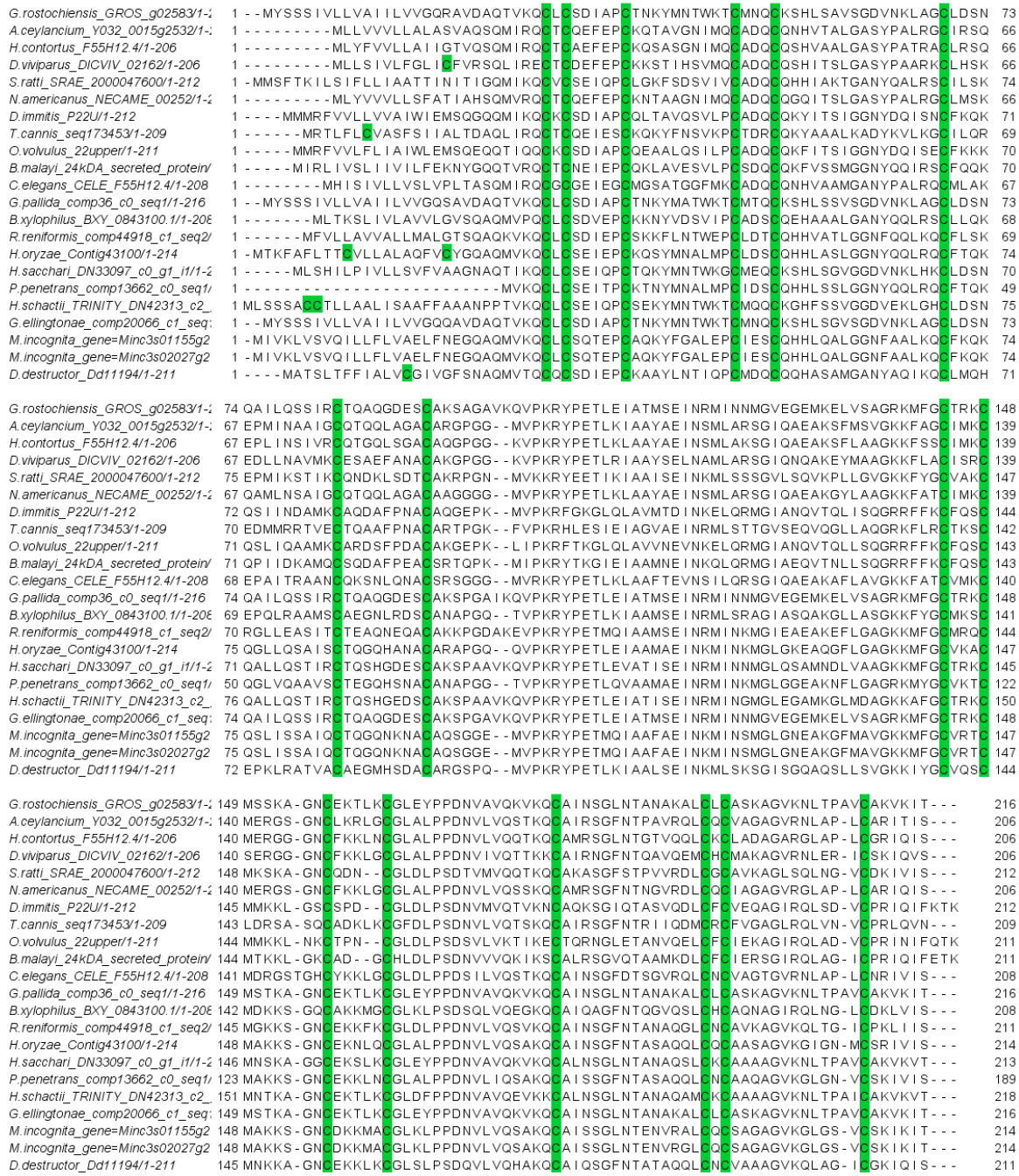


Figure 5.11 – Sequence alignment for homologues of GROS_g02583 across PPN, APN and FLN.

Sequences were aligned using Jalview v-2.10.3 following MUSCLE with default parameters. 16 conserved cysteines have been highlighted in green. Strong conservation of this gene can be seen across all tested species of nematodes.

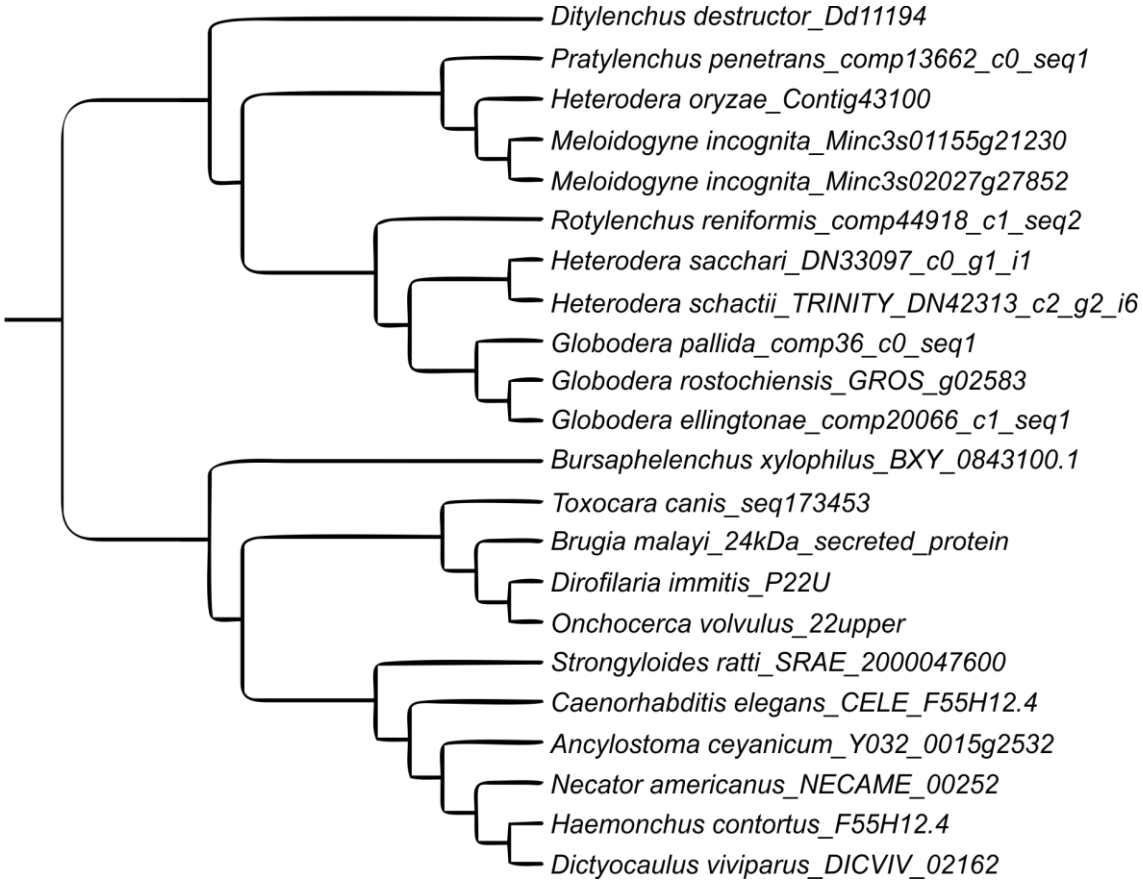


Figure 5.12 – GROS_g02583 alignment tree

A phylogenetic tree was created to see evolutionary relationships based on sequence alignments between GROS_g02583 homologues in various nematode species. The tree resembles nematode clade taxonomy and does not highlight anything unusual about the evolution of this gene.

Thiol groups (-SH) can interact with reactive oxygen species, detoxifying ROS by becoming oxidised. Such effects can be seen in peroxiredoxins and thioredoxins (Harris & Hansen, 2012). Small, cysteine rich proteins such as GROS_g02583 therefore have the potential to act as sponges for ROS, detoxifying an otherwise harmful environment. Ellman's reagent (5,5'-dithiobis-2-nitrobenzoic acid, DNTB) binds to free thiol groups, detectable by absorbance of light at 412 nm. Recombinant GROS_g02583 was obtained from RFP pull downs from apoplastic fluid following expression with the ApSiPR vector. Conventional recombinant expression/purification in *E. coli* could not be used here due to the high rare codon usage in this protein. An Ellman's reagent assay using recombinant GROS_g02583-RFP showed that no free cysteines were present in the protein (**Figure 5.13**). Ellman's reagent binds to free thiol groups measurable by

absorbance of visible light at 412 nm. GROS_g02583_1 and GROS_g02583_2 (10 x concentrated) showed negligible absorbance and so show no free thiols. Denaturing the protein with DTT and heat allowed binding of the reagent and a measurable increase in absorbance at 412 nm.

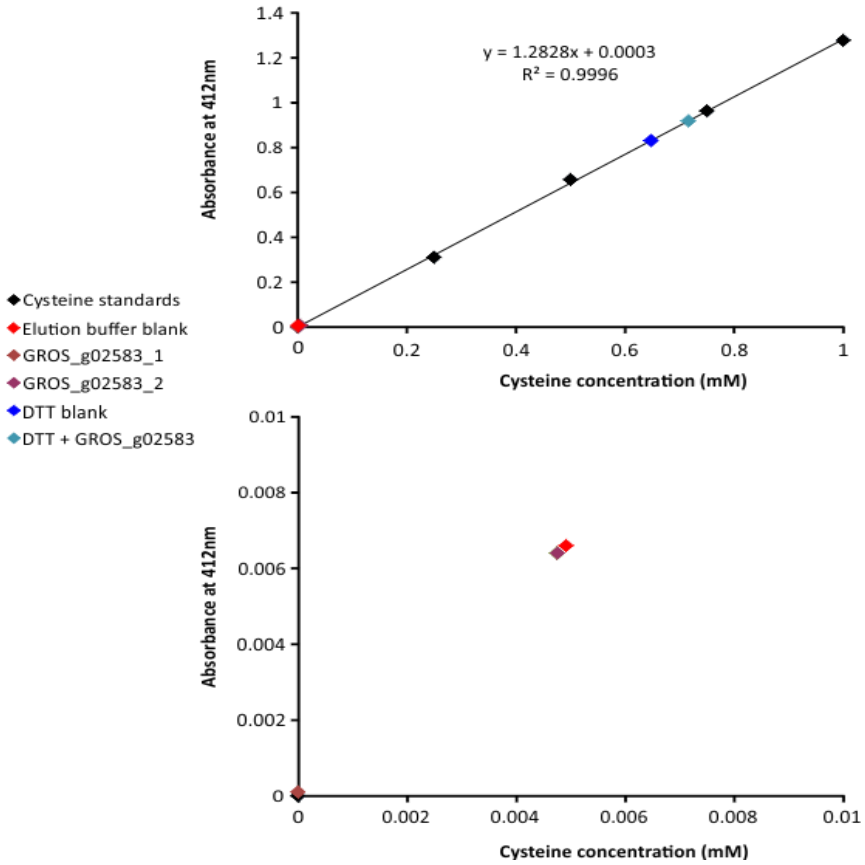


Figure 5.13 – Ellman's assay for quantifying free cysteines in GROS_g02583.
Top) Cysteine standards allowing cysteine concentration to be calculated in recombinant GROS_g02583 at two concentrations. When denatured with DTT, cysteines are bound by Ellman's reagent showing increase in cysteine concentration compared to the blank. Bottom) zoomed in lower region of the top graph. Fluorescence at 412 nm in recombinant GROS_g02583 is negligible or less caused by the elution buffer suggesting that there are no free cysteines in the native protein.

A ROS interaction assay was carried out to test the ability of GROS_g02583 to detoxify ROS produced in a defence response to the bacterial flagellin protein flg-22. No flg-22 controls, 'H_2O', were used to ensure there was no other source of ROS initiation. All repeats of the experiment showed a clearly heightened luminescence when GROS_g02583 was sent to the apoplast. Heightened luminescence in response to RFP

was also seen, but was minimal compared to response to recombinant GROS_g02583. There was little response to the recombinant protein when expressed within the plant cell (Figure 5.14).

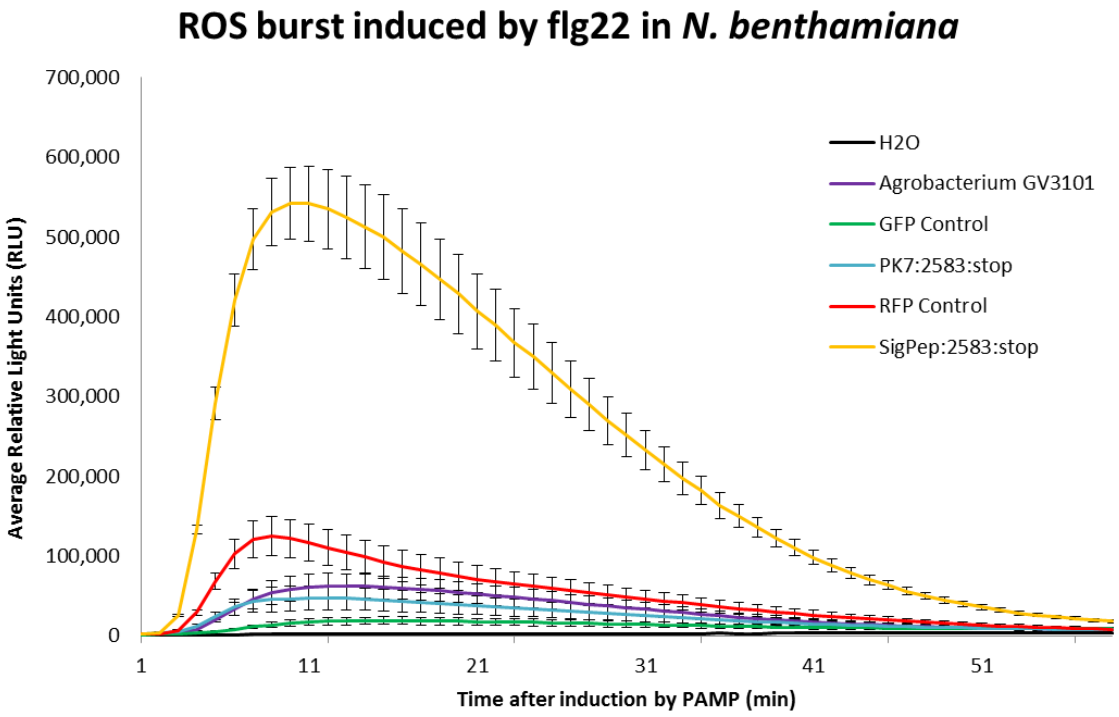


Figure 5.14 – ROS burst induced by flg-22 in *N. benthamiana* infiltrated by Agrobacterium expressing recombinant GROS_g02583 constructs.
Drastically increased luminescence can be seen in response to flg-22 when recombinant GROS_g02583 is present in the apoplast. The plant response to flg-22 was otherwise typical, peaking within 10-15 minutes and reducing to a plateau afterwards. Error bars represent standard error.

The increased response to flg-22 when in the presence of GROS_g02583 was initially thought to be indicative of a priming response. Therefore, the ability of GROS_g02583 to induce a ROS burst alone was tested. Due to difficulties in recombinantly expressing the protein using bacteria the protein was obtained from an apoplastic fluid pulldown. The ROS response in *N. benthamiana* was also tested in response to apoplastic fluid containing GROS_g02583 combined with flg-22. No ROS burst was seen in response to GROS_g02583. Additionally, the heightened ROS burst in response to flg-22 usually seen when GROS_g02583 is expressed in leaf material was not seen. The response to flg-22 is delayed and luminescence signal is not as typically clear, possibly due to amount of

other material present inside the apoplastic fluid (Figure 5.15). Ideally this needs to be tested with purified recombinant GROS_g02583.

To further narrow down how GROS_g02583 is causing an increased plant response to

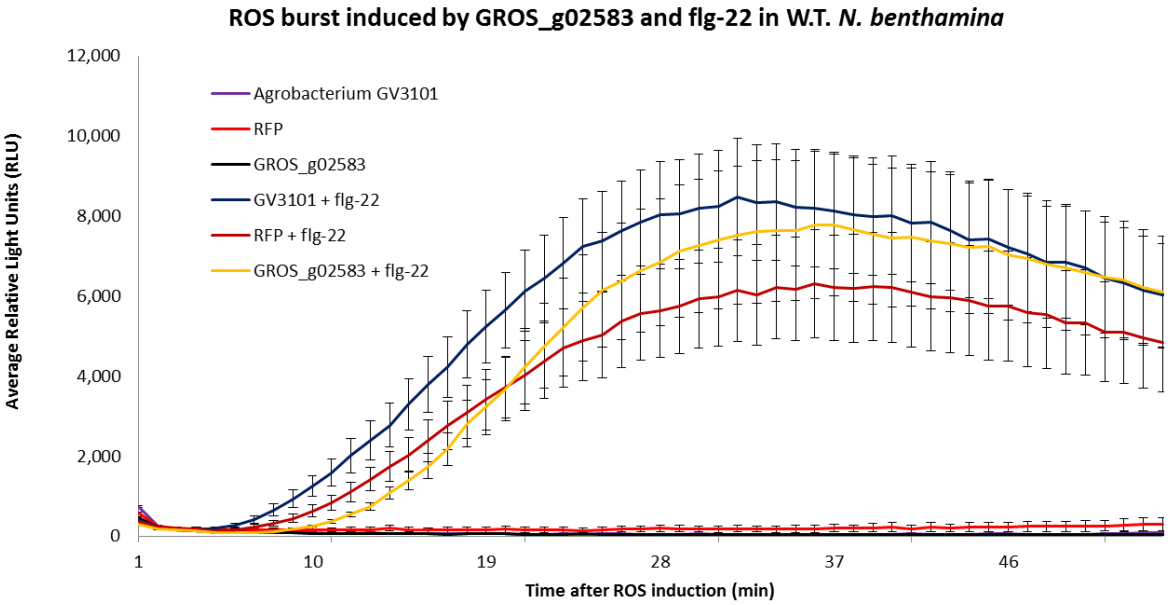


Figure 5.15 – ROS burst induced by GROS_g02583 and flg-22 in W.T. *N. benthamiana*. Apoplastic fluid from *N. benthamiana* infiltrated with *Agrobacterium* expressing Signal-peptide::GROS_g02583::stop was used to induce a ROS burst in unmodified, wildtype *N. benthamiana* leaf discs. Recombinant GROS_g02583 in apoplastic fluid does not induce a ROS burst alone. Additionally, when combined with flg-22 the heightened luminescence seen when GROS_g02583 was transiently expressed is not visible. Error bars represent standard error.

flg-22, another assay was carried out to check if the increased response was specific to PAMPs (flg-22) or if it can be seen with oligogalacturonide DAMPs as used in section **5.3.2**. This assay tested ROS burst responses in plants agroinfiltrated with either unmodified *Agrobacterium*, *Agrobacterium* expressing an RFP control or *Agrobacterium* expressing recombinant GROS_g02583 targeted to the apoplast. Leaf discs were exposed to oligogalacturonides and luminescence was recorded (Figure 5.16). The previously seen, large response when flg-22 was applied to leaf discs expressing GROS_g02583 was not seen here.

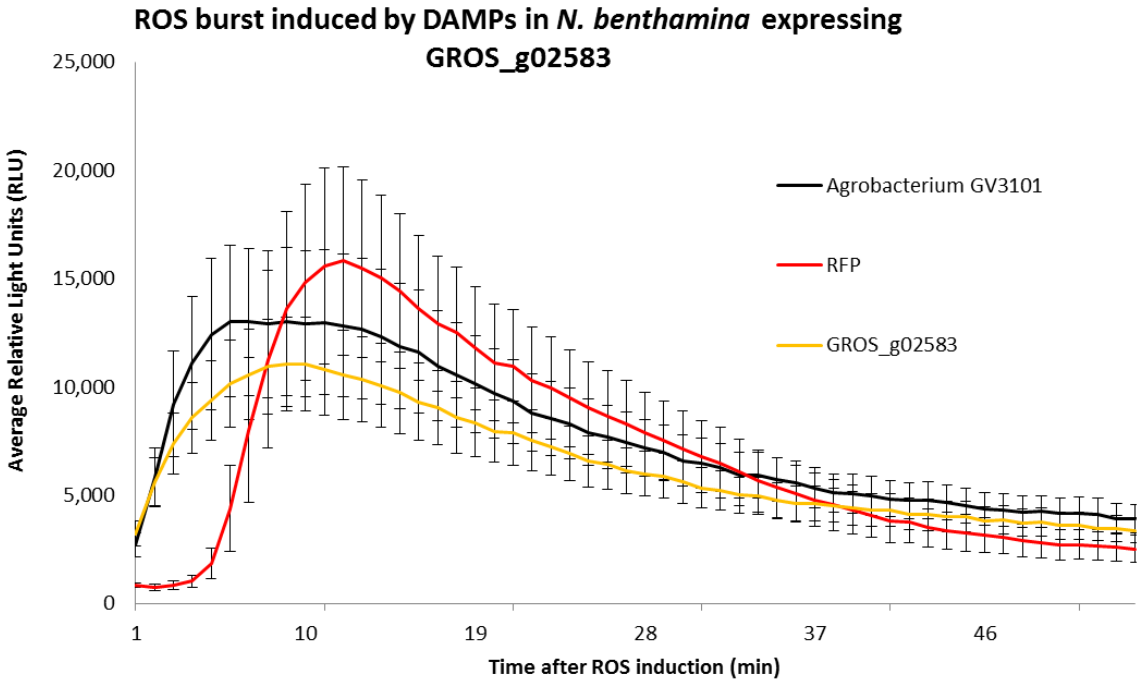


Figure 5.16 – ROS burst induced by oligogalacturonides (DAMPs) in *N. benthamiana* infiltrated by Agrobacterium expressing recombinant GROS_g02583.
Damage associated ROS bursts were initiated using oligogalacturonides. In the presence of GROS_g02583 there was no significant increase of detectable ROS as seen when ROS bursts are initiated with flg-22. Error bars represent standard error.

Leaves of *N. benthamiana* were infiltrated with several GROS_g02583 constructs and controls to test for cell death in response to GROS_g02583 over a longer period of time (Figure 5.17). Two weeks after infection, no cell death was seen in any leaves tested across all original infiltration concentrations.

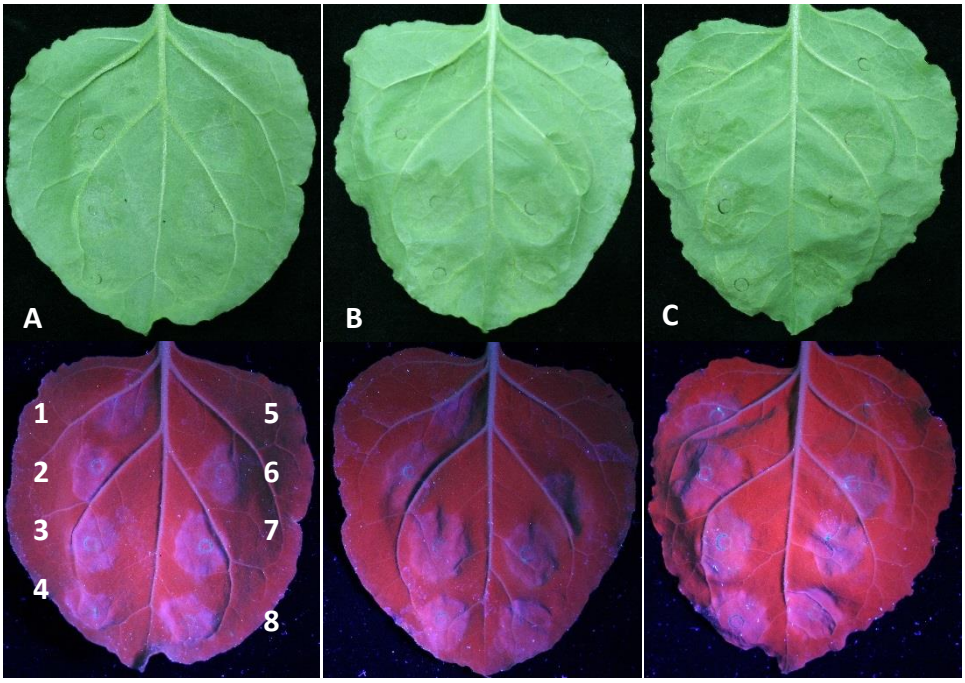


Figure 5.17 – Cell death in leaves of *N. benthamiana* in response to GROS_g02583
No cell death was seen in infected leaves when imaged two weeks after infiltration. **A, B, C** – original infiltration concentrations OD_{600} 0.1, 0.2, 0.3 respectively. **Top –** leaves under white light, **bottom –** leaves under UV light. **1 –** Agrobacterium GV3101, **2 –** GROS_g02583::GFP, **3 –** RFP, **4 –** Signal Peptide::GROS_g02583::STOP, **5 –** Infiltration buffer blank, **6 –** GFP, **7 –** GROS_g02583::STOP, **8 –** Signal Peptide::GROS_g02583::RFP.

5.3.4 Expanding use of a new protocol.

The surface protein labelling method was also used on parasitic stage nematodes that had been collected from potato roots three weeks after infection of the host. Although no usable protein data was collected here it became apparent that this method could be used to identify surface proteins at other PCN life-stages. This could give great insight into further interactions between the parasite and the host after establishing the feeding site. This period is key for the nematode as the expansive growth by the bloating female is possibly the most destructive stage of PCN. For this reason, the surface of the nematode is expected to change at an increased rate to defend against plant responses to a more apparent attack.

Parasitic stage nematodes were surface biotinylated followed by conjugation of biotin with streptavidin-dylight. Parasitic stage surface proteins were bound by biotin similarly to juvenile surface proteins (Figure 5.18). Additionally, it was noted that using this

method it could be possible to check surface proteins between moults (Figure 5.18 C). No usable protein data was collected due to the number of nematodes needed for analysis.

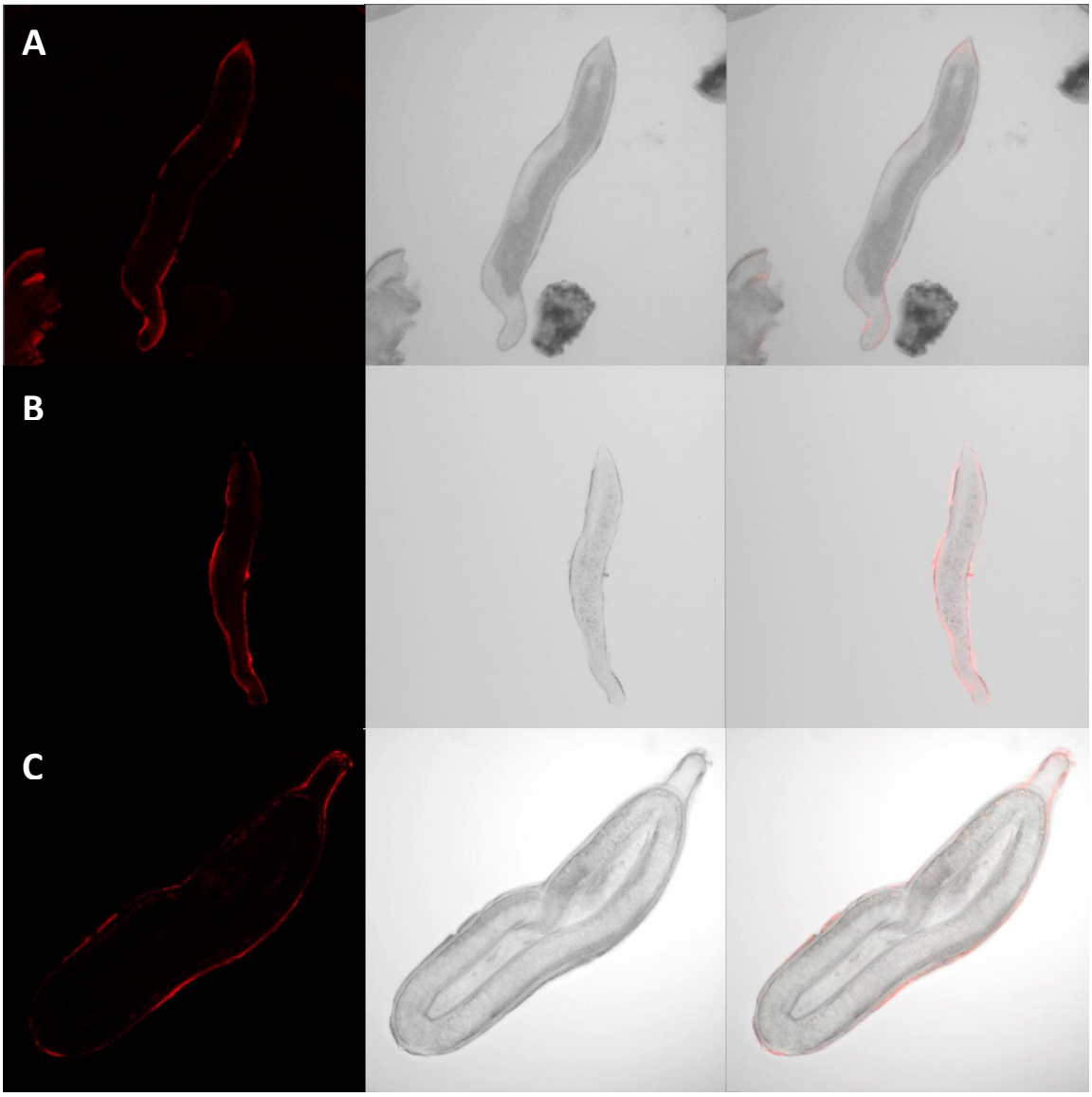


Figure 5.18 – Surface protein biotinylation of parasitic stage *G. rostochiensis*
Surface protein labelling protocols are applicable to both juveniles and parasitic stages. As seen by section **C** with enough nematodes it could be possible to identify the surface proteins between moults, if large scale lifecycle synchronisation was be achieved.

5.3.4.1 Hypodermis associated upstream motif analysis

As explained in the introduction to this chapter there has been recent work identifying upstream promoter motifs that are associated with genes encoding proteins secreted by the nematode (DOGbox, STATAWAARS). Collagens are expressed primarily in the

hypodermis as they are the main component of the cuticle. The *C. elegans* hypodermal collagen (Col-19) was used to identify *G. rostochiensis* collagens based on homology using BLAST. Cuticular collagens are expected to have a signal peptide (Page & Johnstone, 2007), therefore any predicted *G. rostochiensis* collagens that did not have a signal peptide were removed from the list. Collagen associated signal peptides, such as the one associated with *C. elegans* Col-19, are not predictable under the default settings for SignalP. Therefore, the D-cutoff values were lowered from 0.45 to 0.3 to reduce the chances of false negatives. These genes were added to a list of genes expressed in the hypodermis as described in this chapter.

The list was further expanded by adding hypodermal genes localised by undergraduate honours project student Jim Nieuwesteeg. Jim used the same surface protein mass spectrometry data set as used to identify proteins in **Table 5.1**. However, unlike these proteins Jim allowed identified proteins to become candidate proteins even if they lacked a detectable signal peptide. This allowed identification and localisation of three more hypodermis protein candidates expressed in the hypodermis (**Figure 5.19**). The list of genes used for upstream motif detection is shown in **Table 5.3**.

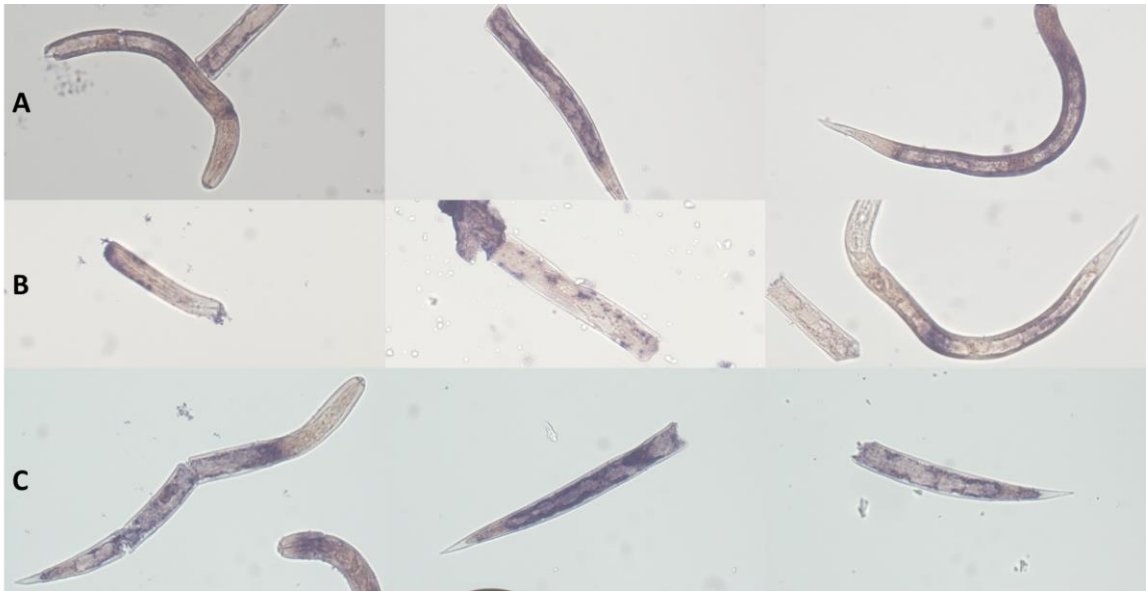


Figure 5.19 – in situ hybridisation of surface protein candidates lacking a detectable signal peptide.

Gene candidates were identified by undergraduate student Jim Nieuwesteeg using the surface protein extraction mass spectrometry data. Unlike the work in this chapter these candidates did not have signal peptides. **A** – GROS_g00815, 32kDa beta-galactoside-binding lectin. **B** – GROS_g01645, Galactoside-binding lectin. **C** – GROS_g12431, Cyclophilin-type peptidyl-prolyl cis-trans isomerase.

Table 5.3 – PCN genes expressed in the hypodermis

Upstream regions for genes expressing the proteins in this table were extracted and analysed for a DNA motif associated with expression in the hypodermis. Expression of genes in this table has been localised to the hypodermis of *G. rostochiensis*. Collagens are predicted to be solely expressed in the hypodermis as they are the primary component of the cuticle. Collagens in this table are only those that are have a predicted signal peptide.

Gene name	Protein function	Reference
GROS_g02490	Glutathione peroxidase	(Jones *et al.*, 2004)
GROS_g014202	Fatty acid and retinol binding protein	(Prior *et al.*, 2001)
GROS_g01193	Peroxiredoxin	(Robertson *et al.*, 2000)
GROS_g02583	Hypothetical protein	**Section 5.3.3**
GROS_g01153	Galectin	**Section 5.3.2**
GROS_g06693	Metalloproteinase inhibitor	**Section 5.3.1**
GROS_g10494	Fatty acid oxidation complex	**Section 5.3, figure 4**

GROS_g01391	Heat shock protein	**Section 5.3, figure 4**
GROS_g05049	Fibronectin	**Section 5.3, figure 4**
GROS_g00815	32kDa beta-galactoside-binding lectin	**Figure 5.19**
GROS_g01645	Galactoside-binding lectin	**Figure 5.19**
GROS_g12431	Cyclophilin-type peptidyl-prolyl cis-trans isomerase	**Figure 5.19**
GROS_g00254	Collagen alpha-1 (IV) chain	
GROS_g01379	Nematode cuticle collagen domain protein	
GROS_g02261	Nematode cuticle collagen domain protein	
GROS_g03922	Nematode cuticle collagen domain protein	
GROS_g04486	Putative cuticular collagen	
GROS_g07249	Nematode cuticle collagen domain-containing protein	
GROS_g07336	Putative cuticular collagen	
GROS_g08577	Nematode cuticular collagen N-terminal domain-containing protein	
GROS_g10748	Putative cuticular collagen	
GROS_g10749	Collagen 5	
GROS_g11030	Cuticular collagen sqt-1	
GROS_g11113	Putative cuticular collagen	
GROS_g11737	Nematode cuticle collagen triple helix repeat protein	
GROS_g11822	Nematode cuticle collagen domain protein	
GROS_g12337	Cuticle collagen lon-3	

Upstream regions for genes in **Table 5.3** were extracted and conserved motifs were calculated between the hypodermal gene list and DOGbox motif containing genes using HOMER.

Upstream regions of 250bp, 500bp, 750bp or 1000bp were individually extracted and analysed for upstream motifs. Any strong motif identified within 500bp upstream should also be identified in regions 750bp or 1000bp upstream but not necessarily 250bp upstream. There were no noticeable sequence imbalances between test genes (**Table 5.3**) and DOGbox control genes. However, one novel motif of interest stood out (**Figure 5.20**). The motif was present in 56% of test upstream regions compared to a negligible 4% of the control regions. This increased to being present in 59% of sequences when the upstream region was expanded from 750bp to 900bp. The motif was searched back against the *G. rostochiensis* genome to pull out all genes with the motif upstream, occurrences of the motif per gene were counted and signal peptides were predicted as before with SignalP (**Figure 5.20**). A slight enrichment for signal peptides was noticed with 17% of sequences that had the motif also having a predicted signal peptide. Annotations for genes containing both the identified motif and a predicted signal peptide can be found in (**7.10 Proteins with predicted hypodermis upstream motif and predicted signal peptide**). 93 of the genes identified to contain the upstream motif were also previously identified in the surface protein extraction mass spectrometry data. A summary of these genes and their annotations can be found in **7.11 Proteins with predicted hypodermis upstream motif also found in surface protein extraction mass spectrometry data**.

```
>TTGCACTCGT
0.001   0.001   0.061   0.937
0.001   0.001   0.061   0.937
0.123   0.123   0.753   0.001
0.062   0.683   0.131   0.124
0.753   0.185   0.001   0.061
0.124   0.690   0.062   0.124
0.001   0.069   0.001   0.929
0.001   0.773   0.103   0.123
0.131   0.185   0.642   0.042
0.061   0.001   0.001   0.937
```

Number of motifs	Number of genes in genome with motif	Number of genes with signal peptide and motif	Number of genes with motif also identified in surface protein mass spectrometry data
1	1964	333	82
2	209	37	9
3	25	6	2
4	1	0	0
5	4	1	0
7	1	0	0

Figure 5.20 – Information for motif upstream of hypodermis genes

Top – graphic of motif showing probabilities of nucleotide per site as per nucleotide probabilities suggested by HOMER, shown in **centre**. **Bottom** – table showing number of sequences containing motifs, motifs and signal peptides or containing motifs and identified in surface protein pulldowns.

5.4 Discussion

The methods used here for surface protein identification were able to provide hits to previously identified surface proteins such as GpX and FAR-1. These act as good positive controls for the methods. Backing up protein identification with localisation of gene expression to the hypodermis gives greater confidence in the results. However, this method is not perfect. Not all transcripts were localised to the hypodermis by *in situ* hybridisation showing that lysis of nematodes during biotinylation can give rise to non-target proteins being present in final protein extractions. Additionally, there is a chance that identified surface protein candidates could be overlooked when processing the mass spectrometry data. For example, all proteins in **Table 5.1** contain a predicted signal peptide for classical secretion. However, as seen with the previously identified peroxiredoxin, not all surface proteins have this detectable signal peptide (Robertson *et al.*, 2000). Furthermore, proteins identified here had extremely high levels of confidence in peptide recognition, lowering these values would increase the number of candidates. Mass spectrometry can only match protein hits to sequences present in a database. If the protein is not in that database or is incorrectly entered into the database, then no match will be made between identified peptide and nematode protein. With the example of GROS_g02583, if protein extractions had been carried out on *G. pallida*, peptide hits for this protein would have been unassigned as the gene had not been correctly mapped to the *G. pallida* transcriptome. Identifying proteins by mass spectrometry in this way is therefore only as reliable as the protein database used for the analysis.

GROS_g06693, previously identified as CRI-2, was described as an oesophageal gland protein (Wang *et al.*, 2001). Since this publication CRI-2 has been recorded as an oesophageal gland protein. However, it was never detected in the gland cells with *in situ* hybridisation. Here, GROS_g06693 was localised to the hypodermis by ISH. The *C. elegans* homologue of this protein regulates response to lipopolysaccharides (Alper *et al.,* 2008). In animal-parasitic species CRI-2 homologues exhibit an unknown function ultimately modulating host innate immunity, possibly by establishing an anti-

inflammatory environment (Cantacessi *et al.*, 2013). However, this protein is predicted to have domains used for inhibiting tissue metalloproteinases. This family of proteins is capable of inhibiting matrix metallopeptidases (MMPs), membrane anchored metalloproteinases and disintegrin-metalloproteinases (Brew & Nagase, 2010). MMPs are documented to control cell death in plants, functioning against microbial infection (Zimmermann *et al.*, 2016; Zhao *et al.*, 2017). Therefore, a role inhibiting these functions is understandable during nematode invasion. As this protein was previously found in excretory/secretory products and has now been localised to the hypodermis, GROS_g06693 was a strong candidate for a defensive surface coat protein. Due to difficulties in recombinantly expressing this protein, potential interacting host proteins were not identified. However, by recombinantly expressing the protein within leaves of *N. benthamiana* with *Agrobacterium*, the ability of GROS_g06693 to moderate host defences could be monitored through use of a ROS assay. Results from these assays were mostly inconclusive. When data from all the assays for GROS_g06693 were combined and averaged they show that there is no significant difference in ROS production with or without the presence of GROS_g06693 inside the cell or apoplast. This is expected as metalloproteinases are not known to be key components of plant defences in response to PAMPs or DAMPs. Instead, it is possible that GROS_g06693 acts to inhibit MMPs attempting to facilitate cell death, keeping potential feeding sites available for an extended time.

Attempts to identify MMPs that may be inhibited by GROS_g06693 using RFP tag pull downs from apoplastic fluid extractions were unsuccessful. This is possibly due to MMPs associated with plant cell death being membrane anchored. They may therefore not have been extracted within the apoplastic fluid. Alternatively, the large size of the RFP tag used could inhibit binding of the inhibitor with its target MMP. Alternate techniques such as yeast-2-hybrid could be used to identify protein interactions between the inhibitor and target MMP. However, apoplastic protein libraries for yeast-2-hybrid analysis were not readily available at the time of this work.

Although the protein being inhibited by GROS_g06693 was not identified other information about this protein has now been obtained. The transcript has now been localised to the hypodermis. Combining this with identification in surface protein pull downs and previous work identifying the protein in excretory/secretory products gives stronger evidence for GROS_g06693 being found in the surface coat of PCN. GROS_g06693 does not directly interfere with host PAMP or DAMP triggered defences but instead has the potential to inhibit host-initiated cell death that could otherwise halt the parasite lifecycle.

The galectin, GROS_g01153, was a strong candidate for further analysis from the mass spectrometry data as it is easy to see a potential function for this protein within the PCN life cycle. Typically, galectins would be expected to be expressed within the gut or oesophageal gland (Dubreuil *et al.*, 2007), where they would have interactions aiding with digestion of host carbohydrates. However, GROS_g01153 has a predicted signal peptide and its mRNA was clearly localised to the hypodermis of *G. rostochiensis*. One role for a hypodermally secreted lectin would be binding to the nematode's own surface glycocalyx, rendering the parasites surface glycosylation undetectable to the host. Galectin binding to nematode surfaces has been widely described before, however, this work was mainly carried out to identify the presence of the surface glycocalyx (McClure & Stynes, 1988; Davis *et al.*, 1989; Aroch *et al.*, 2017). An alternative function for a surface released galectin could be associated with removing host damage associated molecular patterns. This would allow the nematode to reduce the ability of the host to induce a defence response upon nematode facilitated destruction of root tissues when travelling to a feeding site.

Using cell wall break down products to induce a plant defence response showed a reduction in detectable ROS in leaf discs expressing GROS_g01153 within the cell. This can be directly applied to the damaging trail taken by PCN through the host roots. If cell walls are broken, releasing DAMPs as the nematode is moving through the root both intra- and inter-cellular space would become combined. Therefore, GROS_g01153 only visibly functioning when expressed within the plant cell could simply reflect more

favourable conditions for this protein function within the intra-cellular environment. However, in this assay there is no cell damage causing mixing of cytoplasm and apoplast. If GROS_g01153 is directly binding to cell wall products then sending the protein to the apoplast could mean the protein binds to the abundance of cell wall material present in the intercellular space. Therefore, upon introduction of oligogalacturonide DAMPs in the assay, there is little GROS_g01153 available that is not already bound to a polysaccharide.

Unfortunately, purified recombinant GROS_g01153 could not be obtained. The high rare codon usage in this protein caused premature termination of translation of the protein even in tRNA optimised *E. coli* cell lines (BL21). If recombinant protein had been obtained it would have been possible to assay carbohydrate binding. Binding to carbohydrates that are detected as DAMPs would suggest that this protein does indeed decrease plant defence responses by reducing the amount of DAMPs available for detection by the host.

It is worth noting that once the protein identification pipeline restrictions were eased slightly by allowing proteins to become candidates regardless of the presence of a detectable signal peptide, two further lectins were identified in surface protein extractions. These lectins (GROS_g00815 and GROS_g01645) have expression patterns heightened at the juvenile life-stage and were also both localised to hypodermis of *G. rostochiensis*. Although these galectins did not undergo any further analysis after localisation, their presence in both the hypodermis and surface protein extraction data suggests that an array of lectins may be present on the *G. rostochiensis* surface. In addition to these galectins, once the requirement for a signal peptide presence had been removed, peroxiredoxin could also be found in the protein mass spectrometry data, further demonstrating the validity of the technique.

The predicted protein GROS_g02583 is a good candidate for a surface coat protein identified by this new technique. Although an exact role for this protein has not been identified, the work carried out here has begun to characterise this interesting pioneer protein. GROS_g02583 was cloned from cDNA after identification in surface coat protein

pull downs. This protein is highly conserved across many nematodes, both parasitic and free living. However, unlike free living nematode homologues of this protein, GROS_g02583 shows large upregulation of gene expression towards later stages of nematode development. Sixteen cysteines are conserved across all homologues of GROS_g02583 in all examined nematodes. These cysteines all appear to form di-sulphide bridges with other cysteines within the protein suggesting their function is structural rather than mediating interactions with other molecules. The high level of conservation suggests that this protein has a key role within nematode biology.

Within a plant setting, GROS_g02583 expressed by *Agrobacterium* aggregated unless in the apoplast. This possibly reflects the more favourable conditions for this protein in an apoplastic environment, for example, a harsher more acidic pH found within the apoplast may aid regular protein folding (Grignon & Sentenac, 1991). Similarly, disulphide bonds are found most commonly in extracellular proteins, aggregation of GROS_g02583 in the cytoplasm may represent these bonds not forming properly (Hatahet *et al.,* 2010). GROS_g02583 does not appear to contribute to plant cell death. However, when present in the apoplast, plant defence responses to flg-22 are hugely increased. The same effect was not seen when defence responses were induced by DAMPs. Originally, this large response to flg-22 when in the presence of GROS_g02583 was thought to be due to a plant priming response caused by GROS_g02583. However, when tested, GROS_g02583 does not appear to induce plant defences alone.

There is the potential to expand these surface protein extraction protocols to other nematodes and to other life-stages of PCN. As with any nematode protein identification by mass spectrometry, procuring large quantities of nematodes is key to obtaining reliable results. While attempting to extract sufficient numbers of parasitic stage PCN for surface coat extractions, it became clear that contamination with host material could drastically affect protein identification. Nematode shaped plant material was often collected with parasitic nematode samples, although very little of this will match to predicted nematode proteins, the presence of these proteins alone is enough to skew

the mass spectrometer into identifying these proteins rather than target sequences. This data then accumulates as unassigned protein hits.

Upstream regions of proteins expressed in the hypodermis were extracted and analysed for conserved motifs. It was assumed that any motif identified here would differ from the DOG box motif associated with dorsal gland expression. If such a motif could be found then other genes expressed in the hypodermis could be withdrawn from the genome data, this could include new candidates for defensive surface coat proteins. A potential motif was suggested (TTGCACTCGT). The motif was chosen after being consistently identified in over half of the target hypodermis sequences. As this motif is suggested for all hypodermal genes (secreted and non-secreted) it is not possible to support motif findings with the presence of signal peptide enrichment unlike when searching for gland cell motifs (DOG box). Ultimately, hypodermal expression and the presence of this motif could be correlated by localising gene expression for genes containing the motif with *in situ* hybridisation. However, the high number of sequences returned as containing this motif and the overall sequence noise within the motif reduce the reliability of the identified sequence. Similarly, not all genes identified in this chapter as being present in the hypodermis, for example GROS_g02583, contain any copies of the motif upstream.

It is still likely that an upstream motif corresponding to expression in the hypodermis of nematodes does exist. The work in this chapter has significantly increased the number of genes known to be expressed in the hypodermis allowing the first searches for such a motif to be carried out. However, the current list of genes may be still too small to reliably identify any motifs. The current gene list (**Table 5.3**) could be expanded by adding genes identified in surface protein mass spectrometry data (**Table 5.1**) but without localising the gene expression there is a strong chance of identifying false positives. It is hoped that with continued research in this area such a motif could eventually be identified, allowing for nematode surface coat proteins to be identified from genomic data.

5.5 References

ALFANO, J.R. & COLLMER, A. 2004. TYPE III SECRETION SYSTEM EFFECTOR PROTEINS: Double Agents in Bacterial Disease and Plant Defense. *Annual Review of Phytopathology*, **42**, 385–414.

ALPER, S., LAWS, R., LACKFORD, B., BOYD, W.A., DUNLAP, P., FREEDMAN, J.H. & SCHWARTZ, D.A. 2008. Identification of innate immunity genes and pathways using a comparative genomics approach. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 7016–7021.

ANDREI, C., DAZZI, C., LOTTI, L., TORRISI, M.R., CHIMINI, G. & RUBARTELLI, A. 1999. The secretory route of the leaderless protein interleukin 1beta involves exocytosis of endolysosome-related vesicles. *Molecular biology of the cell*, **10**, 1463–1475.

AROCH, I., AROGETI, I., MARCOVICS, A., SPIEGEL, Y. & LAVY, E. 2017. In vitro lectin binding to the outer surface of Spirocerca lupi at different life-stages. *Veterinary Parasitology*, **235**, 94–99.

BAUTERS, L., HAEGEMAN, A., KYNDT, T. & GHEYSEN, G. 2014. Analysis of the transcriptome of Hirschmanniella oryzae to explore potential survival strategies and host-nematode interactions. *Molecular Plant Pathology*, **15**, 352–363.

BENDTSEN, J.D., JENSEN, L.J., BLOM, N., VON HEIJNE, G. & BRUNAK, S. 2004. Feature-based prediction of non-classical and leaderless protein secretion. *Protein Engineering Design and Selection*, **17**, 349–356.

BLANC-MATHIEU, R., PERFUS-BARBEOCH, L., AURY, J., DA ROCHA, M., GOUZY, J., SALLET, E., MARTIN-JIMENEZ, C., … DANCHIN, E.G.J. 2017. Hybridization and polyploidy enable genomic plasticity without sex in the most devastating plant-parasitic nematodes. Gojobori, T., *ed. PLOS Genetics*, **13**.

BLAXTER, M.L. & MAIZELS, R.M. 1992. Nematode Surface Coats: Actively Evading Immunity. *Parasitology Today,* **8**, 243-247.

BOLGER, A.M., LOHSE, M. & USADEL, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, **30**, 2114–2120.

BREW, K. & NAGASE, H. 2010. The tissue inhibitors of metalloproteinases (TIMPs): an ancient family with structural and functional diversity. *Biochimica et biophysica*

acta, **1803**, 55–71.

C. ELEGANS SEQUENCING CONSORTIUM. 1998. Genome sequence of the nematode C. elegans: a platform for investigating biology. *Science (New York, N.Y.)*, **282**, 2012–2018.

CANTACESSI, C., HOFMANN, A., PICKERING, D., NAVARRO, S., MITREVA, M. & LOUKAS, A. 2013. TIMPs of parasitic helminths - a large-scale analysis of high-throughput sequence datasets. *Parasites & vectors*, **6**, 156.

CAPELLA-GUTIÉRREZ, S., SILLA-MARTÍNEZ, J.M. & GABALDÓN, T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, **25**, 1972–1973.

COLEMAN, J., INUKAI, M. & INOUYE, M. 1985. Dual functions of the signal peptide in protein transfer across the membrane. *Cell*, **43**, 351–360.

COTTON, J.A., BENNURU, S., GROTE, A., HARSHA, B., TRACEY, A., BEECH, R., DOYLE, S.R., … LUSTIGMAN, S. 2017. The genome of Onchocerca volvulus, agent of river blindness. *Nature Microbiology*, **2**, 16216.

COTTON, J.A., LILLEY, C.J., JONES, L.M., KIKUCHI, T., REID, A.J., THORPE, P., TSAI, I.J., … URWIN, P.E. 2014. The genome and life-stage specific transcriptomes of Globodera pallida elucidate key aspects of plant parasitism by a cyst nematode. *Genome biology*, **15**, R43, 10.1186/gb-2014-15-3-r43.

DAVIES, K.G. & CURTIS, R.H.C. 2011. Cuticle Surface Coat of Plant-Parasitic Nematodes. *Annual Review of Phytopathology*, **49**, 135–156.

DAVIS, E.L., KAPLAN, D.T., DICKSON, D.W. & MITCHELL, D.J. 1989. Root Tissue Response of Two Related Soybean Cultivars to Infection by Lectin-treated Meloidogyne spp. *Journal of nematology*, **21**, 219–228.

DUBREUIL, G., MAGLIANO, M., DELEURY, E., ABAD, P. & ROSSO, M.N. 2007. Transcriptome analysis of root-knot nematode functions induced in the early stages of parasitism. *New Phytologist*, **176**, 426–436.

ENDO, B.Y. & WYSS, U. 1992. Ultrastructure of cuticular exudations in parasitic juvenileHeterodera schachtii, as related to cuticle structure. *Protoplasma*, **166**, 67–77.

ESPADA, M., EVES-VAN DEN AKKER, S., MAIER, T., VIJAYAPALANI, P., BAUM, T., MOTA, M. & JONES,

J.T. 2018. STATAWAARS: a promoter motif associated with spatial expression in the major effector-producing tissues of the plant-parasitic nematode Bursaphelenchus xylophilus. *BMC Genomics*, **19**, 553.

EVES-VAN DEN AKKER, S., LAETSCH, D.R., THORPE, P., LILLEY, C.J., DANCHIN, E.G.J., DA ROCHA, M., RANCUREL, C., … JONES, J.T. 2016. The genome of the yellow potato cyst nematode, Globodera rostochiensis, reveals insights into the basis of parasitism and virulence. *Genome Biology*, **17**, 124.

EVES-VAN DEN AKKER, S., LILLEY, C.J., YUSUP, H.B., JONES, J.T. & URWIN, P.E. 2016. Functional C-TERMINALLY ENCODED PEPTIDE (CEP) plant hormone domains evolved de novo in the plant parasite Rotylenchulus reniformis. *Molecular Plant Pathology*, **17**, 1265–1275.

FIORETTI, L., PORTER, A., HAYDOCK, P.J. & CURTIS, R. 2002. Monoclonal antibodies reactive with secreted–excreted products from the amphids and the cuticle surface of Globodera pallida affect nematode movement and delay invasion of potato roots. *International Journal for Parasitology*, **32**, 1709–1718.

GHEDIN, E., WANG, S., SPIRO, D., CALER, E., ZHAO, Q., CRABTREE, J., ALLEN, J.E., … SCOTT, A.L. 2007. Draft Genome of the Filarial Nematode Parasite Brugia malayi. *Science*, **317**, 1756–1760.

GODEL, C., KUMAR, S., KOUTSOVOULOS, G., LUDIN, P., NILSSON, D., COMANDATORE, F., WROBEL, N., … MÄSER, P. 2012. The genome of the heartworm, Dirofilaria immitis, reveals drug and vaccine targets. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, **26**, 4650–4661.

GRAVATO-NOBRE, M.J., MCCLURE, M.A., DOLAN, L., CALDER, G., DAVIES, K.G., MULLIGAN, B., EVANS, K. & VON MENDE, N. 1999. Meloidogyne incognita Surface Antigen Epitopes in Infected Arabidopsis Roots. *Journal of nematology*, **31**, 212–223.

GRIGNON, C. & SENTENAC, H. 1991. pH and Ionic Conditions in the Apoplast. *Annual Review of Plant Physiology and Plant Molecular Biology*, **42**, 103–128.

GUHATHAKURTA, D., SCHRIEFER, L.A., WATERSTON, R.H. & STORMO, G.D. 2004. Novel transcription regulatory elements in Caenorhabditis elegans muscle genes. *Genome research*, **14**, 2457–2468.

HARRIS, C. & HANSEN, J.M. 2012. Oxidative Stress, Thiols, and Redox Profiles. *In Methods in molecular biology (Clifton, N.J.)*. 325–346.

HARTMANN, S. & LUCIUS, R. 2003. Modulation of host immune responses by nematode cystatins. *International Journal for Parasitology*, **33**, 1291–1302.

HATAHET, F., NGUYEN, V., SALO, K.E.H. & RUDDOCK, L.W. 2010. Disruption of reducing pathways is not essential for efficient disulfide bond formation in the cytoplasm of E. coli. *Microbial cell factories*, **9**, 67.

HÖFLICH, J., BERNINSONE, P., GÖBEL, C., GRAVATO-NOBRE, M.J., LIBBY, B.J., DARBY, C., POLITZ, S.M., HODGKIN, J., HIRSCHBERG, C.B. & BAUMEISTER, R. 2004. Loss of srf-3-encoded Nucleotide Sugar Transporter Activity in Caenorhabditis elegans Alters Surface Antigenicity and Prevents Bacterial Adherence.

HU, G.G., MCCLURE, M.A. & SCHIMITT, M.E. 2000. Origin of a Meloidogyne incognita Surface Coat Antigen. *Journal of nematology*, **32**, 174–182.

IBERKLEID, I., VIEIRA, P., DE ALMEIDA ENGLER, J., FIRESTER, K., SPIEGEL, Y. & HOROWITZ, S.B. 2013. Fatty Acid-and Retinol-Binding Protein, Mj-FAR-1 Induces Tomato Host Susceptibility to Root-Knot Nematodes. Ghanim, M., *ed. PLoS ONE*, **8**.

JONES, J.T., REAVY, B., SMANT, G. & PRIOR, A.E. 2004. Glutathione peroxidases of the potato cyst nematode Globodera Rostochiensis. *Gene*, **324**, 47–54.

KARIMI, M., INZÉ, D. & DEPICKER, A. 2002. GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends in plant science*, **7**, 193–195.

KIKUCHI, T., COTTON, J.A., DALZELL, J.J., HASEGAWA, K., KANZAKI, N., MCVEIGH, P., TAKANASHI, T., … BERRIMAN, M. 2011. Genomic insights into the origin of parasitism in the emerging plant pathogen Bursaphelenchus xylophilus. *PLoS pathogens*, **7**.

KIKUCHI, T., EVES-VAN DEN AKKER, S. & JONES, J.T. 2017. Genome Evolution of Plant-Parasitic Nematodes. *Annual Review of Phytopathology*, **55**, 333–354.

KULKARNI, A., HOLZ, A., RÖDELSPERGER, C., HARBECKE, D. & STREIT, A. 2016. Differential chromatin amplification and chromosome complements in the germline of Strongyloididae (Nematoda). *Chromosoma*, **125**, 125–136.

LAING, R., KIKUCHI, T., MARTINELLI, A., TSAI, I.J., BEECH, R.N., REDMAN, E., HOLROYD, N., … COTTON, J.A. 2013. The genome and transcriptome of Haemonchus contortus, a key

model parasite for drug and vaccine discovery. *Genome Biology*, **14**, R88.

LANGMEAD, B. & SALZBERG, S.L. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods*, **9**, 357–359.

LE, X., WANG, X., GUAN, T., JU, Y. & LI, H. 2016. Isolation and characterization of a fatty acid- and retinoid-binding protein from the cereal cyst nematode Heterodera avenae. *Experimental Parasitology*, **167**, 94–102.

LIEBAU, E., WILDENBURG, G., WALTER, R.D., HENKLE-DÜHRSEN, K., WALTER, R.D. & LIEBAU, E. 1994. A novel type of glutathione S-transferase in Onchocerca volvulus. *Infection and immunity*, **62**, 4762–4767.

MAIZELS, R.M. 2013. Toxocara canis: molecular basis of immune recognition and evasion. *Veterinary parasitology*, **193**, 365–374.

MCCARTHY, D.J., CHEN, Y. & SMYTH, G.K. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research*, **40**, 4288–4297.

MCCLURE, M.A. & STYNES, B.A. 1988. Lectin binding sites on the amphidial exudates of meloidogyne. *Journal of nematology*, **20**, 321–326.

MCNULTY, S.N., STRÜBE, C., ROSA, B.A., MARTIN, J.C., TYAGI, R., CHOI, Y., WANG, Q., … MITREVA, M. 2016. Dictyocaulus viviparus genome, variome and transcriptome elucidate lungworm biology and support future intervention. *Scientific Reports*, **6**, 20316.

MEI, Y., WRIGHT, K.M., HAEGEMAN, A., BAUTERS, L., DIAZ-GRANADOS, A., GOVERSE, A., GHEYSEN, G., JONES, J.T. & MANTELIN, S. 2018. The Globodera pallida SPRYSEC Effector GpSPRY-414-2 That Suppresses Plant Defenses Targets a Regulatory Component of the Dynamic Microtubule Network. *Frontiers in Plant Science*, **9**, 1019.

MILNE, I., LINDNER, D., BAYER, M., HUSMEIER, D., MCGUIRE, G., MARSHALL, D.F. & WRIGHT, F. 2009. TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics*, **25**, 126–127.

MODHA, J., ROBERTS, M.C., ROBERTSON, W.M., SWEETMAN, G., POWELL, K.A., KENNEDY, M.W. & KUSEL, J.R. 1999. The surface coat of infective larvae of Trichinella spiralis. *Parasitology*, **118**, 509–522.

NICKEL, W. 2003. The mystery of nonclassical protein secretion. *European Journal of*

Biochemistry, **270**, 2109–2119.

O'LEARY, B.M., RICO, A., MCCRAW, S., FONES, H.N. & PRESTON, G.M. 2014. The infiltration-centrifugation technique for extraction of apoplastic fluid from plant leaves using Phaseolus vulgaris as an example. *Journal of visualized experiments : JoVE*.

PAGE, A.P. & JOHNSTONE, I.L. 2007. The Cuticle. *In WormBook*.

PHANI, V., SHIVAKUMARA, T.N., DAVIES, K.G. & RAO, U. 2017. Meloidogyne incognita Fatty Acid- and Retinol- Binding Protein (Mi-FAR-1) Affects Nematode Infection of Plant Roots and the Attachment of Pasteuria penetrans Endospores. *Frontiers in Microbiology*, **8**, 2122.

PHILLIPS, W.S., EVES-VAN DEN AKKER, S. & ZASADA, I.A. 2017. Draft Transcriptome of Globodera ellingtonae. *Journal of nematology*, **49**, 129–130.

PRIOR, A., JONES, J.T., BLOK, V.C., BEAUCHAMP, J., MCDERMOTT, L., COOPER, A. & KENNEDY, M.W. 2001. A surface-associated retinol- and fatty acid-binding protein (Gp-FAR-1) from the potato cyst nematode Globodera pallida: lipid binding activities, structural analysis and expression pattern. *The Biochemical journal*, **356**, 387–394.

QUINLAN, A.R. & HALL, I.M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.

RIDDLE, D.L., BLUMENTHAL, T., MEYER, B.J. & PRIESS, J.R. 1997. Introduction to C. elegans, The Nematode Surface. In C. elegans II edited by D. L. Riddle. **33**. Accessed from: https://www.ncbi.nlm.nih.gov/books/NBK20022/

ROBERTSON, L., ROBERTSON, W.M., SOBCZAK, M., HELDER, J., TETAUD, E., ARIYANAYAGAM, M.R., FERGUSON, M.A., FAIRLAMB, A. & JONES, J.T. 2000. Cloning, expression and functional characterisation of a peroxiredoxin from the potato cyst nematode Globodera rostochiensis. *Molecular and biochemical parasitology*, **111**, 41–49.

ROBINSON, M.D., MCCARTHY, D.J. & SMYTH, G.K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.

ROBINSON, M.P., ATKINSON, H.J. & PERRY, R.N. 1985. The Effect of Delayed Emergence On Infectivity of Juveniles of the Potato Cyst Nematode Globodera Rostochiensis. *Nematologica*, **31**, 171–178.

SCHWARZ, E.M., HU, Y., ANTOSHECHKIN, I., MILLER, M.M., STERNBERG, P.W. & AROIAN, R. V. 2015. The genome and transcriptome of the zoonotic hookworm Ancylostoma ceylanicum identify infection-specific gene families. *Nature Genetics*, **47**, 416–422.

SERENA, C., CALVO, E., CLARES, M.P., DIAZ, M.L., CHICOTE, J.U., BELTRÁN-DEBON, R., FONTOVA, R., RODRIGUEZ, A., GARCÍA-ESPAÑA, E. & GARCÍA-ESPAÑA, A. 2015. Significant in vivo anti-inflammatory activity of Pytren4Q-Mn a superoxide dismutase 2 (SOD2) mimetic scorpiand-like Mn (II) complex. *PloS one*, **10**.

SHAH, S.J., ANJAM, M.S., MENDY, B., ANWER, M.A., HABASH, S.S., LOZANO-TORRES, J.L., GRUNDLER, F.M.W. & SIDDIQUE, S. 2017. Damage-associated responses of the host contribute to defence against cyst nematodes but not root-knot nematodes. *Journal of experimental botany*, **68**, 5949–5960.

SHAMSELDEAN, M.M., PLATZER, E.G. & GAUGLER, R. 2007. Role of the surface coat of Romanomermis culicivorax in immune evasion. *Nematology*, **9**, 17–24.

SHARON, E., SPIEGEL, Y., SALOMON, R. & CURTIS, R.H.C. 2002. Characterization of Meloidogyne javanica surface coat with antibodies and their effect on nematode behaviour. *Parasitology*, **125**, 177–185.

SOMMER, A., RICKERT, R., FISCHER, P., STEINHART, H., WALTER, R.D. & LIEBAU, E. 2003. A Dominant Role for Extracellular Glutathione S-Transferase from Onchocerca volvulus Is the Production of Prostaglandin D 2. Infection and immunity, **71**, 3603–3606.

SOUZA, C., LI, S., LIN, A.Z., BOUTROT, F., GROSSMANN, G., ZIPFEL, C. & SOMERVILLE, S.C. 2017. Cellulose-Derived Oligomers Act as Damage-Associated Molecular Patterns and Trigger Defense-Like Responses. *Plant Physiology*, **173**, 2383–2398.

SPIEGEL, Y. & MCCLURE, M.A. 1995. The surface coat of plant-parasitic nematodes: chemical composition, origin, and biological role-a review. *Journal of nematology*, **27**, 127–134.

SPIEGEL, Y., MOR, M. & SHARON, E. 1996. Attachment of Pasteuria penetrans Endospores to the Surface of Meloidogyne javanica Second-stage Juveniles. *Journal of nematology*, **28**, 328–334.

TANG, L., OU, X., HENKLE-DÜHRSEN, K. & SELKIRK, M.E. 1994. Extracellular and cytoplasmic

CuZn superoxide dismutases from Brugia lymphatic filarial nematode parasites. *Infection and immunity*, **62**, 961–967.

TANG, Y.T., GAO, X., ROSA, B.A., ABUBUCKER, S., HALLSWORTH-PEPIN, K., MARTIN, J., TYAGI, R., … MITREVA, M. 2014. Genome of the human hookworm Necator americanus. *Nature Genetics*, **46**, 261–269.

TIAN, B., YANG, J. & ZHANG, K. 2007. Bacteria used in the biological control of plant-parasitic nematodes: populations, mechanisms of action, and future prospects. *FEMS Microbiology Ecology*, **61**, 197–213.

VAN DER FITS, L., DEAKIN, E.A., HOGE, J.H.C. & MEMELINK, J. 2000. The ternary transformation system: Constitutive virG on a compatible plasmid dramatically increases Agrobacterium-mediated plant transformation. *Plant Molecular Biology*, **43**, 495–502.

VIEIRA, P., EVES-VAN DEN AKKER, S., VERMA, R., WANTOCH, S., EISENBACK, J.D. & KAMO, K. 2015. The Pratylenchus penetrans Transcriptome as a Source for the Development of Alternative Control Strategies: Mining for Putative Genes Involved in Parasitism and Evaluation of in planta RNAi. *PLOS ONE*, **10**.

WANG, X., ALLEN, R., DING, X., GOELLNER, M., MAIER, T., DE BOER, J.M., BAUM, T.J., HUSSEY, R.S. & DAVIS, E.L. 2001. Signal Peptide-Selection of cDNA Cloned Directly from the Esophageal Gland Cells of the Soybean Cyst Nematode Heterodera glycines. *Molecular Plant-Microbe Interactions*, **14**, 536–544.

WANG, Y., WU, L., LIU, X., WANG, S., EHSAN, M., YAN, R., SONG, X., XU, L. & LI, X. 2017. Characterization of a secreted cystatin of the parasitic nematode Haemonchus contortus and its immune-modulatory effect on goat monocytes. *Parasites & vectors*, **10**, 425.

WHISSON, S.C., BOEVINK, P.C., MOLELEKI, L., AVROVA, A.O., MORALES, J.G., GILROY, E.M., ARMSTRONG, M.R., … BIRCH, P.R.J. 2007. A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature*, **450**, 115–118.

ZHAO, P., ZHANG, F., LIU, D., IMANI, J., LANGEN, G. & KOGEL, K. 2017. Matrix metalloproteinases operate redundantly in Arabidopsis immunity against necrotrophic and biotrophic fungal pathogens. *PLOS ONE*, **12**.

ZHENG, J., PENG, D., CHEN, L., LIU, H., CHEN, F., XU, M., JU, S., RUAN, L. & SUN, M. 2016. The Ditylenchus destructor genome provides new insights into the evolution of plant parasitic nematodes. *Proceedings. Biological sciences*, **283**.

ZHU, X., KORHONEN, P.K., CAI, H., YOUNG, N.D., NEJSUM, P., VON SAMSON-HIMMELSTJERNA, G., BOAG, P.R., … GASSER, R.B. 2015. Genetic blueprint of the zoonotic pathogen Toxocara canis. *Nature Communications*, **6**, 6145.

ZIMMERMANN, D., GOMEZ-BARRERA, J.A., PASULE, C., BRACK-FRICK, U.B., SIEFERER, E., NICHOLSON, T.M., PFANNSTIEL, J., STINTZI, A. & SCHALLER, A. 2016. Cell Death Control by Matrix Metalloproteinases. *Plant physiology*, **171**, 1456–1469.

6. General discussion

6.1 Recapitulation

The overall aim of this thesis was to expand our current knowledge of structures related to the protection of developing PCN juveniles. Primarily this was done by characterising molecular aspects of the PCN eggshell. The eggshell has been overlooked in recent years with research favouring analysis of parasite-host interactions through identification and functional characterisation of effector proteins. Research into PCN hatching is often orientated around identifying hatching elicitors originating from host plants and not how these molecules could interact with the parasite. Physiological studies have allowed the nature of the changes that are induced during hatching to be characterised (Perry, 2002). However, until now it has not been possible to analyse the functional basis of hatching at the molecular level. Identifying and localising proteins within the eggshell could help develop a better understanding of the PCN hatching mechanisms. Lipid extractions from eggshells could offer further information into how the permeability barrier is formed and how this could relate in response to host cues.

The juvenile surface coat is a barrier between the nematode and its immediate surroundings (Davies & Curtis, 2011). If the parasite is to reach and remain active at an appropriate feeding site this barrier must be adaptive and help prevent recognition by the host or suppress responses if recognition does occur. Identifying proteins present on the surface of the juvenile that protect the nematode during infection of the host would allow for a better understanding of these host-parasite interactions.

6.2 Elaboration

Mass spectrometry has played a large role in this research. The improvement of this technology over the last two decades has increased sensitivity of spectrometers allowing identification of proteins on a much finer scale. Smaller sample sizes were required for sufficient protein detection. Protein identification by mass spectrometry is

only as powerful as the proteomic database that peptides are compared to. Differences of a few amino acids in the database could result in identified peptides not being associated to a protein. For example, if the surface coat protein samples had been searched against the *G. pallida* transcriptome instead of the *G. rostochiensis* one then the hypothetical protein GROS_g02583 would never have been identified as this gene was not mapped in the original transcriptome. Fortunately advances including cost reduction and increased reliability in long sequence reads means genome technology is becoming more available and may not be an issue in future work.

Combining these advanced technologies with software that integrates identified peptide data with a species-specific protein database makes the process more user friendly.

The availability and type of spectrometer possibly restricted some aspects of the work. Firstly, when carrying out GC-MS to identify FAMEs, contaminants from column bleed were noticeable. Ultimately, a new column would solve this issue, however, this comes with additional costs. An AB-Sciex QTRAP 4000 was used for identification of lipids by ESI-MS-MS. Contamination in this machine was also noticeable, possibly due to its age. Newer, more refined QTRAP spectrometers exist, but it is unfeasible to replace such a machine easily.

6.2.1 The PCN surface coat

Nematode surface coat extractions are difficult. Conditions that are too disruptive will rupture the nematode resulting in sample contamination with internal body contents whereas conditions not harsh enough will not strip enough detectable material from the surface coat. This is reflected in studies of secreted proteins from other nematodes. For example, Shinya *et al.* (2013) analysed the secretome of the pine wilt nematode *Bursaphelenchus xylophilus*. Only 625 of the 1515 “secreted” proteins identified in this study had a predicted signal peptide and the list of secreted proteins included numerous obvious housekeeping proteins that are clearly unlikely to be part of the secreted protein profile of this nematode, even allowing for non-classical secretion pathways. If the nematodes are fixed in paraformaldehyde before surface coat extractions then

surface proteins will become cross linked, potentially modifying how the peptide will be recognised by the mass spectrometer. However, carrying out the extractions on active juveniles will result in contamination with proteins from oesophageal glands, skewing the focus of the mass spectrometer. Biotin labelling of the surface coat was developed to allow labelling of surface proteins which could then be pulled down from a solution after various extraction techniques.

Glutathione peroxidase (GROS_g02490) and fatty acid and retinol binding protein (Gr-FAR-1, GROS_g14202) are both known to be on the surface of various plant-parasitic nematodes (Prior *et al.*, 2001; Jones *et al.*, 2004). Identification of both of these proteins in the surface coat extractions performed here provided a good indication that the developed method is extracting surface proteins. Using antibodies raised against Gp-FAR-1 (*G. pallida*) allowed confirmation that Gr-FAR-1 is also localised to the surface coat. Although this is not surprising, it is the first time FAR-1 has been localised to the surface coat of *G. rostochiensis*.

Protein extractions were carried out 5 separate times. Peptides identified by mass spectrometry had a confidence interval of 95% and total protein confidence interval of 99% as decided by the default settings on Scaffold 4 proteome software. Candidate proteins were selected if the protein was identified in at least 3 protein extractions and had a signal peptide. Reducing these strict requirements for candidate proteins would have increased the number of potential surface proteins but may have allowed false positive signals to be labelled as potentially genuine. Expression of 16 genes was localised by *in situ* hybridisation. Of these 16, 11 were expressed in the hypodermis, providing reassurance that the criteria used to identify candidate proteins were appropriate.

Gene expression was not localised to the hypodermis for all candidate genes. This suggested that some lysis of nematodes had occurred throughout the surface coat extraction process. However, this did result in collection of a library of images for *in situ* hybridisation staining. Stable transformation of many parasitic nematode species is currently not viable meaning fluorescently fused proteins cannot be recombinantly

produced and tracked within the nematode. Therefore, localisation of gene expression through use of *in situ* hybridisation has become an indispensable technique. It is hoped that this library of images could be used to aid identification of gene localisation in future work.

Conditions for reliably inducing plant defences against cell wall break down products were developed. This offered the ability to test novel host-parasite interactions through investigation of host mediated carbohydrate sequestering. Hiding host DAMPs through binding cell wall break down products is a novel host-pathogen interaction supported by the identification of PCN surface associated lectins. Testing the response of oligogalacturonide induced ROS assays should also be investigated in the presence of additionally identified lectins that lacked a signal peptide (GROS_g00815 and GROS_g01645). This could allow identification of a wider variety of surface associated, DAMP response reducing, lectins. Ultimately carbohydrate binding assays would need to be carried out to confirm that these galectins are interacting with cell wall breakdown products. These assays would also require purified recombinant proteins which were unobtainable here.

A novel protein that drastically increases the host ROS response was also identified in nematode surface protein extractions. Although an exact function is still unknown due to the lack of detectable, known, protein domains, GROS_g02583 shared an extremely conserved sequence between all investigated species of nematodes. Modelling this protein structure against already solved crystal structures was unsuccessful due to the lack of homology between GROS_g02583 and other proteins. Clean recombinant expression of GROS_g02583 could allow the crystal structure to be determined. Once a structure has been determined it could be possible to determine how the highly conserved cysteine arrangement may play a role in the pathogen-host or pathogen-external environment interaction. Ultimately this might suggest why there is a sharp increase in expression of GROS_g02583 after invasion of the host and potentially identification of novel domains within the protein.

An upstream motif associated with expression of genes in the hypodermis was proposed. This motif was present in almost 60% of genes known to be expressed in the hypodermis compared to less than 5% of genes associated with oesophageal gland expression. Expression of genes containing the predicted motif in the nematode hypodermis needs to be tested. This can be done with *in situ* hybridisation. Genes showing multiple copies of the motif should initially be tested. However, there would also be value in localising expression of genes with just one copy of the motif. Although gene expression in the hypodermis and signal peptide presence are indicators for proteins being present in the nematode surface coat, not all proteins with these characteristics have the same fate. Ultimately all identified surface protein candidates expressed in the hypodermis would need to be localised to the surface with specific antibodies.

6.2.2 The PCN eggshell

Eggshell protein extractions suggested the presence of an eggshell associated annexin (GROS_g03104). This was later confirmed through immuno-localisation. Eggshell annexin knock-downs suggested a role for the protein associated with the eggshell permeability barrier.

Although experiments to begin characterisation of the eggshell annexin were carried out there is still much to learn about this protein. Use of the functional antibodies that were synthesised could allow exact localisation of the protein within the eggshell. Labelling with immuno-gold secondary antibodies and imaging with an electron microscope could be used to confirm protein presence in the lipoprotein permeability barrier beneath the chitin layer.

Functional testing of the apparently *Globodera* specific motif within this protein could show whether this motif is related to the PCN hatching cascade. This motif has been identified at a predicted calcium binding site in GROS_g03104. However, this now poses further questions such as, can this binding site still bind calcium or perhaps has the binding site modified to allow binding of another, larger, molecule? An exact crystal structure for GROS_g03104 should be determined. Further work investigating ligand

binding and identifying conformational changes in the protein in the presence and absence of calcium and host hatching factors will allow greater understanding of this protein's role within the eggshell.

Although chondroitin proteoglycans were identified in eggshell protein extractions and have been previously identified in eggshells of *C. elegans*, they have not yet been localised to the PCN eggshell (Olson *et al.*, 2012). Antibodies could be synthesised to localise these proteins. Previous work on eggshell chondroitin proteoglycans has shown that altering these proteins disrupts the hierarchical deployment of the eggshell, disrupting juvenile formation. Therefore knock downs of this protein to identify protein function are unlikely to be useful.

Large scale expression of proteins identified from eggshell protein extractions could be combined with techniques such as thermal shift to identify protein responses to host root diffusate. The benefits of this could be two-fold. Firstly, identifying proteins directly responding to root diffusates would aid identification of active components within the hatching factors. Secondly, identifying these protein responses would give a greater insight into whether hatching does mimic a ligand-receptor response.

With increased time, refining of testing for recombinantly expressed proteins could be carried out. For example, thermal shift assays to detect the response of the eggshell annexin GROS_g03104 to calcium were attempted but results were inconclusive due to the high hydrophobicity of this protein. Optimisation of reaction buffers and protein concentrations could allow more usable results to be obtained. Attempts at nematode protein expression using *E. coli* were often unsuccessful due to the high number of rare codons in *G. rostochiensis* proteins. With more time, expression of candidate proteins could be further investigated trialing different expression in different organisms or codon optimisation. Other proteins identified in eggshell lipid extractions such as lipase (GROS_g11707) or GH30 family protein (GROS_g13621) could be tested for their ability to bind or degrade PCN eggshells.

The methods used here could also have implications for other parasitic nematode species. Investigating eggshells of *H. glycines* might highlight proteins reflecting zinc dependency in this species. Moving away from cyst nematodes, identification of animal-parasitic nematode eggshell proteins could allow development of eggshell targeting nematicides or hatching related inhibitors. Here, there is no cyst wall surrounding the eggs meaning chemical inhibition may have greater effect.

Lipid extractions from PCN eggshells identified the presence of tri- and di- acylglycerides (DAGs and TAGs) and various phosphatidic acids, supporting previous work (Gibson *et al.*, 1995). Fatty acid methyl ester analysis showed a large difference between fatty acids from juvenile PCN compared to PCN eggshells. Higher occurrences of saturated and long chained fatty acids reflects the impermeable properties of the eggshell. Although di- and tri- acylglycerides were present, it is unlikely that these neutral lipids are found within lipid layers due to the lack of polarity needed for forming a strong bilayer. Instead, DAGs and TAGs are more likely to be present inside the perivitelline fluid and used by the non-parasitising, developing embryo as a source of fatty acids.

Strong binding between the eggshell annexin and PI species, especially PI_3P, was seen. Although the PI_3P head group was possibly identified from eggshell lipid extractions, the required fatty acids were not found.

6.2.2.1 Ascarosides in PCN

Having been identified in *Ascaris* eggshells, the potential properties of these glycolipids suggested that the ascarosides were likely to also be found in PCN eggshells. However, no ascarosides were detected in PCN eggshell lipid extractions. Similarly, no ascarylose sugar was detected in *G. rostochiensis* cysts, juveniles, eggshell lipid extractions or liquid culture used to hatch juveniles.

Currently four proteins are known to be associated with stepwise chain shortening of long fatty acids for ascaroside production, ACOX-1, MAOC-1, DHS-28 and DAF-22 (von Reuss *et al.*, 2012). These shortened fatty acids must then be combined with an ascarylose sugar to form the ascaroside. However, no ascarosyl-transferase has been

identified in *C. elegans* (Witting *et al.*, 2018). Until an ascarosyl-transferase has been identified, no PCN orthologue can be found. It is possible that in PCN this protein has been lost or mutated to favour a different hexose/sugar reflecting the lack of identifiable ascarylose in *G. rostochiensis*. It is also possible that this transferase does not exist. Instead a glucosyl-transferase could combine shortened fatty acids and this glycolipid is subsequently modified into an ascaroside by another unidentified sugar modifying protein.

ACOX-1 is responsible for the first step of ascaroside side chain production where it dehydrogenates the α-β position, producing unsaturation in the side chains. *C. elegans* mutants in ACOX-1 are no longer able to produce ascarosides with side chains less than 9 carbons (von Reuss *et al.*, 2012). In PCN, the closest match to *C. elegans* ACOX-1 is GROS_g11667 which shares just 33.74% sequence homology (**Table 6.1**). The mostly divergent sequence found in *G. rostochiensis* could explain the lack of identifiable ascarosides in PCN samples.

Table 6.1 – Homology scores between proteins associated with ascaroside side chain formation in *C. elegans* and *G. rostochiensis*

Orthologous genes identified through BLASTP. Homology scores were calculated by Jalview after protein alignment by the MUSCLE server using default settings.

	***C. elegans* ascaroside side chain producing proteins**			
Orthologous PPN gene	**ACOX-1**	**MAOC-1**	**DHS-28**	**DAF-22**
GROS_g11667	33.74			
GROS_g01444		47.22		
GROS_g12109			62.07	
GROS_g02700				68.69

Life-stage expression of ACOX-1 orthologues in both *G. pallida* and *G. rostochiensis* show upregulation of the gene during early stages of host parasitism. However, no upregulation is seen in cysts which could represent a need for the protein to produce ascarosides during egg formation or mate attraction. This protein is expressed at relatively low levels throughout the life cycle suggesting the protein is not in high demand (Figure 6.1) (Eves-van den Akker *et al.*, 2016).

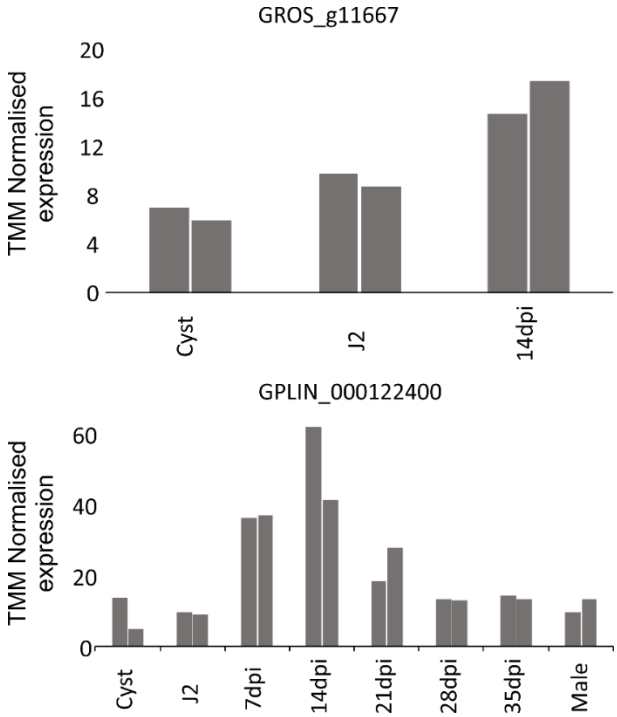


Figure 6.1 – Life stage expression of ACOX-1 orthologues in *G. rostochiensis* and *G. pallida*.
For both PCN species, ACOX-1 expression is highest at early parasitic stages and not at stages relating to mating or egg production

6.2.3 A new PCN eggshell schematic

Structural information was gathered about the *G. rostochiensis* eggshell by Perry *et al.* (1982) and a general tylenchid eggshell schematic was later created by Perry & Trett (1986). Until now, this has represented the majority of our information about the *G. rostochiensis* eggshell. Using the information gathered in this thesis a new schematic for the PCN eggshell can be created.

The first change would be the presence of the CpG layer in place of the outer lipid layer. This reflects what is seen in *C. elegans* eggshells.

Inner lipid layers can also now be clarified. Neutral lipids such as TAGs are present in eggshell lipid extractions. However, it is unlikely that these lipids are present in the lipid layer to this quantity due to the lack of amphipathic properties of TAGs. PCN eggshell lipid barriers likely comprise of phosphatidic acid species with low degrees of unsaturation. Unlike *Ascaris spp.*, ascarosides are not part of the PCN eggshell, although the presence of alternative fatty-acyl glycosides was suggested from the lipid mass spectrometry data. Suggested annexin lipid binding properties also highlights the possible presence of phosphatidylinositol species in the lipid layer.

An exact place for the eggshell annexin within the schematic is difficult to determine. The annexin shows no transmembrane domains, so it is unlikely to be found within a lipid layer. However, the lipid binding function of the annexin suggests it would still be associated with a lipid layer. Previous work suggests that eggshell calcium binding sites are found beneath the chitin layer (Clarke & Perry, 1985). This would locate the annexin to the surface of the inner lipid layers. It is unclear whether annexins are associated with all inner lipid layers or restricted to just one layer.

A new schematic for the PCN eggshell has been made (**Figure 6.2**). It is hoped that this schematic can be built upon with PCN eggshell research in the future.

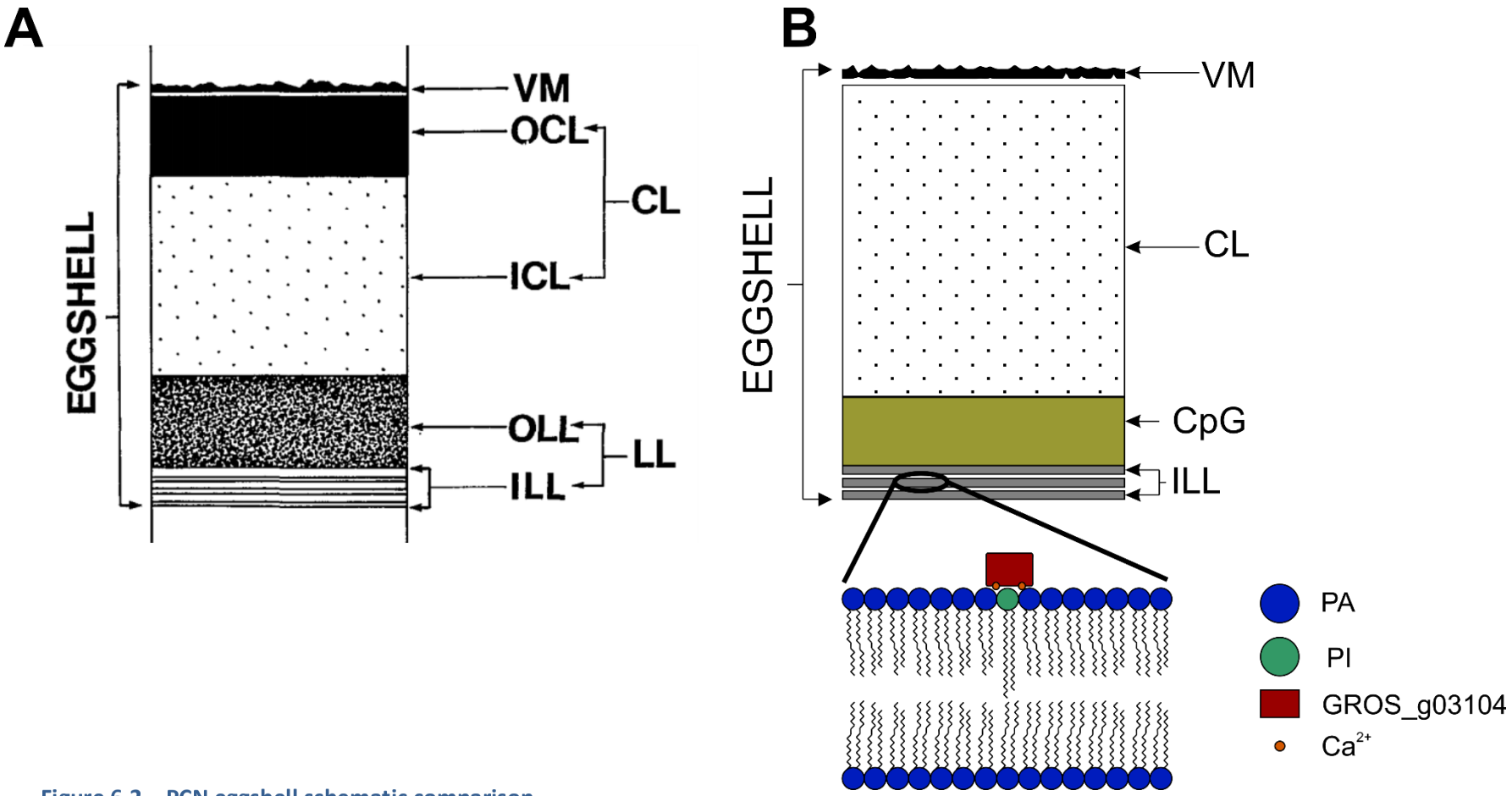


Figure 6.2 – PCN eggshell schematic comparison

A) eggshell schematic taken from Perry and Trett 1986. *G. rostochiensis* eggshells were proposed to consist of three layers. An outer vitelline membrane, a single chitin layer and an inner lipid membrane (**Perry 1982**). B) A new eggshell schematic proposed after research in this thesis showing the addition of a CpG layer and a more defined inner lipid layer. **VM** – vitelline membrane, **OCL** – outer chitin layer, **ICL** – inner chitin layer, **CL** – chitin layer, **OLL** – outer lipid layer, **ILL** – inner lipid layer, **CpG** – chondroitin proteoglycan layer, **PA** – phosphatidic acid, **PI** – phosphatidyl inositol, **GROS_g03104** – *G. rostochiensis* eggshell annexin.

6.3 References

CLARKE, A.J. & PERRY, R.N. 1985. Egg-shell calcium and the hatching of Globodera rostochiensis. *International Journal for Parasitology*, **15**, 511–516.

DAVIES, K.G. & CURTIS, R.H.C. 2011. Cuticle Surface Coat of Plant-Parasitic Nematodes. *Annual Review of Phytopathology*, **49**, 135–156.

EVES-VAN DEN AKKER, S., LAETSCH, D.R., THORPE, P., LILLEY, C.J., DANCHIN, E.G.J., DA ROCHA, M., RANCUREL, C., … JONES, J.T. 2016. The genome of the yellow potato cyst nematode, Globodera rostochiensis, reveals insights into the basis of parasitism and virulence. *Genome Biology*, **17**, 124.

GIBSON, D.M., MOREAU, R.A., MCNEIL, G.P. & BRODIE, B.B. 1995. Lipid Composition of Cyst Stages of Globodera rostochiensis. *Journal of nematology*, **27**, 304–311.

JONES, J.T., REAVY, B., SMANT, G. & PRIOR, A.E. 2004. Glutathione peroxidases of the potato cyst nematode Globodera Rostochiensis. *Gene*, **324**, 47–54.

OLSON, S.K., GREENAN, G., DESAI, A., MÜLLER-REICHERT, T. & OEGEMA, K. 2012. Hierarchical assembly of the eggshell and permeability barrier in C. Elegans. *Journal of Cell Biology*, **198**, 731–748.

PERRY, R.N. 2002. Hatching. *In* Lee, D.L., *ed. The Biology of Nematodes*. Taylor & Francis, 147–165.

PERRY, R.N. & TRETT, M.W. 1986. Ultrastructure of the eggshell of Heterodera schachtii and H. glycines (Nematoda: Tylenchida). *Revue de nematologie*.

PERRY, R.N., WHARTON, D.A. & CLARKE, A.J. 1982. The structure of the egg-shell of Globodera rostochiensis (Nematoda: Tylenchida). *International Journal for Parasitology*, **12**, 481–485.

PRIOR, A., JONES, J.T., BLOK, V.C., BEAUCHAMP, J., MCDERMOTT, L., COOPER, A. & KENNEDY, M.W. 2001. A surface-associated retinol- and fatty acid-binding protein (Gp-FAR-1) from the potato cyst nematode Globodera pallida: lipid binding activities, structural analysis and expression pattern. *The Biochemical journal*, **356**, 387–394.

SHINYA, R., MORISAKA, H., KIKUCHI, T., TAKEUCHI, Y., UEDA, M. & FUTAI, K. 2013. Secretome Analysis of the Pine Wood Nematode Bursaphelenchus xylophilus Reveals the

Tangled Roots of Parasitism and Its Potential for Molecular Mimicry. *PLoS ONE*, **8**.

VON REUSS, S.H., BOSE, N., SRINIVASAN, J., YIM, J.J., JUDKINS, J.C., STERNBERG, P.W. & SCHROEDER, F.C. 2012. Comparative Metabolomics Reveals Biogenesis of Ascarosides, a Modular Library of Small-Molecule Signals in *C. elegans*. *Journal of the American Chemical Society*, **134**, 1817–1824.

WITTING, M., HASTINGS, J., RODRIGUEZ, N., JOSHI, C.J., HATTWELL, J.P.N., EBERT, P.R., VAN WEEGHEL, M., … CASANUEVA, O. 2018. Modeling Meets Metabolomics—The WormJam Consensus Model as Basis for Metabolic Studies in the Model Organism Caenorhabditis elegans. *Frontiers in Molecular Biosciences*, **5**, 96.

7. Supplementary figures

7.1 Primer table, use and T_m

Synthesised Primers For Gene Amplification				
Primer Name	**Sequence (5′➔ 3′)**	**Application**	**T_m (˚C)**	**Additional Notes**
g02490ISHF	TGGATGAGACGACGAGATGG	*In situ* hybridisation	59.4	
g02490ISHR	TAAAGGCACGTTGGTCAAGG	*In situ* hybridisation	57.3	
g01153ISHF	CAAGCTGATCCCGTACAAGC	*In situ* hybridisation	59.4	
g01153ISHR	AAAGCTGGACTGGTAACACG	*In situ* hybridisation	57.3	
g10494ISHF	GGGCAAGGTGATGAAGCAAA	*In situ* hybridisation	57.3	
g10494ISHR	TTCTTGTTTATGCCGCTCCG	*In situ* hybridisation	57.3	
g07793ISHF	ATTTGCCAATGTGCTGAGGG	*In situ* hybridisation	57.3	
g07793ISHR	TGTTGAAGTTGACAGGTGCG	*In situ* hybridisation	57.3	
g06693ISHF	ATTTCCCGAGCGTGTCAATG	*In situ* hybridisation	57.3	
g06693ISHR	GCCAGTTAGTCGTCTTGGTT	*In situ* hybridisation	57.3	
g01391ISHF	GATGAGGCGGAAGAGGAAGA	*In situ* hybridisation	59.4	
g01391ISHR	AAGGCACGAGAAGAGAAGCA	*In situ* hybridisation	57.3	
g02583ISHF	TGATGTGAACAAGTTGGCCG	*In situ* hybridisation	60.0	
g02583ISHR	TACTTCCTTAACCACAGCCG	*In situ* hybridisation	60.0	
g05049ISHF	TCTGAACAAGATGCCCTCGT	*In situ* hybridisation	60.0	
g05049ISHR	CAACCACTGCCCCTTCTCAT	*In situ* hybridisation	62.0	
g01970ISHF	GCCCACCAATTATCCGAACC	*In situ* hybridisation	62.0	
g01970ISHR	AATCGGCCGTTTTCACACAG	*In situ* hybridisation	60.0	
g01882ISHF	CAATTCCAGGCCGAAGTCAG	*In situ* hybridisation	62.0	
g01882ISHR	CCGGCTGTTCCTTTTAGTGT	*In situ* hybridisation	60.0	
g12305ISHF	TGAACAGCCACAGACATTGC	*In situ* hybridisation	60.0	
g12305ISHR	CGTTATGGGCTTCGAACTCC	*In situ* hybridisation	62.0	
g02910ISHF	CTCCGCCAACAACTTTCCAA	*In situ* hybridisation	60.0	
g02910ISHR	CACGCCTTTCACGGTTCATT	*In situ* hybridisation	60.0	
g13102ISHF	GTTCCCAACGCTTTGTGACT	*In situ* hybridisation	60.0	

g13102ISHR	ACGCGTGTAACTGTGTTTCT	*In situ* hybridisation	58.0	
g09658ISHF	CGCGGATTGGAAAATGAGGT	*In situ* hybridisation	60.0	
g09658ISHR	TGCCTCTCAACACTCCTCAC	*In situ* hybridisation	62.0	
g03960ISHF	ATGGACAACGCCTTCCAGTA	*In situ* hybridisation	60.0	
g03960ISHR	CTCCGACTTCTAGCGACAAC	*In situ* hybridisation	62.0	
g06000ISHF	GCAAAACGTTCACTTGTGCC	*In situ* hybridisation	60.0	
g06000ISHR	ACCACGGTTTAGTTTCGCTT	*In situ* hybridisation	58.0	
g14202ISHF	CACCGGAGCCAGTTTTATCG	*In situ* hybridisation	59.4	
g14202ISHR	CGCTTCCATTGCCTGTTCTT	*In situ* hybridisation	57.3	
g01153TOPOF	ACCATGGGTGAAGAACCTGAGGAGAAGCTG	TOPO Cloning	59.3	For cloning whole gene, addition of Kozak sequence
g01153TOPOR+S	CATTGGTAGTTGCTGCGAGCTACT	TOPO Cloning	52.3	For cloning whole gene, includes stop codon
g01153TOPOR-S	GCATTGGTAGTTGCTGCGAGCT	TOPO Cloning	51.6	For cloning whole gene, no stop codon
g02583TOPOF	ACCATGCAGACCGTGAAGCAGTGCTTGTGCTC	TOPO Cloning	61.9	For cloning whole gene, addition of Kozak sequence
g02583TOPOR+S	GACACACGCGTTTCCAGTTTTAGTGGATT	TOPO Cloning	55.0	For cloning whole gene, includes stop codon
g02583TOPOR-S	GACACACGCGTTTCCAGTTTTAGTGG	TOPO Cloning	54.4	For cloning whole gene, no stop codon
g06693TOPOF	ACCATGTGTCAATGCCAATGGGCTCCACC	TOPO Cloning	59.3	For cloning whole gene, addition of Kozak sequence
g06693TOPOR+S	CGTCCTTAAACTGGTAGTTACGTCGACT	TOPO Cloning	54.8	For cloning whole gene, includes stop codon
g06693TOPOR-S	CTTCGTCCTTAAACTGGTAGTTACGTCG	TOPO Cloning	54.8	For cloning whole gene, no stop codon
g01153Nde1F	ATACATATGGGTGAAGAACCTGAGGAGAAGCTG	pET-15b Cloning	58.1	For cloning whole gene, addition of Nde1 restriction site
g01153BamH1R	ATAGGATCCTCATCGAGCGTCGTTGATGGTTAC	pET-15b Cloning	59.3	For cloning whole gene, addition of BamH1 restriction site

g02583Nde1F	ATACATATGCAGACCGTGAAGCAGTGCTTGTGCTC	pET-15b Cloning	60.5	For cloning whole gene, addition of Nde1 restriction site
g02583BamH1R	CACGGATCCTTAGGTGATTTTGACCTTTGCGCAC	pET-15b Cloning	60.6	For cloning whole gene, addition of BamH1 restriction site
g06693Nde1F	ATACATATGTGTCAATGCCAATGGGCTCCACC	pET-15b Cloning	58.0	For cloning whole gene, addition of Nde1 restriction site
g06693BamH1R	ATAGGATCCTCAGCTGCATTGATGGTCA	pET-15b Cloning	54.8	For cloning whole gene, addition of BamH1 restriction site
g03104Nde1F	ATACATATGATGTCAAACGCCGCC	pET-15b Cloning	50.6	For cloning whole gene, addition of Nde1 restriction site
g03104BamH1R	ATAGGATCCTCAATTTCCTTTGATCAGCG	pET-15b Cloning	53.6	For cloning whole gene, addition of BamH1 restriction site
g11707Nde1F	ATACATATGACCATTTCCGAGCCGTTCGC	pET-15b Cloning	56.4	For cloning whole gene, addition of Nde1 restriction site
g11707BamH1R	ATAGGATCCTCAGGCCGATAACAGTTTCAG	pET-15b Cloning	56.5	For cloning whole gene, addition of BamH1 restriction site
g13621Nde1F	CGCCATATGGAATGGAAGCCGTGCAAAAAGC	pET-15b Cloning	59.3	For cloning whole gene, addition of Nde1 restriction site
g13621BamH1	GCGGGATCCTCATTCTTTCAACGCCTTTTTTGGC	pET-15b Cloning	60.6	For cloning whole gene, addition of BamH1 restriction site

g03104_BspH1_F	AAAATCATGATGTCAAACGCCGCCA	pEHISTEV cloning	62.0	For cloning whole gene, addition of BspH1 restriction site
g11707_BspH1_F	ATAATCATGACCATTTCCGAGCCGTTCGC	pEHISTEV cloning	65.0	For cloning whole gene, addition of BspH1 restriction site
g13621_BspH1_F	CCCATCATGAAAATGGAATGGAAGCCGTGCAAA	pEHISTEV cloning	66.0	For cloning whole gene, addition of BspH1 restriction site
g06693_BspH1_F	ACAATCATGAAAATGTGTCAATGCCAATGGGC	pEHISTEV cloning	63.0	For cloning whole gene, addition of BspH1 restriction site
g02583_BspH1_F	ACAATCATGAAAATGCAGACCGTGAAGCAGTGCT	pEHISTEV cloning	67.0	For cloning whole gene, addition of BspH1 restriction site
T7-F	AATACGACTCACTATAGGG	Colony PCR	47.6	pET-15b vector
T7-R	CTAGTTATTGCTCAGCGG	Colony PCR	50.2	pET-15b vector
M13F	GTAAAACGACGGCCAG	Colony PCR	53.0	pGEM-Teasy, TOPO vectors
M13R	CAGGAAACAGCTATGAC	Colony PCR	49.0	pGEM-Teasy, TOPO vectors
SPE1-SigPep-For	CGGTACTAGTATGAAGACTAATCTTTTTC	Gateway Cloning	55.0	Creation of RFP gateway vector with apoplast targeting signal peptide
SPE1-SigPep-Rev	CGGTACTAGTGGCCGAGGAT	Gateway Cloning	60.0	Creation of RFP gateway vector with apoplast targeting signal peptide
P35s-For	AAGGAAGTTCATTTCATTTGGAGAGGA	Gateway Colony PCR	58.0	
T35s-Rev	CAACACATGAGCGAAACCCTATAAGAA	Gateway Colony PCR	59.0	

Xho1_3104RNAi_F	ACACTCGAGGCCGCCAAAAACATTCTCGG	RNAi	59.3	Creation of GROS_g03104 RNAi vector in pHannibal/pART-27
Kpn1_3104RNAi_R	ACAGGTACCAATCAATTCCTCTAAATCGCCGG	RNAi	58.0	Creation of GROS_g03104 RNAi vector in pHannibal/pART-27
Xba1_3104RNAi_F	ACATCTAGAGCCGCCAAAAACATTCTCGG	RNAi	56.4	Creation of GROS_g03104 RNAi vector in pHannibal/pART-27
HindIII_3104RNAi_R	ACAAAGCTTAATCAATTCCTCTAAATCGCCGG	RNAi	55.5	Creation of GROS_g03104 RNAi vector in pHannibal/pART-27
FRAG_1_R	TCGTCTTACACATCACTTGTCA	RNAi	55.0	Colony PCR to check insertion into pHannibal
FRAG_2_F	AGTCGAACATGAATAAACAAGGTAACA	RNAi	56.0	Colony PCR to check insertion into pHannibal
Kana_ntpII_F	ATGACTGGGCACAACAGACAATCGGCTGCT	Transgenics	68.8	Checking for kanamycin resistance in GROS_g03104 transgenic line
Kana_ntpII_R	CCGCTATCAGGACATAGCGTTGGCTACCCG	Transgenics	69.7	Checking for kanamycin resistance in GROS_g03104 transgenic line
stEF1alpha_F	CCAGAAGAAGGGAAAGTGAA	Semi-qPCR	56.9	SQRTPCR control
stEF1alpha_R	CAACAAAAGCAAAAGAAAACAG	Semi-qPCR	56.9	SQRTPCR control
GAPDH_For	GTGATTAGCAACGCTTCGTG	RNAi qPCR control	57.3	Nematode gene for qPCR control
GAPDH_Rev	GTCATCAGCCCTTCGATGAT	RNAi qPCR control	57.3	Nematode gene for qPCR control
3104RNAi_F	TTTTTCGGCATTGGCAATCT	RNAi qPCR	53.2	qPCR of GROS_g03104 in knock down nematodes
3104RNAi_R	TGTACTGGTCTTTGATCTCG	RNAi qPCR	55.3	qPCR of GROS_g03104 in knock down nematodes

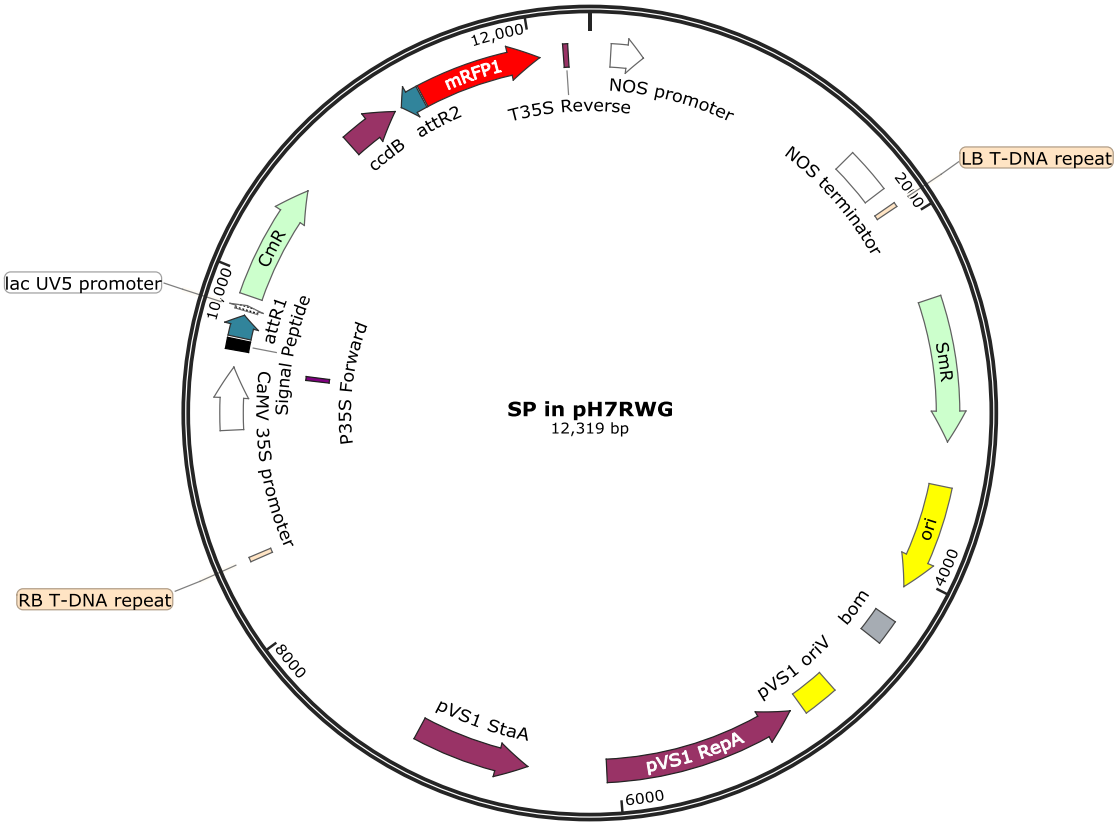
7.2 ApSiPR vector

ApSiP-RWG binary Gateway destination vector

*p35S::**ApoplastSigPep**::GW::mRFP::t35S* for RFP fusion

Gateway vector with spectinomycin resistance. ApSiP-RWG fuses an N- teminal apoplastic targeting signal peptide to the protein of interest with further C- terminal fusion to RFP for imaging or purification.

Reference for pH7RWG2 backbone and for apoplast targeting signal peptide from SP in pK7FWG2: Karimi *et al.* 2002 Trends in Plant Science May 7(5): 193-195 "Gateway vectors for *Agrobacterium*-mediated plant transformation"



Sequencing over signal peptide

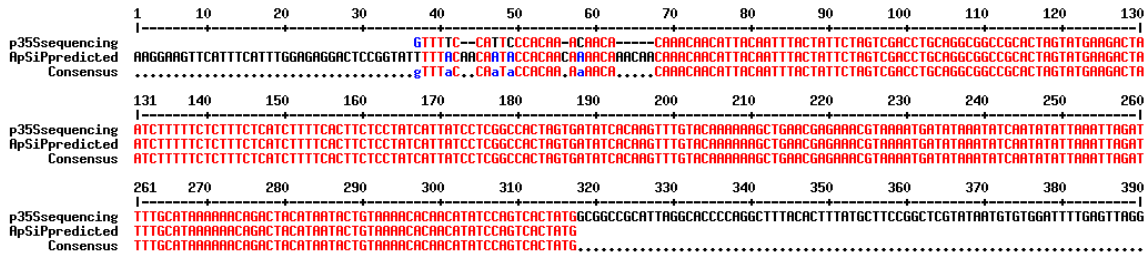
>ApSiPpredicted (p35s->attR1 end)

AAGGAAGTTCATTTCATTTGGAGAGGACTCCGGTATTTTTACAACAATACCACAACAAAACAAA
CAACAAACAACATTACAATTTACTATTCTAGTCGACCTGCAGGCGGCCGCACTAGTATGAAGAC
TAATCTTTTTTCTCTTTCTCATCTTTTCACTTCTCCTATCATTATCCTCGGCCACTAGTGATATC
ACAAGTTTGTACAAAAAAGCTGAACGAGAAACGTAAAATGATATAAATATCAATATATTAAATT
AGATTTTGCATAAAAAACAGACTACATAATACTGTAAAACACAACATATCCAGTCACTATG

>p35Ssequencing

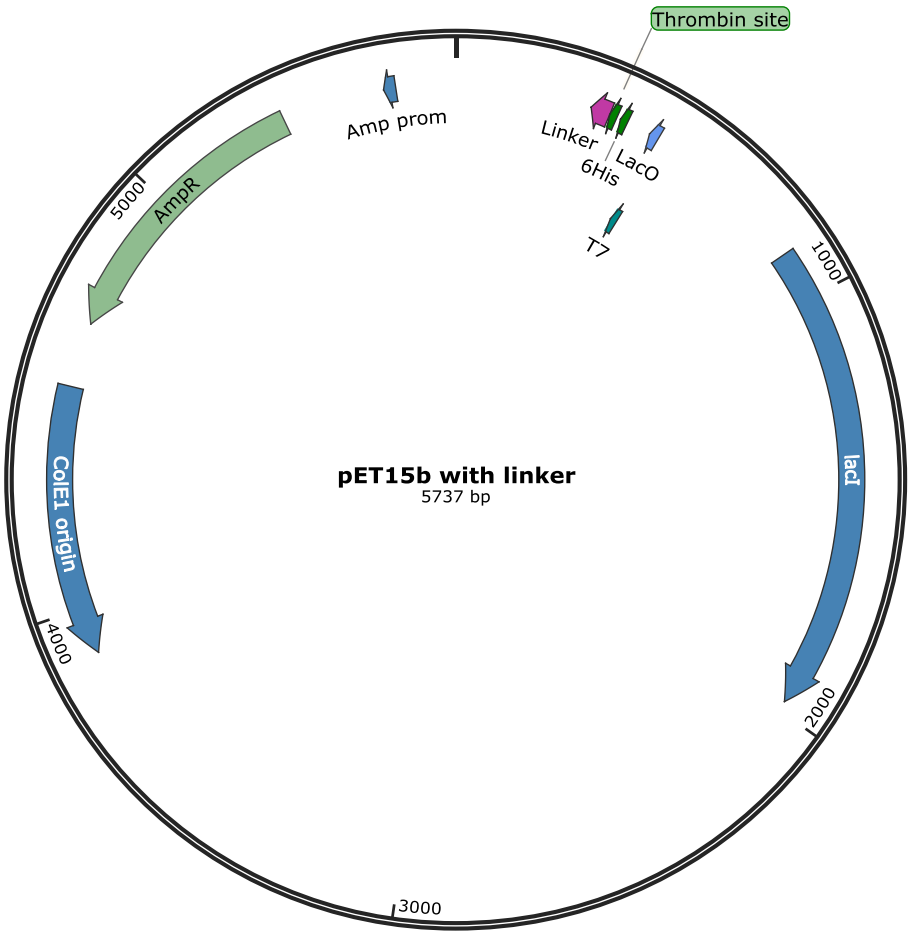
GTTTTCCATTCCCACAAACAACACAAACAACATTACAATTTACTATTCTAGTCGACCTGCAGGC
GGCCGCACTAGTATGAAGACTAATCTTTTTTCTCTTTCTCATCTTTTCACTTCTCCTATCATTAT
CCTCGGCCACTAGTGATATCACAAGTTTGTACAAAAAAGCTGAACGAGAAACGTAAAATGATAT
AAATATCAATATATTAAATTAGATTTTGCATAAAAAACAGACTACATAATACTGTAAAACACAA
CATATCCAGTCACTATGGCGGCCGCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTC
GTATAATGTGTGGATTTTGAGTTAGGATCCGTCGAGATTTTCAGGAGCTAAGGAAGCTAAAATG
GAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTG
AGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTT
TTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGC
CTGATGAATGCTCATCCGGAATTCCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATA
GTGNTCACCCTTGTTACACCGTTTTCCATGAGCAACTGAAACGTTTTCATCGCTCTGGAGTGAA
TACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCAGGATGTGGCGTGTTACGGTGAAA
CCTGGCCTATTNCCTAAGGGTTATGAGAATATGTTTTTCGTCCAGCNATCCNGGGNGAGTTCNC
NGTTTG

Allignment showing correct sequence for signal peptide, in frame with attR1

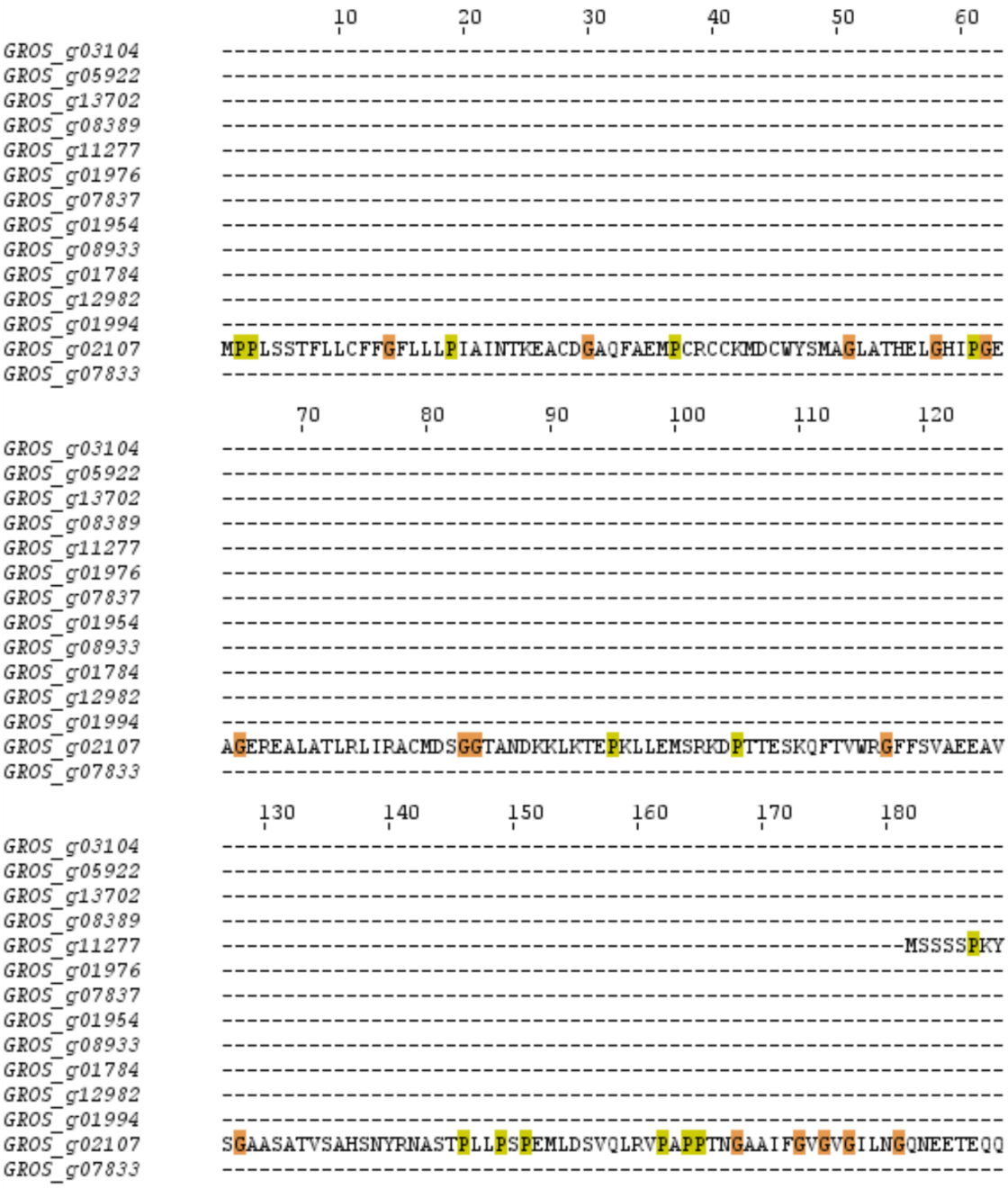


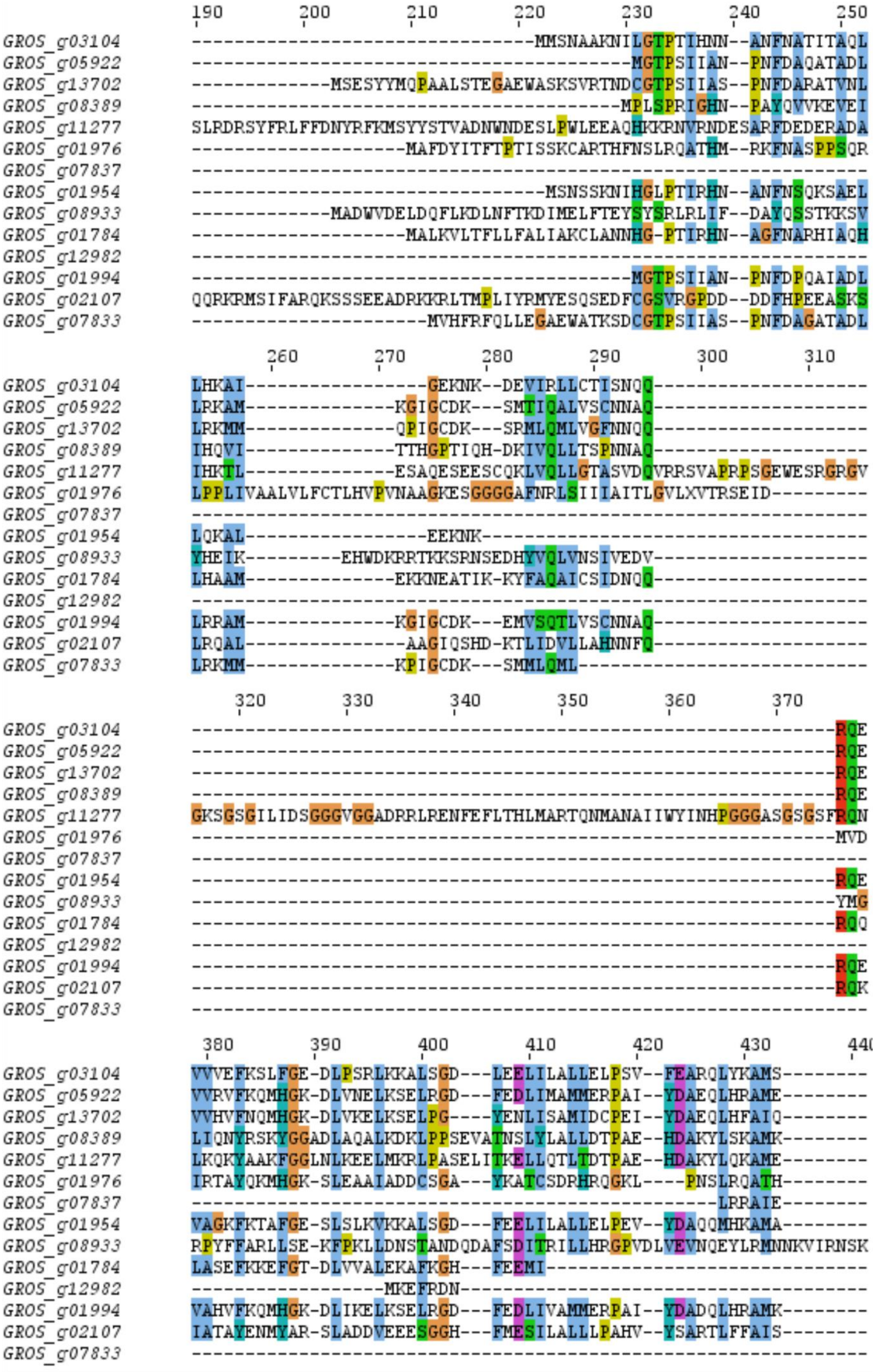
7.3 pET15b with linker vector

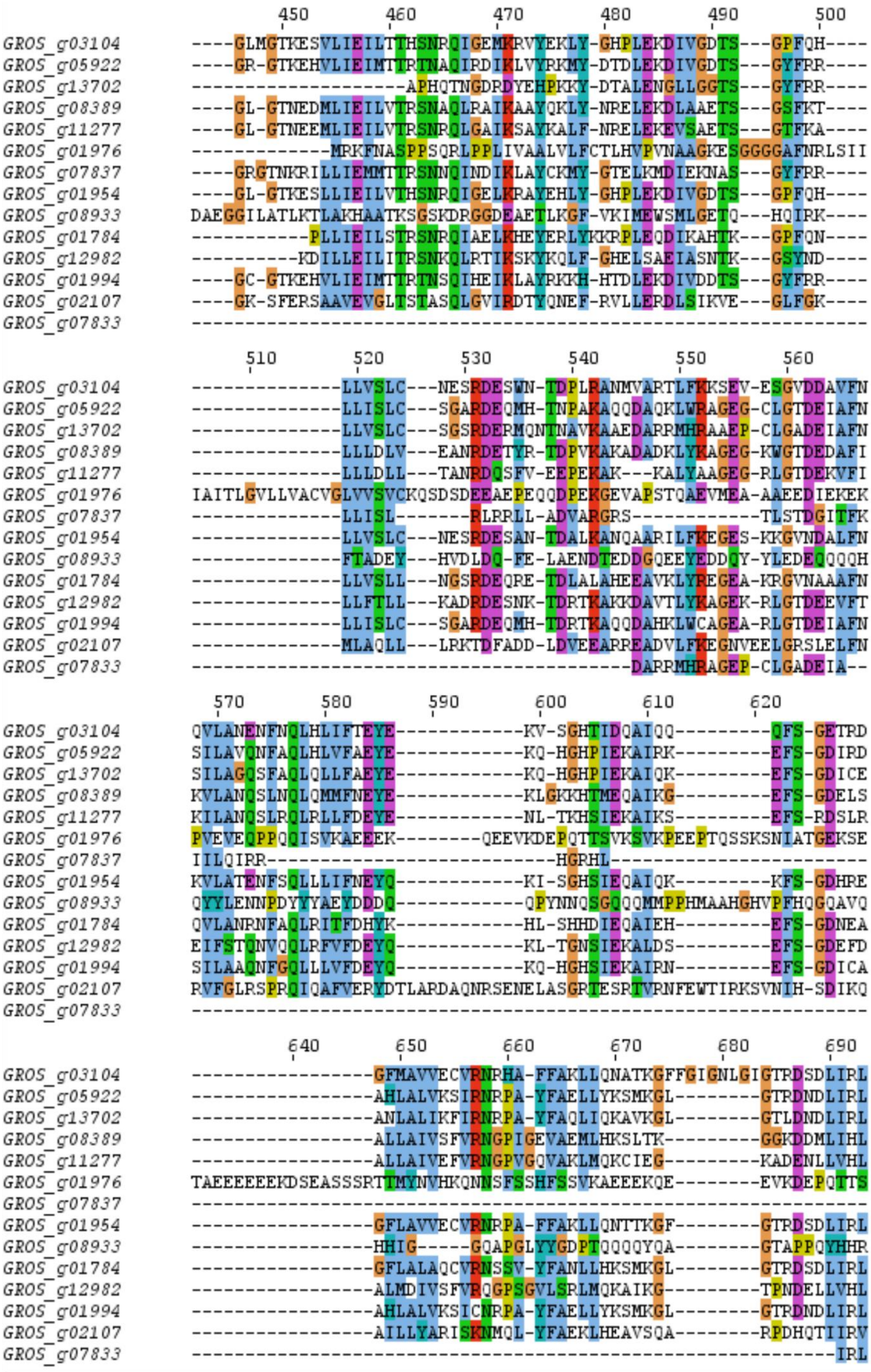
Protein expression vector based of commercial pET15b. However, an added linker region has been inserted downstream of the thrombin cleavage site to give a larger number of enzyme restriction site possibilities (Louise Major, TKS group). The N-terminus of the protein is linked to a 6xHis tag for nickel column purification which can be cleaved using thrombin if required.

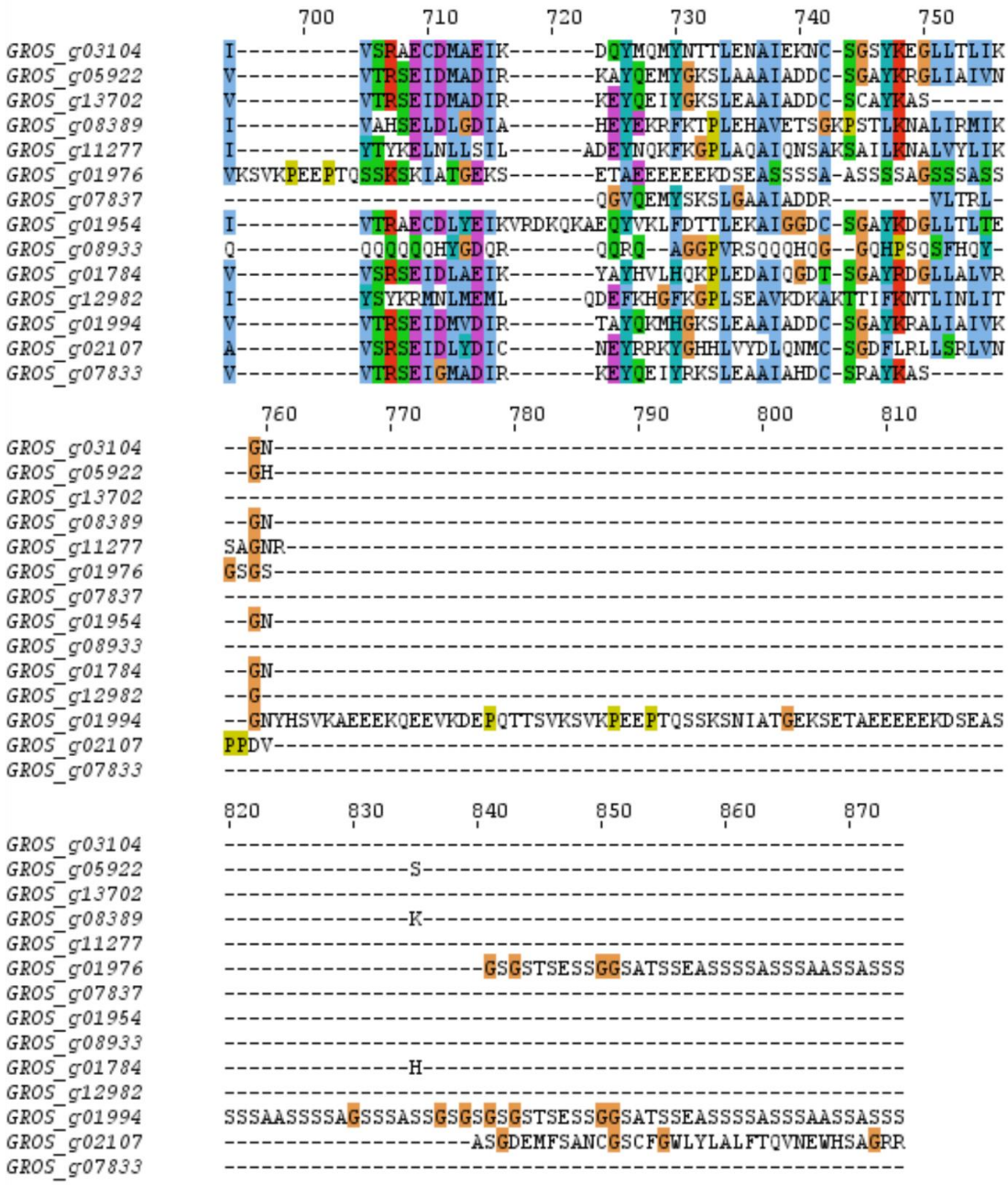


7.4 Alignment of annexins from *G. rostochiensis*

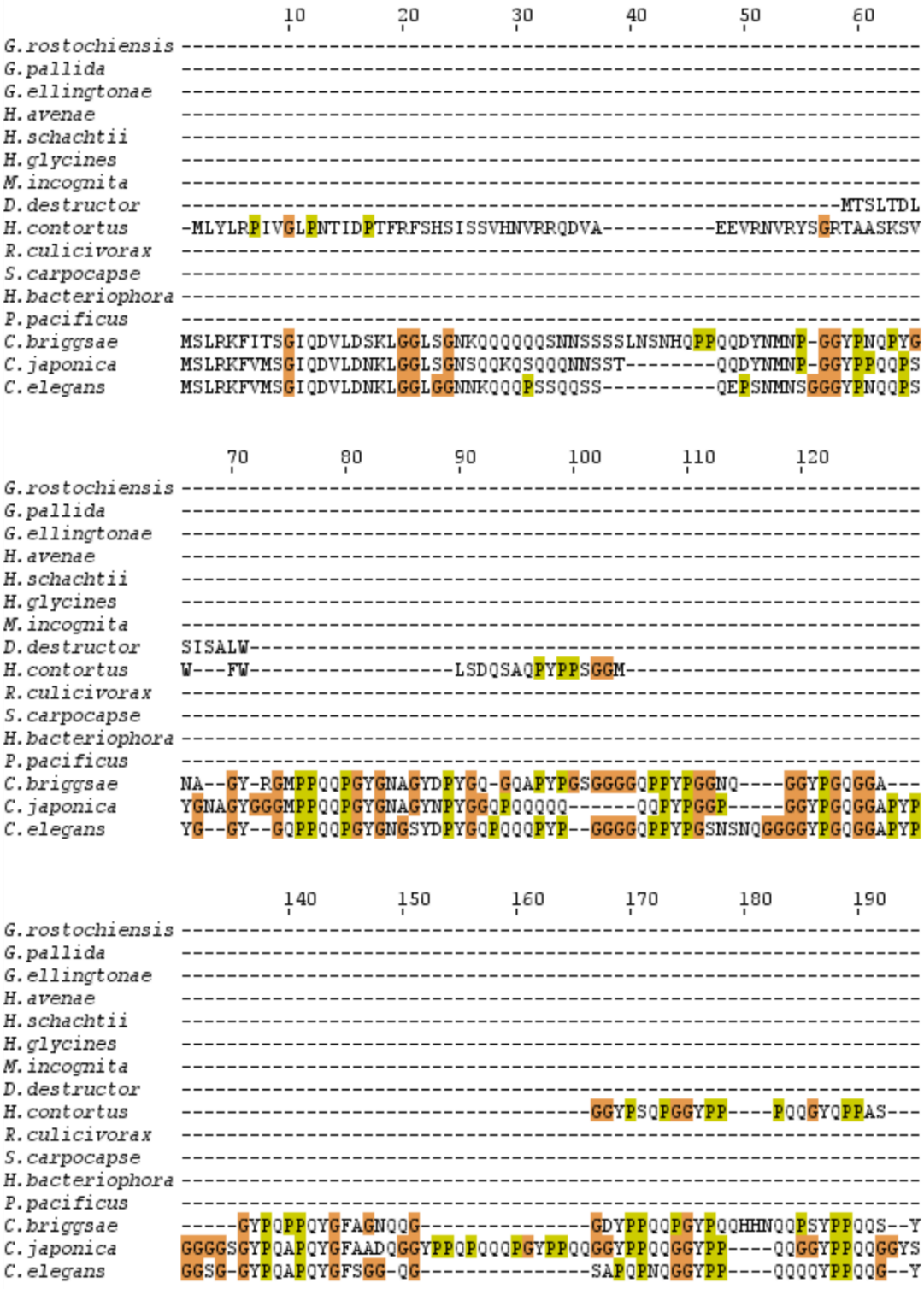


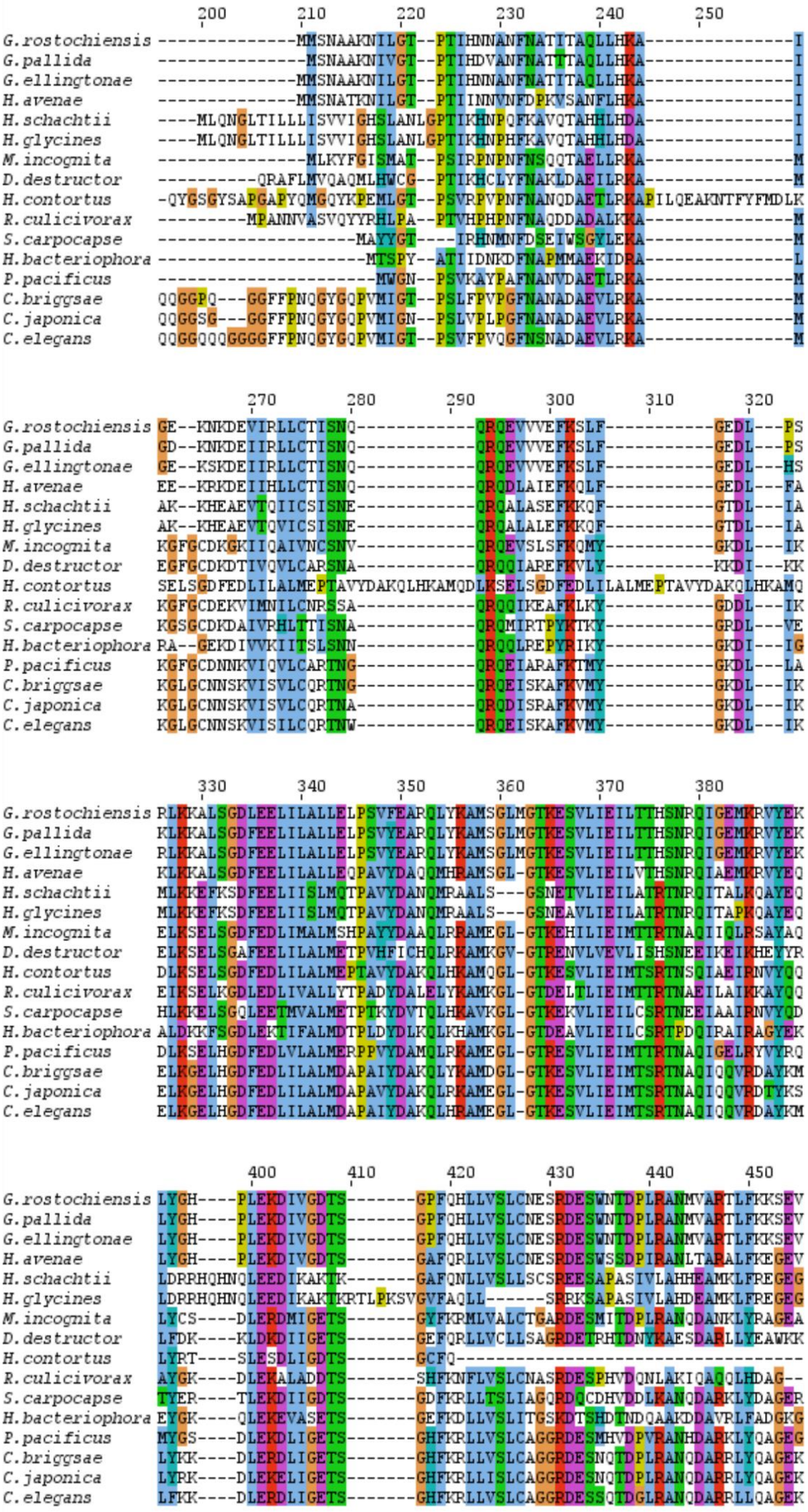

190 200 210 220 230 240 250
GROS_g03104 ----------MMSNAAKNILGTPTIHNN--ANFNATITAQL
GROS_g05922 ----------MGTPSIIAN--PNFDAQATADL
GROS_g13702 ----------MSESYYMQPAALSTEGAEWASKSVRTNDCGTPSIIAS--PNFDARATVNL
GROS_g08389 ----------MPLSPRIGHN--PAYQVVKEVEI
GROS_g11277 SLRDRSYFRLFFDNYRFKMSYYSTVADNWNDESLPWLEEAQHKKRNVRNDESARFDEDERADA
GROS_g01976 ----------MAFDYITFTPTISSKCARTHFNSLRQATHM--PKFNASPPSQR
GROS_g07837 ----------
GROS_g01954 ----------MSNSSKNIHGLPTIPHN--ANFNSQKSAEL
GROS_g08933 ----------MADWVDELDQFLKDLNFTKDIMELFTEYSYSRLRLIF--DAYQSSTKKSV
GROS_g01784 ----------MALKVLTFLLFALIAKCLANNHG-PTIPHN--AGFNAPHIAQH
GROS_g12982 ----------
GROS_g01994 ----------MGTPSIIAN--PNFDPQAIADL
GROS_g02107 QQRKRMSIFARQKSSSEEADRKKRLTMPLIYRMYESQSEDFCGSVRGPDD--DDFHPEEASKS
GROS_g07833 ----------MVHFRFQLLEGAEWATKSDCGTPSIIAS--PNFDAGATADL
260 270 280 290 300 310
GROS_g03104 LHKAI----------GEKNK--DEVIRLLCTISNQQ----------
GROS_g05922 LPKAM----------KGIGCDK---SMTIQALVSCNNAQ----------
GROS_g13702 LPKMM----------QPIGCDK---SRMLQMLVGFNNQQ----------
GROS_g08389 IHQVI----------TTHGPTIQH-DKIVQLLTSPNNAQ----------
GROS_g11277 IHKTL----------ESAQESEESCQKLVQLLGTASVDQVRRSVAPRPSGEWESRGRGV
GROS_g01976 LPPLIVAALVLFCTLHVPVNAAGKESGGGGAFNRLSIIIAITLGVLXVTRSEID---------
GROS_g07837 ----------
GROS_g01954 LQKAL----------EEKNK----------
GROS_g08933 YHEIK---------EHWDKRRTKKSRNSEDHYVQLVNSIVEDV----------
GROS_g01784 LHAAM----------EKKNEATIK-KYFAQAICSIDNQQ----------
GROS_g12982 ----------
GROS_g01994 LPRAM----------KGIGCDK---EMVSQTLVSCNNAQ----------
GROS_g02107 LRQAL----------AAGIQSHD-KTLIDVLLAHNNFQ----------
GROS_g07833 LPKMM----------KPIGCDK---SMMLQML----------
320 330 340 350 360 370
GROS_g03104 ----------RQE
GROS_g05922 ----------RQE
GROS_g13702 ----------RQE
GROS_g08389 ----------RQE
GROS_g11277 GKSGSGILIDSGGGVGGADRRLRENFEFLTHLMARTQNMANAIIWYINHPGGGASGSGSFRQN
GROS_g01976 ----------MVD
GROS_g07837 ----------
GROS_g01954 ----------RQE
GROS_g08933 ----------YMG
GROS_g01784 ----------RQQ
GROS_g12982 ----------
GROS_g01994 ----------RQE
GROS_g02107 ----------RQK
GROS_g07833 ----------
380 390 400 410 420 430 440
GROS_g03104 VVVEFKSLFGE-DLPSRLKKALSGD---LEELILALLELPSV--FEARQLYKAMS---------
GROS_g05922 VVRVFKQMHGK-DLVNELKSELRGD---FEDLIMAMMERPAI--YDAEQLHRAME---------
GROS_g13702 VVHVFNQMHGK-DLVKELKSELPG----YENLISAMIDCPEI--YDAEQLHFAIQ---------
GROS_g08389 LIQNYRSKYGGADLAQALKDKLPPSEVATNSLYLALLDTPAE--HDAKYLSKAMK---------
GROS_g11277 LKQKYAAKFGGLNLKEELMKRLPASELITKELLQTLTDTPAE--HDAKYLQKAME---------
GROS_g01976 IRTAYQKMHGK-SLEAAIADDCSGA---YKATCSDPHRQGKL----PNSLRQATH---------
GROS_g07837 ----------LPRAIE---------
GROS_g01954 VAGKFKTAFGE-SLSLKVKKALSGD---FEELILALLELPEV--YDAQQMHKAMA---------
GROS_g08933 RPYFFARLLSE-KFPKLLDNSTANDQDAFSDITRILLHRGPVDLVEVNQEYLPMNNKVIPNSK
GROS_g01784 LASEFKKEFGT-DLVVALEKAFKGH---FEEMI----------
GROS_g12982 ----------MKEFRDN----------
GROS_g01994 VAHVFKQMHGK-DLIKELKSELRGD---FEDLIVAMMERPAI--YDADQLHRAMK---------
GROS_g02107 IATAYENMYAR-SLADDVEEESGGH---FMESILALLLPAHV--YSARTLFFAIS---------
GROS_g07833 ----------

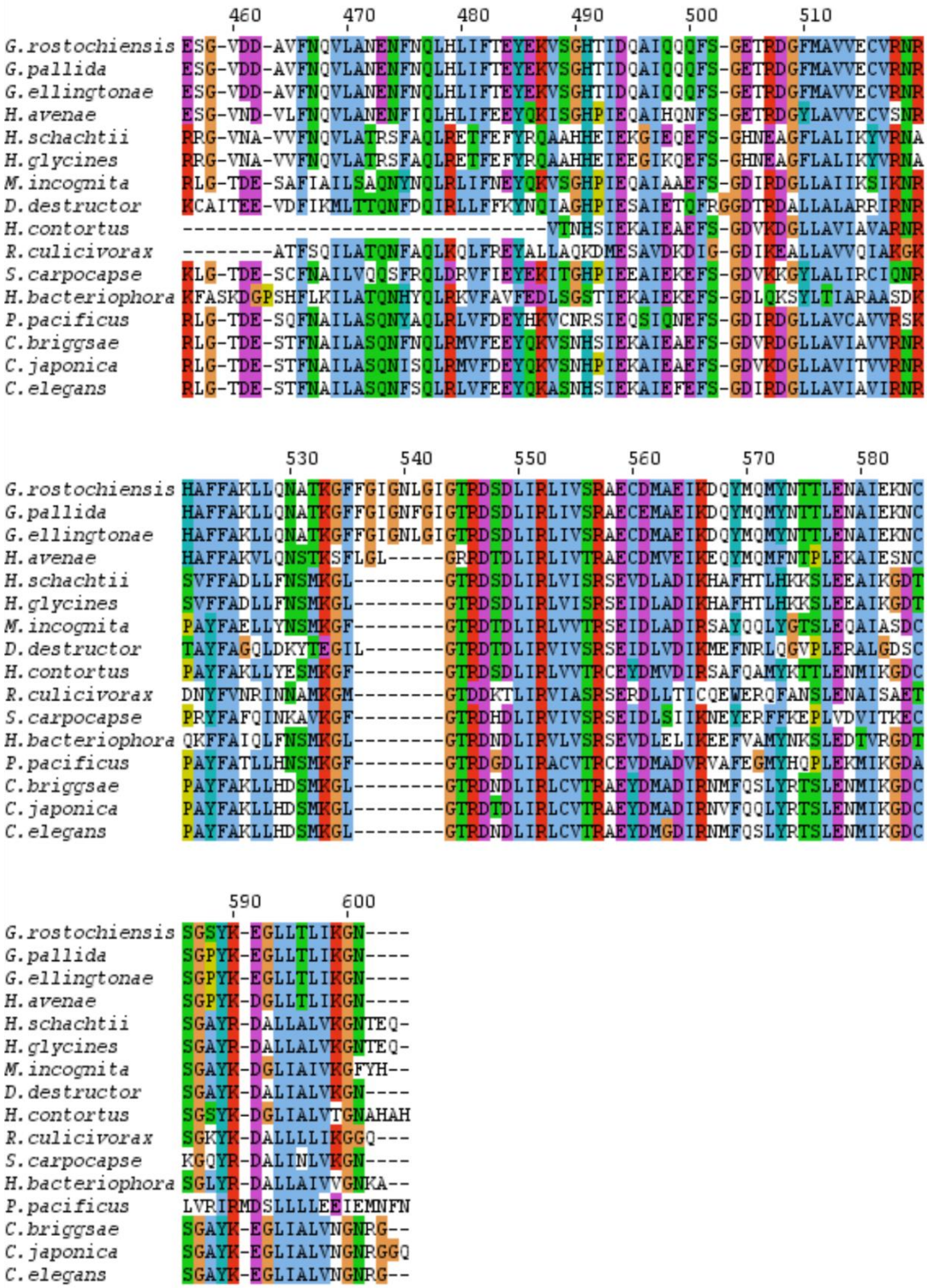

450 460 470 480 490 500
GROS_g03104 ----GLMGTKESVLIEILTTHSNRQIGEMKRVYEKLY-GHPLEKDIVGDTS---GPFQH----
GROS_g05922 ----GR-GTKEHVLIEIMTTRTNAQIRDIKLVYRKMY-DTDLEKDIVGDTS---GYFRR----
GROS_g13702 --------------------APHQTNGDRDYEHPKKY-DTALENGLLGGTS---GYFRR----
GROS_g08389 ----GL-GTNEDMLIEILVTRSNAQLRAIKAAYQKLY-NRELEKDLAAETS---GSFKT----
GROS_g11277 ----GL-GTNEEMLIEILVTRSNRQLGAIKSAYKALF-NRELEKEVSAETS---GTFKA----
GROS_g01976 -------------MRKFNASPPSQRLPPLIVAALVLFCTLHVPVNAAGKESGGGGAFNRLSII
GROS_g07837 ----GRGTNKRILLIEMMTTRSNNQINDIKLAYCKMY-GTELKMDIEKNAS---GYFRR----
GROS_g01954 ----GL-GTKESLLIEILVTHSNRQIGELKRAYEHLY-GHPLEKDIVGDTS---GPFQH----
GROS_g08933 DAEGGILATLKTLAKHAATKSGSKDRGGDEAETLKGF-VKIMEWSMLGETQ---HQIRK----
GROS_g01784 -----------PLLIEILSTRSNRQIAELKHEYERLYKKRPLEQDIKAHTK---GPFQN----
GROS_g12982 ----------KDILLEILITRSNKQLRTIKSKYKQLF-GHELSAEIASNTK---GSYND----
GROS_g01994 ----GC-GTKEHVLIEIMTTRTNSQIHEIKLAYRKKH-HTDLEKDIVDDTS---GYFRR----
GROS_g02107 ----GK-SFERSAAVEVGLTSTASQLGVIRDTYQNEF-RVLLEPDLSIKVE---GLFGK----
GROS_g07833 ----------
510 520 530 540 550 560
GROS_g03104 --------------LLVSLC---NESRDESWN-TDPLRANMVARTLFKKSEV-ESGVDDAVFN
GROS_g05922 --------------LLISLC---SGARDEQMH-TNPAKAQQDAQKLWPAGEG-CLGTDEIAFN
GROS_g13702 --------------LLVSLC---SGSRDERMQNTNAVKAAEDAPRMHPAAEP-CLGADEIAFN
GROS_g08389 --------------LLLDLV---EANRDETYR-TDPVKAKADADKLYKAGEG-KWGTDEDAFI
GROS_g11277 --------------LLLDLL---TANRDQSFV-EEPEKAK---KALYAAGEG-RLGTDEKVFI
GROS_g01976 IAITLGVLLVACVGLVVSVCKQSDSDEEAEPEQQDPEKGEVAPSTQAEVMEA-AAEEDIEKEK
GROS_g07837 --------------LLISL-------RLRRLL-ADVARGRS-----------TLSTDGITFK
GROS_g01954 --------------LLVSLC---NESRDESAN-TDALKANQAARILFKEGES-KKGVNDALFN
GROS_g08933 --------------FTADEY---HVDLDQ-FE-LAENDTEDDGQEEYEDDQY-YLEDEQQQQH
GROS_g01784 --------------LLVSLL---NGSRDEQRE-TDLALAHEEAVKLYREGEA-KRGVNAAAFN
GROS_g12982 --------------LLFTLL---KADRDESNK-TDRTKAKKDAVTLYKAGEK-RLGTDEEVFT
GROS_g01994 --------------LLISLC---SGARDEQMH-TDRTKAQQDAHKLWCAGEA-RLGTDEIAFN
GROS_g02107 --------------MLAQLL---LRKTDFADD-LDVEEAPREADVLFKEGNVEELGRSLELFN
GROS_g07833 --DAPPMHPAGEP-CLGADEIA--
570 580 590 600 610 620
GROS_g03104 QVLANENFNQLHLIFTEYE-----------KV-SGHTIDQAIQQ---------QFS-GETRD
GROS_g05922 SILAVQNFAQLHLVFAEYE-----------KQ-HGHPIEKAIRK---------EFS-GDIRD
GROS_g13702 SILAGQSFAQLQLLFAEYE-----------KQ-HGHPIEKAIQK---------EFS-GDICE
GROS_g08389 KVLANQSLNQLQMMFNEYE-----------KLGKKHTMEQAIKG---------EFS-GDELS
GROS_g11277 KILANQSLRQLRLLFDEYE-----------NL-TKHSIEKAIKS---------EFS-RDSLR
GROS_g01976 PVEVEQPPQQISVKAEEEK--------QEEVKDEPQTTSVKSVKPEEPTQSSKSNIATGEKSE
GROS_g07837 IILQIRR-------------------------HGRHL-----------------------------
GROS_g01954 KVLATENFSQLLLIFNEYQ-----------KI-SGHSIEQAIQK---------KFS-GDHRE
GROS_g08933 QYYLENNPDYYYAEYDDDQ-----------QPYNNQSGQQQMMPPHMAAHGHVPFHQGQAVQ
GROS_g01784 QVLANRNFAQLRITFDHYK-----------HL-SHHDIEQAIEH---------EFS-GDNEA
GROS_g12982 EIFSTQNVQQLRFVFDEYQ-----------KL-TGNSIEKALDS---------EFS-GDEFD
GROS_g01994 SILAAQNFGQLLLVFDEYQ-----------KQ-HGHSIEKAIRN---------EFS-GDICA
GROS_g02107 RVFGLRSPRQIQAFVERYDTLARDAQNRSENELASGRTESRTVRNFEWTIRKSVNIH-SDIKQ
GROS_g07833 ----------
640 650 660 670 680 690
GROS_g03104 ----------------GFMAVVECVRNRHA-FFAKLLQNATKGFFGIGNLGIGTRDSDLIRL
GROS_g05922 ----------------AHLALVKSIRNRPA-YFAELLYKSMKGL--------GTRDNDLIRL
GROS_g13702 ----------------ANLALIKFIRNRPA-YFAQLIQKAVKGL--------GTLDNDLIRL
GROS_g08389 ----------------ALLAIVSFVRNGPIGEVAEMLHKSLTK---------GGKDDMLIHL
GROS_g11277 ----------------ALLAIVEFVRNGPVGQVAKLMQKCIEG---------KADENLLVHL
GROS_g01976 TAEEEEEEKDSEASSSRTTMYNVHKQNNSFSSHFSSVKAEEEKQE--------EVKDEPQTTS
GROS_g07837 ----------
GROS_g01954 ----------------GFLAVVECVRNRPA-FFAKLLQNTTKGF--------GTRDSDLIRL
GROS_g08933 ----------------HHIG-----GQAPGLYYGDPTQQQQYQA--------GTAPPQYHHR
GROS_g01784 ----------------GFLALAQCVRNSSV-YFANLLHKSMKGL--------GTRDSDLIRL
GROS_g12982 ----------------ALMDIVSFVRQGPSGVLSRLMQKAIKG---------TPNDELLVHL
GROS_g01994 ----------------AHLALVKSICNRPA-YFAELLYKSMKGL--------GTRDNDLIRL
GROS_g02107 ----------------AILLYARISKNMQL-YFAEKLHEAVSQA--------RPDHQTIIRV
GROS_g07833 ----------


700 710 720 730 740 750
GROS_g03104 I----------VSRAECDMAEIK-------DQYMQMYNTTLENAIEKNC-SGSYKEGLLTLIK
GROS_g05922 V----------VTRSEIDMADIR-------KAYQEMYGKSLAAAIADDC-SGAYKRGLIAIVN
GROS_g13702 V----------VTRSEIDMADIR-------KEYQEIYGKSLEAAIADDC-SCAYKAS------
GROS_g08389 I----------VAHSELDLGDIA-------HEYEKRFKTPLEHAVETSGKPSTLKNALIPMIK
GROS_g11277 I----------YTYKELNLLSIL------ADEYNQKFKGPLAQAIQNSAKSAILKNALVYLIK
GROS_g01976 VKSVKPEEPTQSSKSKIATGEKS-------ETAEEEEEEKDSEASSSSA-ASSSSAGSSSASS
GROS_g07837 ----------------------------QGVQEMYSKSLGAAIADDR--------VLTRL-
GROS_g01954 I----------VTRAECDLYEIKVRDKQKAEQYVKLFDTTLEKAIGGDC-SGAYKDGLLTLTE
GROS_g08933 Q----------QQQQQQHYGDQR-------QQRQ--AGGPVRSQQQHQG--GQHPSQSFHQY-
GROS_g01784 V----------VSRSEIDLAEIK-------YAYHVLHQKPLEDAIQGDT-SGAYRDGLLALVR
GROS_g12982 I----------YSYKPMNLMEML------QDEFKHGFKGPLSEAVKDKAKTTIFKNTLINLIT
GROS_g01994 V----------VTRSEIDMVDIR-------TAYQKMHGKSLEAAIADDC-SGAYKRALIAIVK
GROS_g02107 A----------VSRSEIDLYDIC-------NEYRRKYGHHLVYDLQNMC-SGDFLRLLSRLVN
GROS_g07833 V----------VTRSEIGMADIR-------KEYQEIYRKSLEAAIAHDC-SRAYKAS------
760 770 780 790 800 810
GROS_g03104 --GN----------
GROS_g05922 --GH----------
GROS_g13702 ----------
GROS_g08389 --GN----------
GROS_g11277 SAGNR----------
GROS_g01976 GSGS----------
GROS_g07837 ----------
GROS_g01954 --GN----------
GROS_g08933 ----------
GROS_g01784 --GN----------
GROS_g12982 --G----------
GROS_g01994 --GNYHSVKAEEEKQEEVKDEPQTTSVKSVKPEEPTQSSKSNIATGEKSETAEEEEEKDSEAS
GROS_g02107 PPDV----------
GROS_g07833 ----------
820 830 840 850 860 870
GROS_g03104 ----------
GROS_g05922 ----------S----------
GROS_g13702 ----------
GROS_g08389 ----------K----------
GROS_g11277 ----------
GROS_g01976 ----------GSGSTSESSGGSATSSEASSSSASSSAASSASSS
GROS_g07837 ----------
GROS_g01954 ----------
GROS_g08933 ----------
GROS_g01784 ----------H----------
GROS_g12982 ----------
GROS_g01994 SSSAASSSSAGSSSASSGSGSGSGSTSESSGGSATSSEASSSSASSSAASSASSS
GROS_g02107 ----------ASGDEMFSANCGSCFGWLYLALFTQVNEWHSAGRR
GROS_g07833 ----------

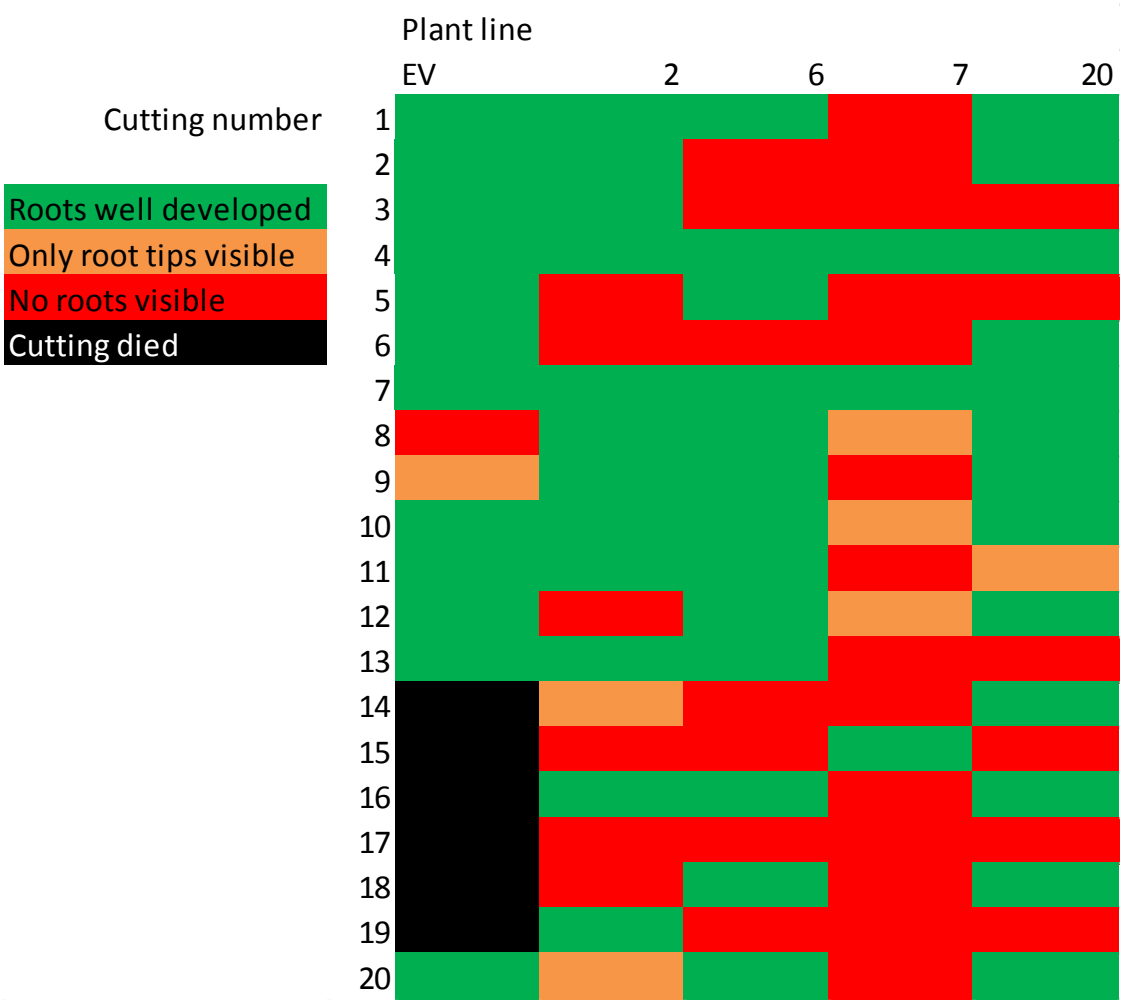
7.5 GROS_g03104 alignment against other nematode annexin orthologues



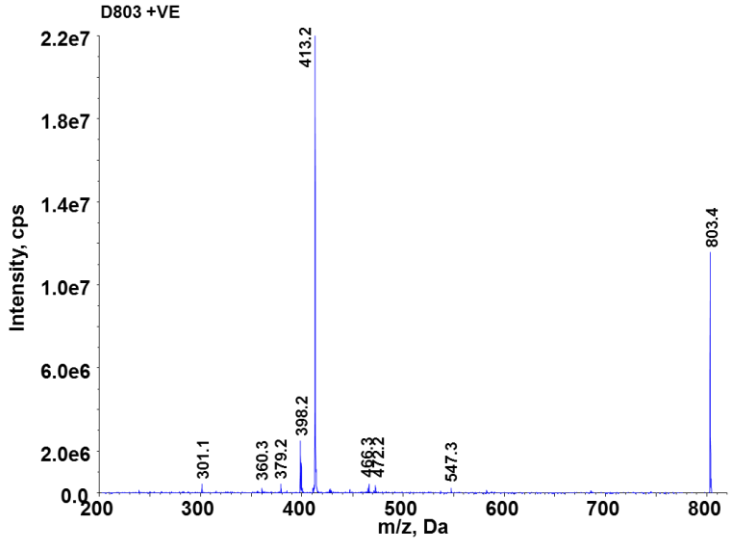

200 210 220 230 240 250
G.rostochiensis ----------MMSNAAKNILGT--PTIHNNANFNATITAQLLHKA----------I
G.pallida ----------MSNAAKNIVGT--PTIHDVANFNATITAQLLHKA----------I
G.ellingtonae ----------MMSNAAKNILGT--PTIHNNANFNATITAQLLHKA----------I
H.avenae ----------MMSNATKNILGT--PTIINNVNFDPKVSANFLHKA----------I
H.schachtii ----MLQNGLTILLLISVVIGHSLANLGPTIKHNPQFKAVQTAHHLHDA----------I
H.glycines ----MLQNGLTILLLISVVIGHSLANLGPTIKHNPHFKAVQTAHHLHDA----------I
M.incognita ----------MLKYFGISMAT--PSIRPNPNFNSQQTAELLPKA----------M
D.destructor ----------QRAFLMVQAQMLHWCG--PTIKHCLYFNAKLDAEILPKA----------M
H.contortus -QYGSGYSAPGAPYQMGQYKPEMLGT--PSVRPVPNFNANQDAETLPKAPILQEAKNTFYFMDLK
R.culicivorax ---------MPANNVASVQYYRHLPA--PTVHPHPNFNAQDDADALKKA----------M
S.carpocapse ----------MAYYGT----IPHNMNFDSEIWSGYLEKA----------M
H.bacteriophora ----------MTSPY--ATIIDNKDFNAPMMAEKIDRA----------L
P.pacificus ----------MWGN--PSVKAYPAFNANVDAETLPKA----------M
C.briggsae QQGGPQ---GGFFPNQGYGQPVMIGT--PSLFPVPGFNANADAEVLPKA----------M
C.japonica QQGGSG---GGFFPNQGYGQPVMIGN--PSLVPLPGFNANADAEVLPKA----------M
C.elegans QQGGQQQGGGGFFPNQGYGQPVMIGT--PSVFPVQGFNSNADAEVLPKA----------M
270 280 290 300 310 320
G.rostochiensis GE--KNKDEVIRLLCTISNQ----------QPQEVVVEFKSLF----------GEDL---PS
G.pallida GD--KNKDEIIRLLCTISNQ----------QPQEVVVEFKSLF----------GEDL---PS
G.ellingtonae GE--KSKDEIIRLLCTISNQ----------QPQEVVVEFKSLF----------GEDL---HS
H.avenae EE--KRKDEIIHLLCTISNQ----------QPQDLAIEFKQLF----------GEDL---FA
H.schachtii AK--KHEAEVTQIICSISNE----------QPQALASEFKKQF----------GTDL---IA
H.glycines AK--KHEAEVTQVICSISNE----------QPQALALEFKKQF----------GTDL---IA
M.incognita KGFGCDKGKIIQAIVNCSNV----------QPQEVSLSFKQMY----------GKDL---IK
D.destructor EGFGCDKDTIVQVLCARSNA----------QPQQIAREFKVLY----------KKDI---KK
H.contortus SELSGDFEDLILALMEPTAVYDAKQLHKAMQDLKSELSGDFEDLILALMEPTAVYDAKQLHKAMQ
R.culicivorax KGFGCDEKVIMNILCNRSSA----------QPQQIKEAFKLKY----------GDDL---IK
S.carpocapse KGSGCDKDAIVRHLTTISNA----------QPQMIRTPYKTKY----------GPDL---VE
H.bacteriophora RA--GEKDIVVKIITSLSNN----------QPQQLREPYRIKY----------GKDI---IG
P.pacificus KGFGCDNNKVIQVLCARTNG----------QPQEIARAFKTMY----------GKDL---LA
C.briggsae KGLGCNNSKVISVLCQRTNG----------QPQEISKAFKVMY----------GKDL---IK
C.japonica KGLGCNNSKVISVLCQRTNA----------QPQDISRAFKVMY----------GKDL---IK
C.elegans KGLGCNNSKVISILCQRTNW----------QPQEISKAFKVMY----------GKDL---IK
330 340 350 360 370 380
G.rostochiensis RLKKALSGDLEELILALLELPSVFEARQLYKAMSGLMGTKESVLIEILTTHSNRQIGEMKRVYEK
G.pallida KLKKALSGDFEELILALLELPSVYEARQLYKAMSGLMGTKESVLIEILTTHSNRQIGEMKRVYEK
G.ellingtonae RLKKALSGDFEELILALLELPSVYEARQLYKAMSGLMGTKESVLIEILTTHSNRQIGEMKRVYEK
H.avenae KLKKALSGDFEELILALLEQPAVYDAQQMHRAMSGL-GTKESVLIEILVTHSNRQIAEMKRVYEQ
H.schachtii MLKKEFKSDFEELIISLMQTPAVYDANQMRAALS---GSNETVLIEILATRTNRQITALKQAYEQ
H.glycines MLKKEFKSDFEELIISLMQTPAVYDANQMRAALS---GSNEAVLIEILATRTNRQITAPKQAYEQ
M.incognita ELKSELSGDFEDLIMALMSHPAYYDAAQLRRAMEGL-GTKEHILIEIMTTRTNAQIIQLPSAYAQ
D.destructor ELKSELSGAFEELILALMETPVHFICHQLPKAMKGV-GTPENVLVEVLISHSNEEIKEIKHEYYR
H.contortus DLKSELSGDFEDLILALMEPTAVYDAKQLHKAMQGL-GTKESVLIEIMTSRTNSQIAEIRNVYQQ
R.culicivorax EIKSELKGDLEDLIVALLYTPADYDALELYKAMKGL-GTDELTLIEIMTTRTNAEILAIKKAYQQ
S.carpocapse HLKKELSGQLEETMVALMETPTKYDVTQLHKAVKGL-GTKEKVLIEILCSRTNEEIAAIRNVYQD
H.bacteriophora ALDKKFSGDLEKTIFALMDTPLDYDLKQLKHAMKGL-GTDEAVLIEILCSRTPDQIRAIRAGYEK
P.pacificus DLRSELHGDFEDLVLALMERPPVYDAMQLRKAMEGL-GTRESVLIEIMTTRTNAQIGELPYVYRQ
C.briggsae ELKGELHGDFEDLILALMDAPAIYDAKQLYKAMDGL-GTKESVLIEIMTSRTNAQIQQVRDAYKM
C.japonica ELKGELHGDFEDLILALMDAPAVYDAKQLRKAMEGL-GTKESVLIEIMTSRTNAQIQQVRDTYKS
C.elegans ELKGELHGDFEDLILALMDAPAIYDAKQLHRAMEGL-GTKESVLIEIMTSRTNAQIQQVRDAYKM
400 410 420 430 440 450
G.rostochiensis LYGH----PLEKDIVGDTS-------GPFQHLLVSLCNESRDESWNTDPLRANMVARTLFKKSEV
G.pallida LYGH----PLEKDIVGDTS-------GPFQHLLVSLCNESRDESWNTDPLRANMVARTLFKKSEV
G.ellingtonae LYGH----PLEKDIVGDTS-------GPFQHLLVSLCNESRDESWNTDPLRANMVARTLFKKSEV
H.avenae LYGH----PLEKDIVGDTS-------GAFQRLLVSLCNESRDESWSSDPIRANLTARALFKEGEV
H.schachtii LDRPHQHNQLEEDIKAKTK-------GAFQNLLVSLLSCSREESAPASIVLAHHEAMKLFREGEG
H.glycines LDRPHQHNQLEEDIKAKTKRTLPKSVGVFAQLL------SPRKSAPASIVLAHDEAMKLFREGEG
M.incognita LYCS----DLERDMIGETS-------GYFKRMLVALCTGARDESMITDPLRANQDANKLYRAGEA
D.destructor LFDK----KLDKDIIGETS-------GEFQRLLVCLLSAGRDETRHTDNYKAESDARLLYEAWKK
H.contortus LYRT----SLESDLIGDTS-------GCFQ----------
R.culicivorax AYGK----DLEKALADDTS-------SHFKNFLVSLCNASRDESPHVDQNLAKIQAQQLHDAG--
S.carpocapse TYER----TLEKDIIGDTS-------GDFKRLLTSLIAGQRDQCDHVDDLKANQDARKLYDAGER
H.bacteriophora EYGK----QLEKEVASETS-------GEFKDLLVSLITGSKDTSHDTNDQAAKDDAVRLFADGKG
P.pacificus MYGS----DLEKDLIGETS-------GHFKRLLVSLCAGGRDESMHVDPVRANHDARKLYQAGEG
C.briggsae LYKK----DLERDLIGETS-------GHFKRLLVSLCAGGRDESNQTDPLRANQDARRLYQAGEK
C.japonica LYRK----DLEKELIGETS-------GHFKRLLISLCAGGRDESNQTDPLRANQDARRLYQAGEK
C.elegans LFKK----DLERDLIGETS-------GHFKRLLVSLCAGGRDESSQTDGLRANQDARRLLQAGEK

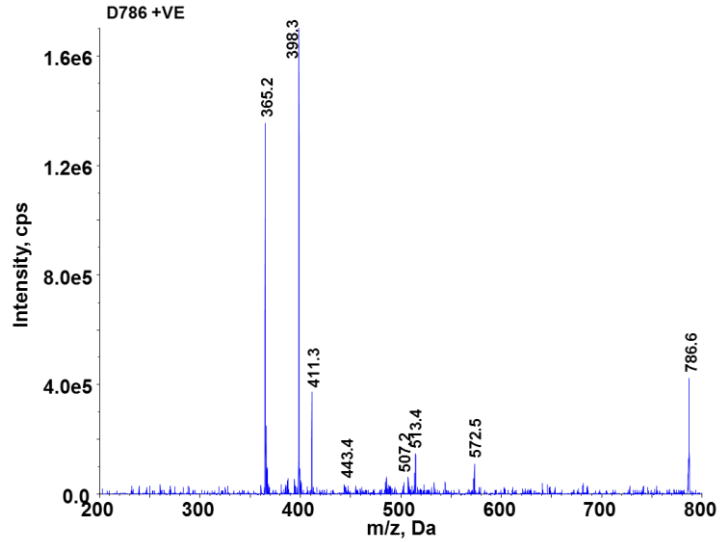

460 470 480 490 500 510
G.rostochiensis ESG-VDD-AVFNQVLANENFNQLHLIFTEYEKVSGHTIDQAIQQQFS-GETRDGFMAVVECVRNR
G.pallida ESG-VDD-AVFNQVLANENFNQLHLIFTEYEKVSGHTIDQAIQQQFS-GETRDGFMAVVECVRNR
G.ellingtonae ESG-VDD-AVFNQVLANENFNQLHLIFTEYEKVSGHTIDQAIQQQFS-GETRDGFMAVVECVRNR
H.avenae ESG-VND-VLFNQVLANENFIQLHLIFEEYQKISGHPIEQAIHQNFS-GETRDGYLAVVECVSNR
H.schachtii RRG-VNA-VVFNQVLATRSFAQLRETFEFYRQAAHHEIEKGIEQEFS-GHNEAGFLALIKYVRNA
H.glycines RRG-VNA-VVFNQVLATRSFAQLRETFEFYRQAAHHEIEEGIKQEFS-GHNEAGFLALIKYVRNA
M.incognita RLG-TDE-SAFIAILSAQNYNQLRLIFNEYQKVSGHPIEQAIAAEFS-GDIRDGLLAIIKSIKNR
D.destructor KCAITEE-VDFIKMLTTQNFDQIRLLFFKYNQIAGHPIESAIETQFRGGDTRDALLALARRIRNR
H.contortus ----------VTNHSIEKAIEAEFS-GDVKDGLLAVIAVAPNR
R.culicivorax --------ATFSQILATQNFAQLKQLFREYALLAQKDMESAVDKDIG-GDIKEALLAVVQIAKGK
S.carpocapse KLG-TDE-SCFNAILVQQSFRQLDRVFIEYEKITGHPIEEAIEKEFS-GDVKKGYLALIRCIQNR
H.bacteriophora KFASKDGPSHFLKILATQNHYQLRKVFAVFEDLSGSTIEKAIEKEFS-GDLQKSYLTIARAASDK
P.pacificus RLG-TDE-SQFNAILASQNYAQLRLVFDEYHKVCNRSIEQSIQNEFS-GDIRDGLLAVCAVVRSK
C.briggsae RLG-TDE-STFNAILASQNFNQLRMVFEEYQKVSNHSIEKAIEAEFS-GDVRDGLLAVIAVVRNR
C.japonica RLG-TDE-STFNAILASQNISQLRMVFDEYQKVSNHPIEKAIEAEFS-GDVKDGLLAVITVVRNR
C.elegans RLG-TDE-STFNAILASQNFSQLRLVFEEYQKASNHSIEKAIEFEFS-GDIRDGLLAVIAVIRNR
530 540 550 560 570 580
G.rostochiensis HAFFAKLLQNATKGFFGIGNLGIGTRDSDLIRLIVSRAECDMAEIKDQYMQMYNTTLENAIEKNC
G.pallida HAFFAKLLQNATKGFFGIGNFGIGTRDSDLIRLIVSRAECEMAEIKDQYMQMYNTTLENAIEKNC
G.ellingtonae HAFFAKLLQNATKGFFGIGNLGIGTRDSDLIRLIVSRAECDMAEIKDQYMQMYNTTLENAIEKNC
H.avenae HAFFAKVLQNSTKSFLGL-----GRRDTDLIRLIVTRAECDMVEIKEQYMQMFNTPLEKAIESNC
H.schachtii SVFFADLLFNSMKGL--------GTRDSDLIRLVISRSEVDLADIKHAFHTLHKKSLEEAIKGDT
H.glycines SVFFADLLFNSMKGL--------GTRDSDLIRLVISRSEIDLADIKHAFHTLHKKSLEEAIKGDT
M.incognita PAYFAELLYNSMKGF--------GTRDTDLIRLVVTRSEIDLADIRSAYQQLYGTSLEQAIASDC
D.destructor TAYFAGQLDKYTEGIL-------GTRDTDLIRVIVSRSEIDLVDIKMEFNRLQGVPLEPALGDSC
H.contortus PAYFAKLLYESMKGF--------GTRDSDLIRLVVTRCEYDMVDIRSAFQAMYKTTLENMIKGDC
R.culicivorax DNYFVNRINMAMKGM--------GTDDKTLIRVIASRSERDLLTTICQEWERQFANSLENAISAET
S.carpocapse PRYFAFQINKAVKGF--------GTRDHDLIRVIVSRSEIDLSIIKNEYERFFKEPLVDVITKEC
H.bacteriophora QKFFAIQLFNSMKGL--------GTRDNDLIRVLVSRSEVDLELIKEEFVAMYNKSLEDTVRGDT
P.pacificus PAYFATLLHNSMKGF--------GTRDGDLIRACVTRCEVDMADVRVAFEGMYHQPLEKMIKGDA
C.briggsae PAYFAKLLHDSMKGL--------GTRDNDLIRLCVTRAEYDMADIRNMFQSLYRTSLENMIKGDC
C.japonica PAYFAKLLHDSMKGL--------GTRDTDLIRLCVTRAEYDMADIRNVFQQLYRTSLENMIKGDC
C.elegans PAYFAKLLHDSMKGL--------GTRDNDLIRLCVTRAEYDMGDIRNMFQSLYRTSLENMIKGDC
590 600
G.rostochiensis SGSYK-EGLLTLIKGN----
G.pallida SGPYK-EGLLTLIKGN----
G.ellingtonae SGPYK-EGLLTLIKGN----
H.avenae SGPYK-DGLLTLIKGN----
H.schachtii SGAYR-DALLALVKGNTEQ-
H.glycines SGAYR-DALLALVKGNTEQ-
M.incognita SGAYK-DGLIAIVKGFYH--
D.destructor SGAYK-DALIALVKGN----
H.contortus SGSYK-DGLIALVTGNAHAH
R.culicivorax SGKYK-DALLLLIKGGQ---
S.carpocapse KGQYR-DALINLVKGN----
H.bacteriophora SGLYR-DALLAIVVGNKA--
P.pacificus LVRIPMDSLLLLEEIEMNFN
C.briggsae SGAYK-EGLIALVNGNRG--
C.japonica SGAYK-EGLIALVNGNRGGQ
C.elegans SGAYK-EGLIALVNGNRG--

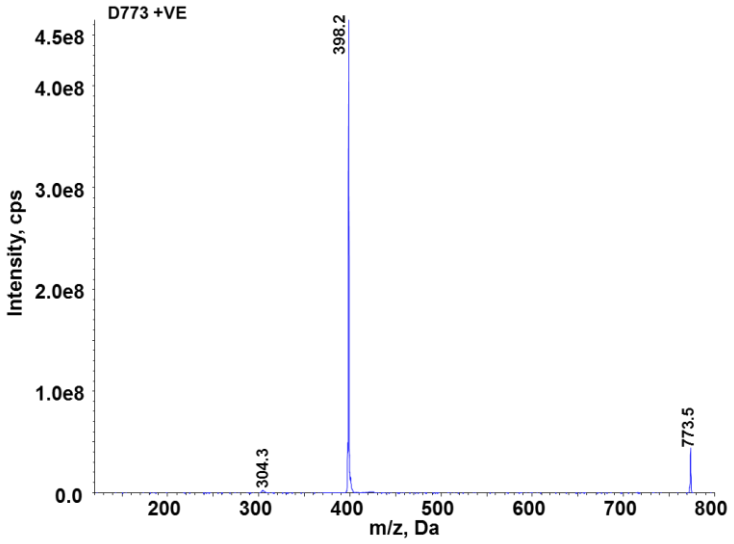
7.6 Scoring root systems of transgenic Desiree cuttings

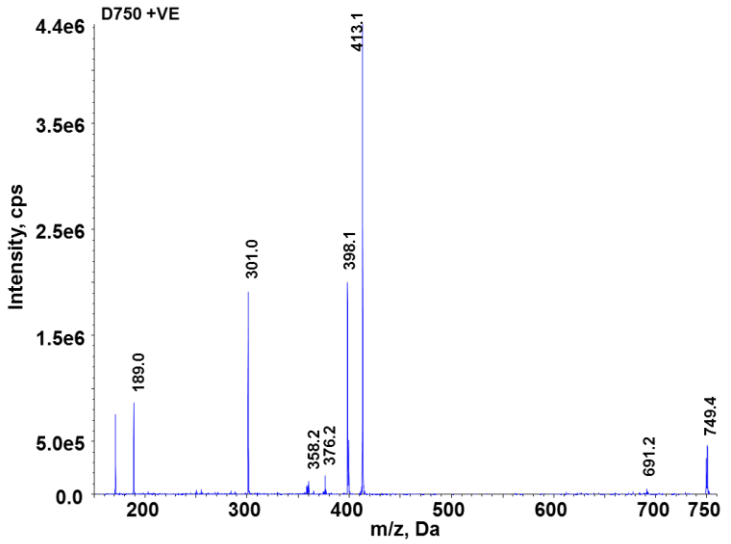


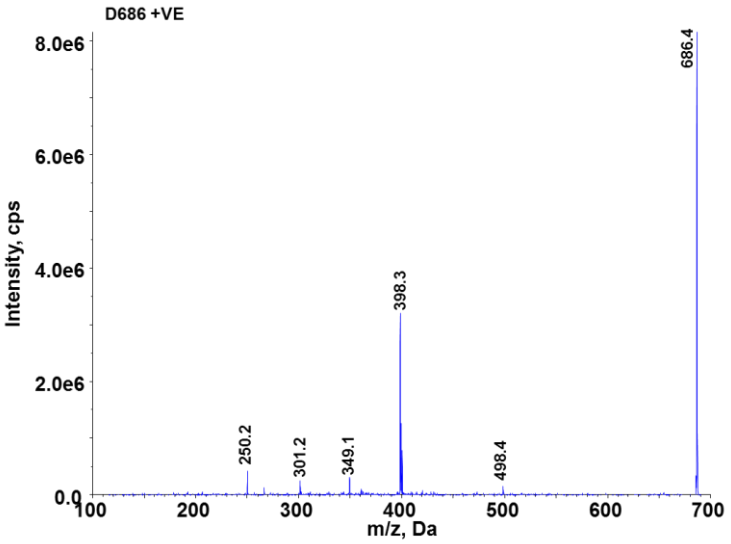
7.7 ESI-MS and MS-MS spectra

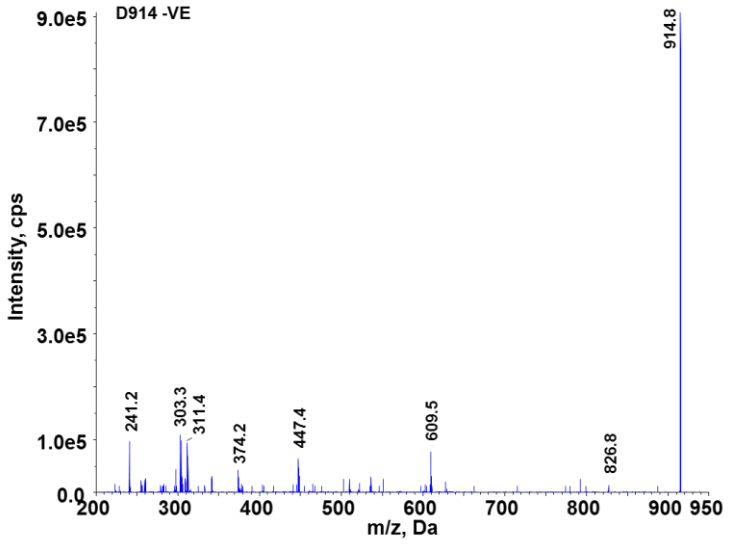


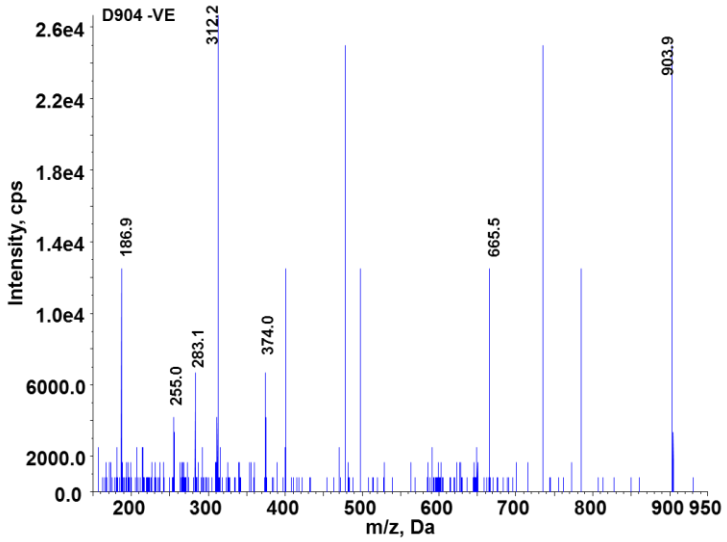


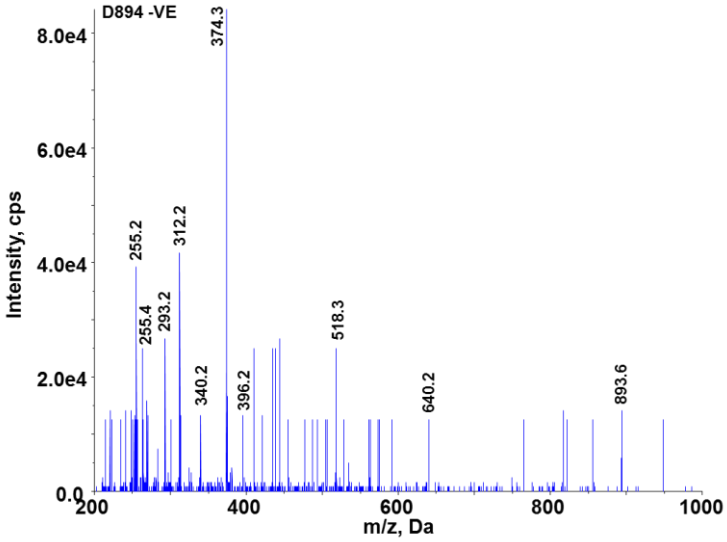


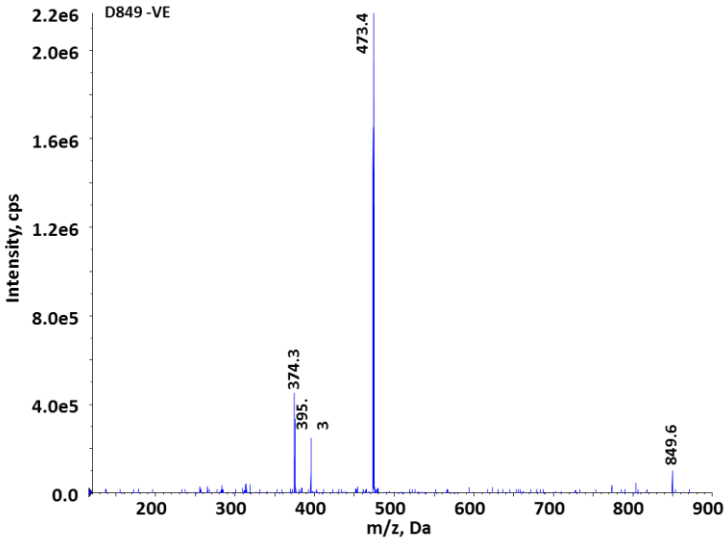

D750 +VE
Intensity, cps
m/z, Da
413.1
301.0
398.1
189.0
358.2
376.2
691.2
749.4

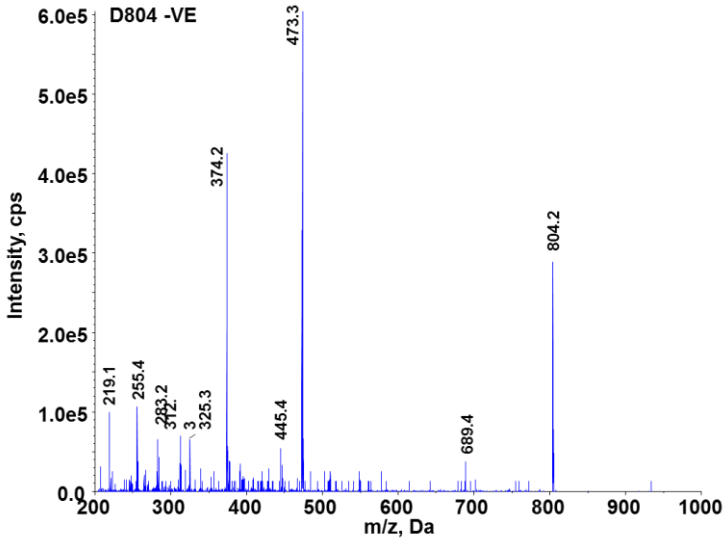

D686 +VE
Intensity, cps
m/z, Da
686.4
398.3
250.2
301.2
349.1
498.4

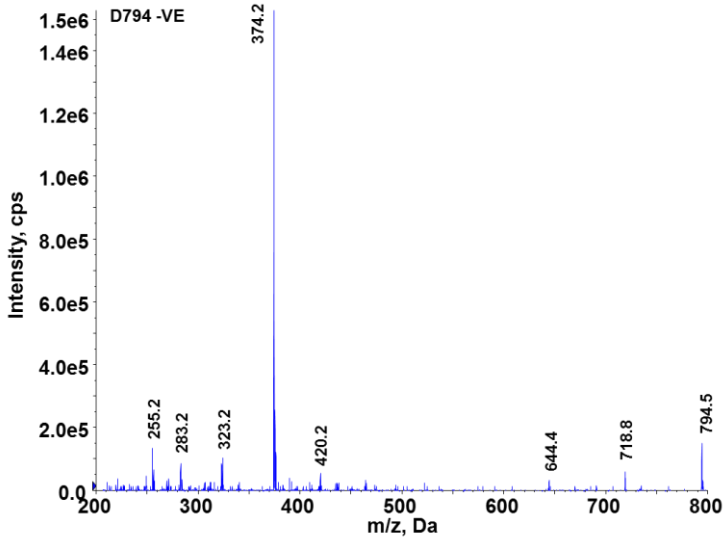

D914 -VE
Intensity, cps
m/z, Da
914.8
241.2
303.3
311.4
374.2
447.4
609.5
826.8

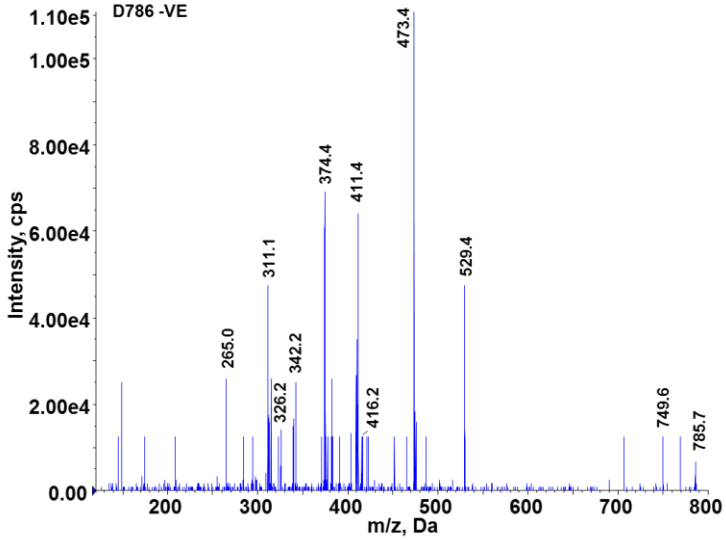

D904 -VE
Intensity, cps
m/z, Da
312.2
903.9
186.9
665.5
283.1
374.0
255.0

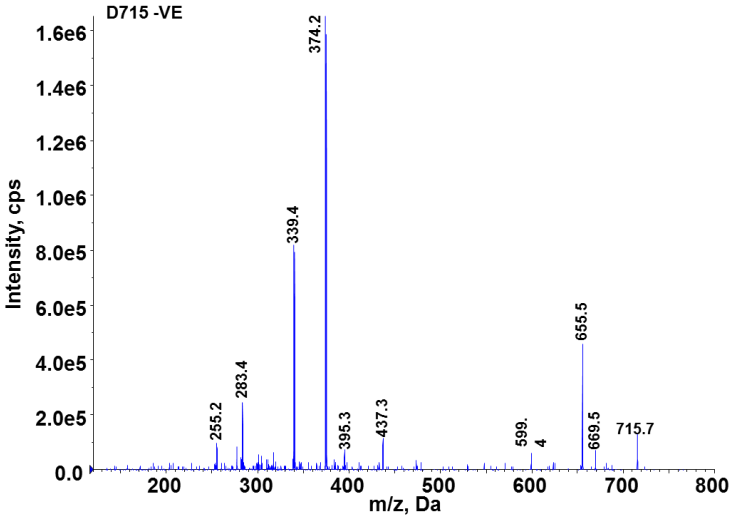

D894 -VE
Intensity, cps
m/z, Da
374.3
255.2
312.2
255.4
293.2
518.3
340.2
396.2
640.2
893.6

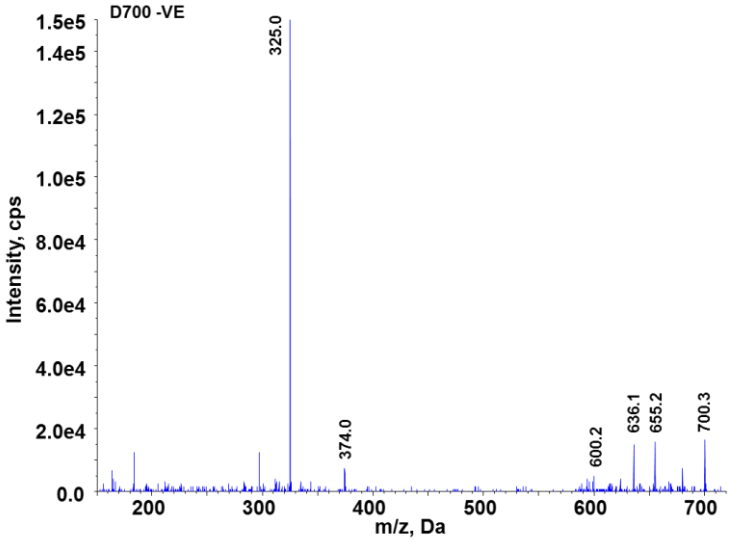

D849 -VE
Intensity, cps
m/z, Da
473.4
374.3
395.
3
849.6

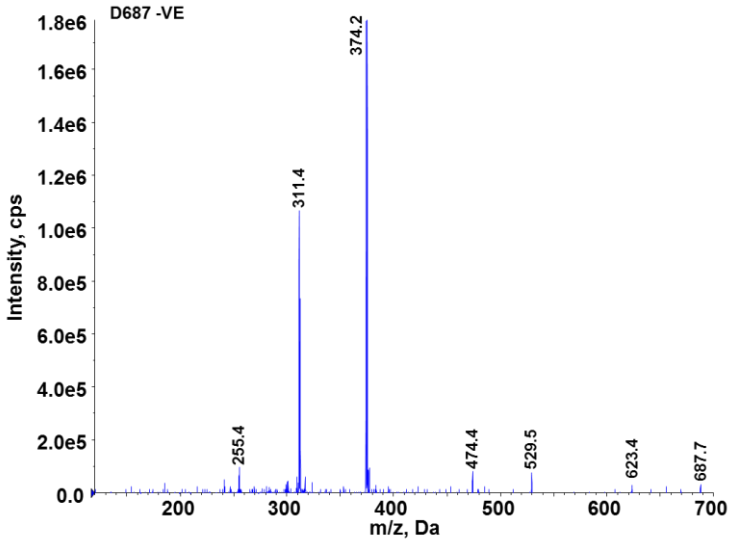

D804 -VE
Intensity, cps
m/z, Da
6.0e5
5.0e5
4.0e5
3.0e5
2.0e5
1.0e5
0.0
200
300
400
500
600
700
800
900
1000
219.1
255.4
283.2
312.3
325.3
374.2
445.4
473.3
689.4
804.2

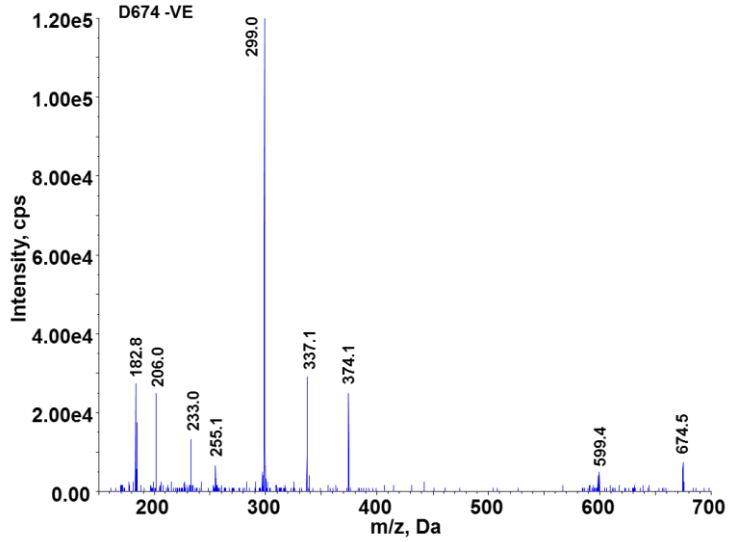

D794 -VE
Intensity, cps
m/z, Da
1.5e6
1.4e6
1.2e6
1.0e6
8.0e5
6.0e5
4.0e5
2.0e5
0.0
200
300
400
500
600
700
800
255.2
283.2
323.2
374.2
420.2
644.4
718.8
794.5

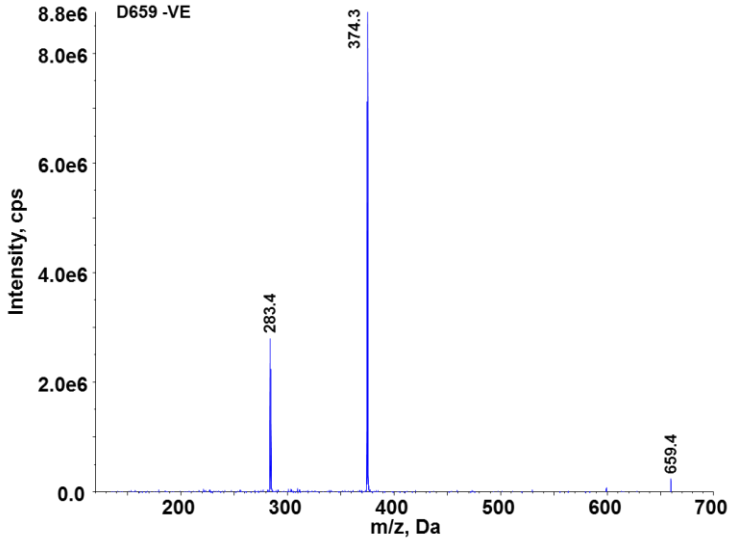

D786 -VE
Intensity, cps
m/z, Da
1.10e5
1.00e5
8.00e4
6.00e4
4.00e4
2.00e4
0.00
200
300
400
500
600
700
800
265.0
311.1
326.2
342.2
374.4
411.4
416.2
473.4
529.4
749.6
785.7

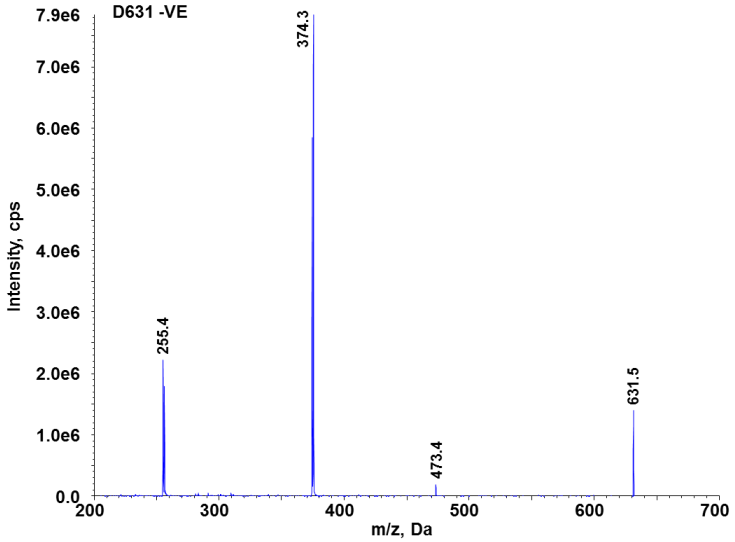

D715 -VE
Intensity, cps
m/z, Da
374.2
339.4
283.4
255.2
395.3
437.3
599.4
655.5
669.5
715.7


D700 -VE
Intensity, cps
m/z, Da
325.0
374.0
600.2
636.1
655.2
700.3

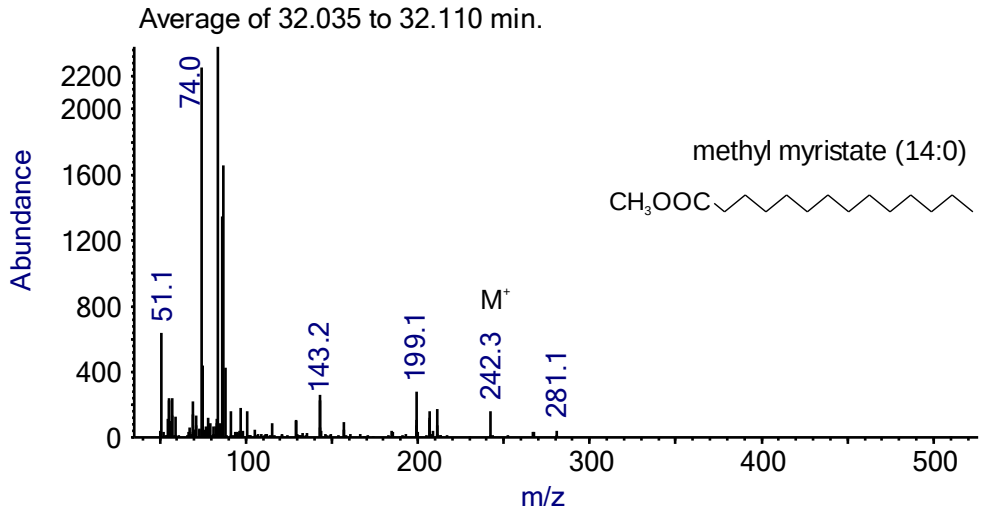

D687 -VE
Intensity, cps
m/z, Da
374.2
311.4
255.4
474.4
529.5
623.4
687.7

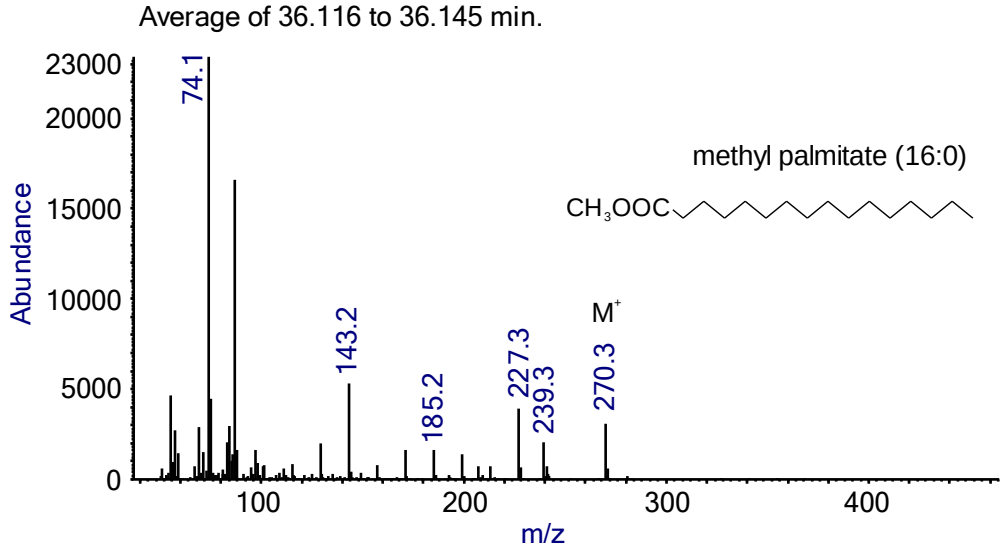

D674 -VE
Intensity, cps
m/z, Da
299.0
182.8
206.0
233.0
255.1
337.1
374.1
599.4
674.5

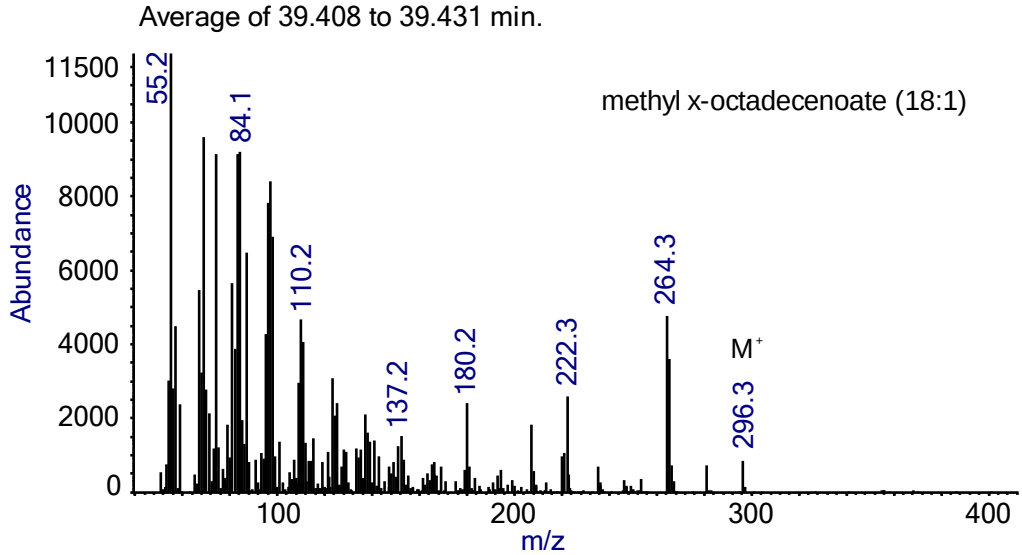

D659 -VE
Intensity, cps
m/z, Da
374.3
283.4
659.4

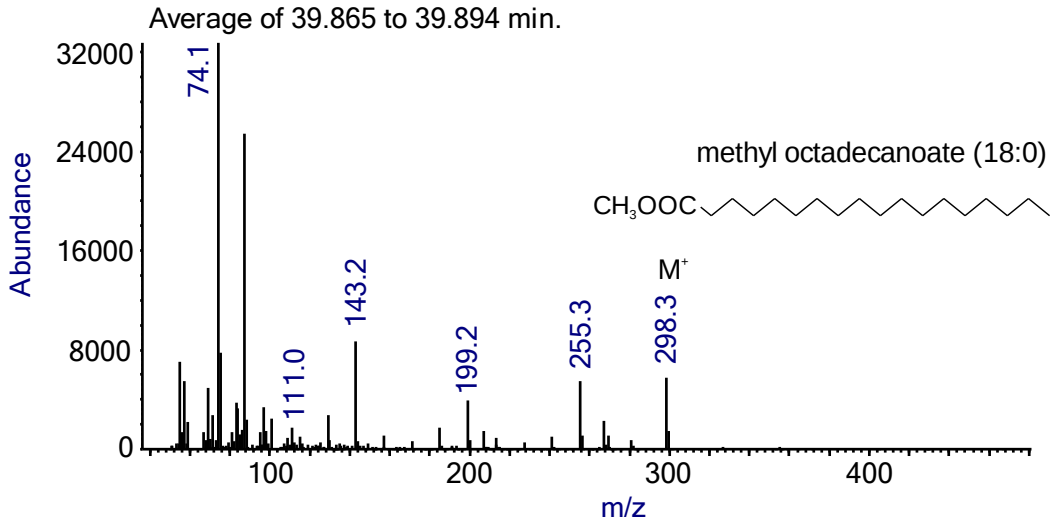

D631 -VE
Intensity, cps
m/z, Da
374.3
255.4
473.4
631.5

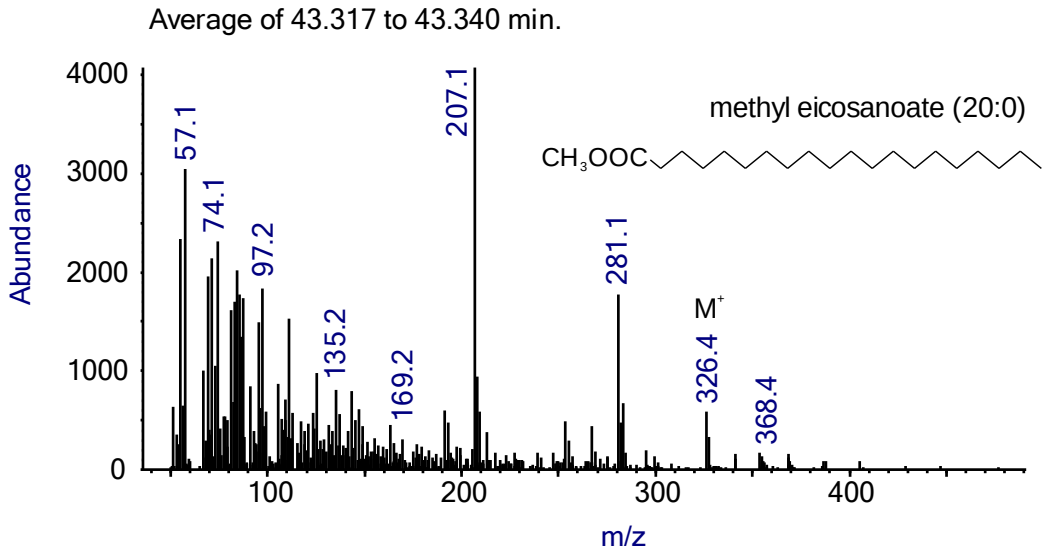
7.8 GC-MS spectra

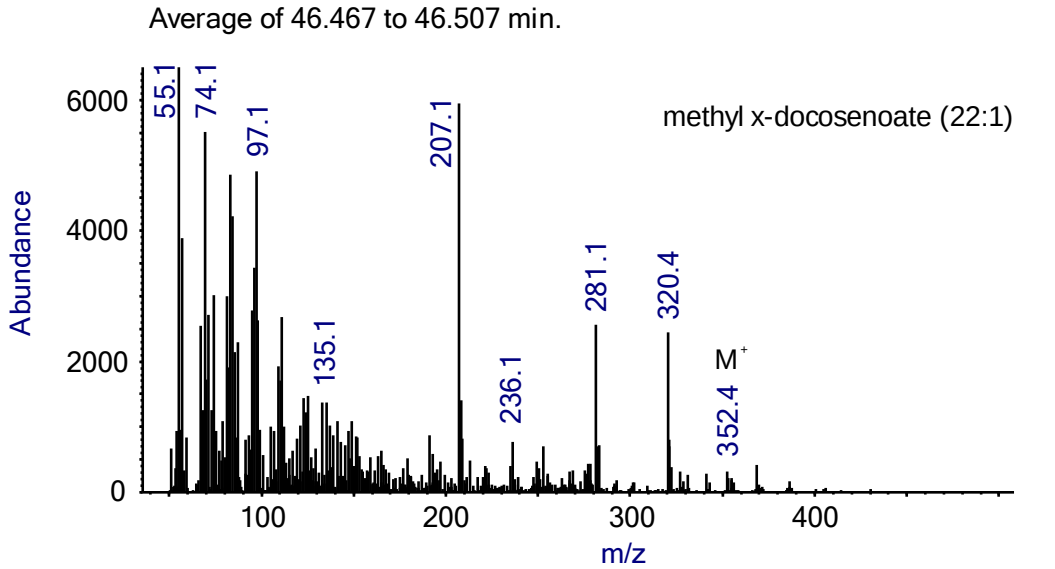


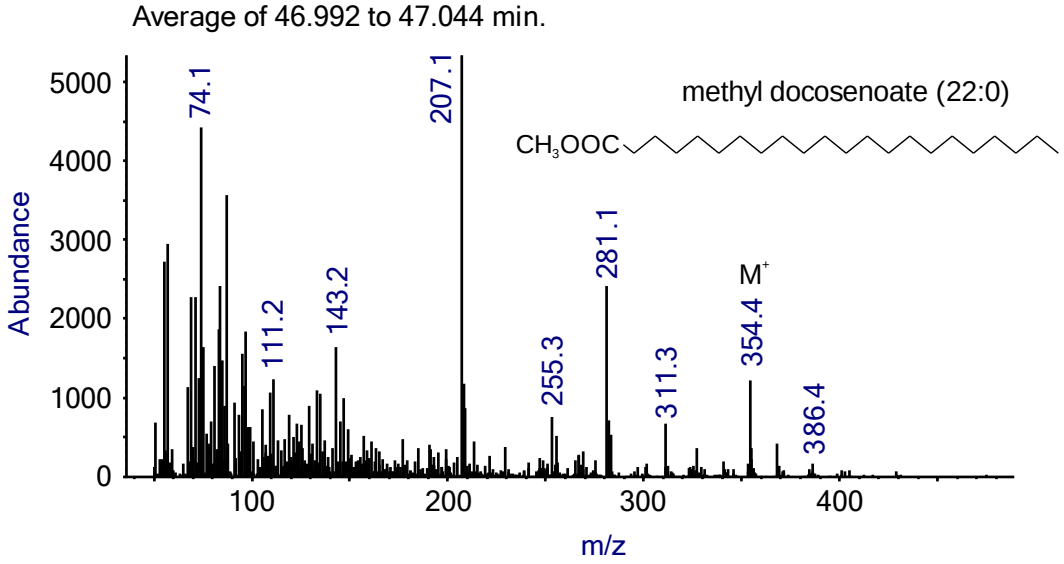




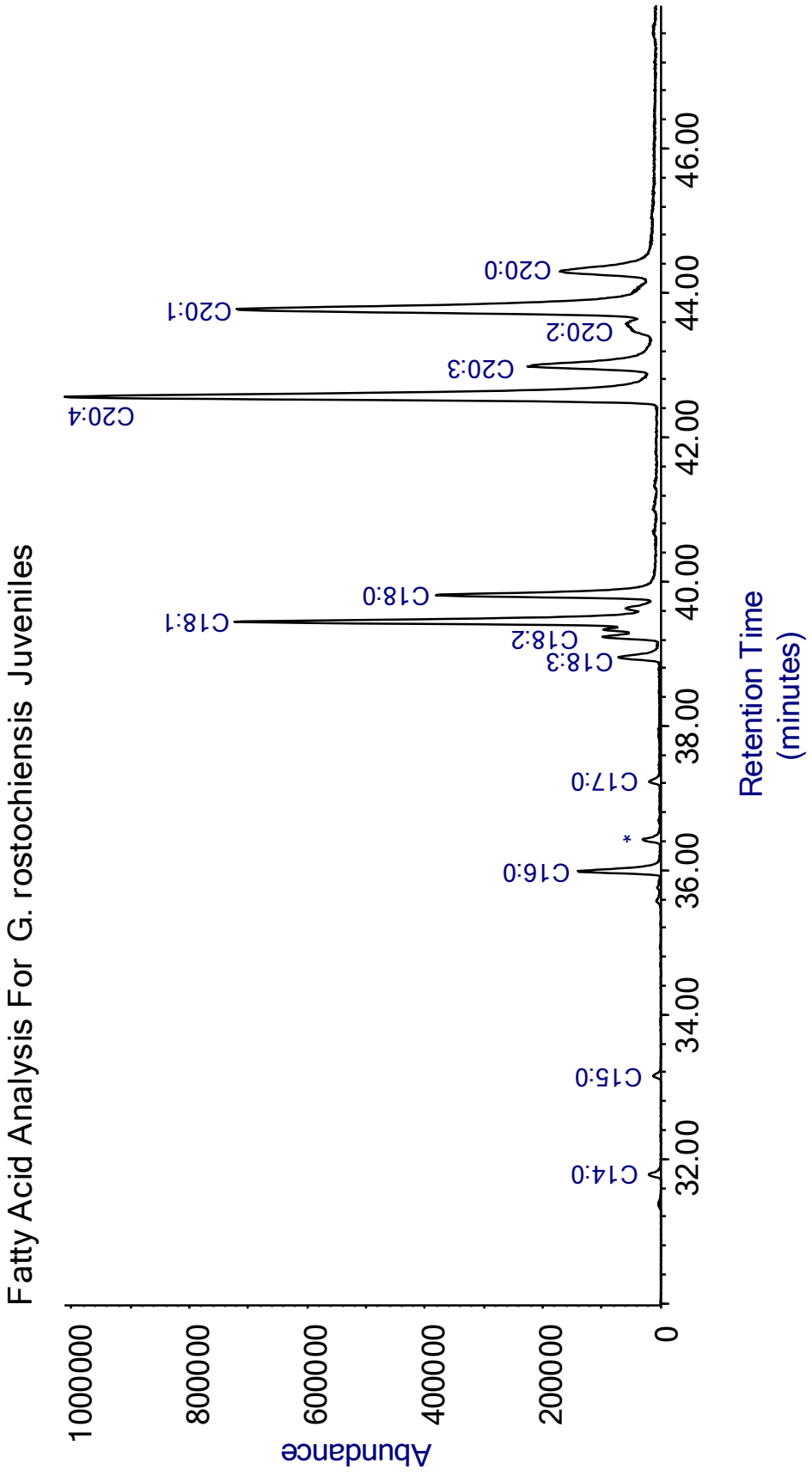

Average of 39.865 to 39.894 min.
Abundance
m/z
methyl octadecanoate (18:0)
CH3OOC
M+
74.1
111.0
143.2
199.2
255.3
298.3


Average of 43.317 to 43.340 min.
Abundance
m/z
methyl eicosanoate (20:0)
CH3OOC
M+
57.1
74.1
97.2
135.2
169.2
207.1
281.1
326.4
368.4


Average of 46.467 to 46.507 min.
Abundance
m/z
methyl x-docosenoate (22:1)
M+
55.1
74.1
97.1
135.1
207.1
236.1
281.1
320.4
352.4


Average of 46.992 to 47.044 min.
methyl docosenoate (22:0)
CH3OOC
Abundance
m/z
74.1
111.2
143.2
207.1
255.3
281.1
311.3
354.4
M+
386.4
5000
4000
3000
2000
1000
0
100
200
300
400

7.9 PCN juvenile FAMEs



7.10 Proteins with predicted hypodermis upstream motif and predicted signal peptide

If annotation column is blank then gene was unannotated in the globodera_rostochiensis.PRJEB13504.WBPS12.annotations.gff3 file

Gene number	Annotation	Number of motifs
GROS_g00016		1
GROS_g00030		1
GROS_g00035		1
GROS_g00082	lys_cragi lysozyme os=crassostrea gigas gn=lysoz pe=2 sv=1	1
GROS_g00084		1
GROS_g00102	txndl_caeel thioredoxin domain-containing protein os=caenorhabditis elegans gn= pe=1 sv=2	1
GROS_g00251		1
GROS_g00254	co4a1_caeel collagen alpha-1 chain os=caenorhabditis elegans gn=emb-9 pe=1 sv=5	1
GROS_g00286		1
GROS_g00304		2
GROS_g00367	toe1_bovin target of egr1 protein 1 os=bos taurus gn=toe1 pe=2 sv=1	1
GROS_g00436		1
GROS_g00437		1
GROS_g00457		3
GROS_g00462		1
GROS_g00549	rap1_caeel ras-related protein rap-1 os=caenorhabditis elegans gn=rap-1 pe=3 sv=1	1
GROS_g00614		1
GROS_g00621		1
GROS_g00666		1
GROS_g00707		1
GROS_g00786	can5_caeel calpain-5 os=caenorhabditis elegans gn=tra-3 pe=1 sv=1	1
GROS_g00799	s38a9_caeel sodium-coupled neutral amino acid transporter 9 homolog os=caenorhabditis elegans gn= pe=3 sv=2	1

GROS_g00820		1
GROS_g00858	npl11_caeel neprilysin-11 os=caenorhabditis elegans gn=nep-11 pe=1 sv=2	1
GROS_g00861	naaa_rat n-acylethanolamine-hydrolyzing acid amidase os=rattus norvegicus gn=naaa pe=1 sv=1	1
GROS_g00872		1
GROS_g00906	gnrr2_claga gonadotropin-releasing hormone ii receptor os=clarias gariepinus pe=2 sv=2	1
GROS_g00963		1
GROS_g00976	cdc2h_plach cell division control protein 2 homolog os=plasmodium chabaudi gn=crk2 pe=3 sv=1	1
GROS_g01012	syg2_caeel synaptogenesis protein syg-2 os=caenorhabditis elegans gn=syg-2 pe=1 sv=1	1
GROS_g01030		1
GROS_g01075		1
GROS_g01083		1
GROS_g01095		1
GROS_g01104		1
GROS_g01105		1
GROS_g01107		1
GROS_g01153	lec3_caeel 32 kda beta-galactoside-binding lectin lec-3 os=caenorhabditis elegans gn=lec-3 pe=2 sv=3	1
GROS_g01164	ubp44_danre ubiquitin carboxyl-terminal hydrolase 44 os=danio rerio gn=usp44 pe=2 sv=1	1
GROS_g01182		1
GROS_g01189		1
GROS_g01272		1
GROS_g01291	klf7_human krueppel-like factor 7 os=homo sapiens gn=klf7 pe=2 sv=1	1
GROS_g01307		1
GROS_g01320	rbgpr_caeel rab3 gtpase-activating protein regulatory subunit os=caenorhabditis elegans gn=rbg-2 pe=3 sv=1	1
GROS_g01335	ykr5_caeel probable g-protein coupled receptor os=caenorhabditis elegans gn= pe=3 sv=3	1

GROS_g01365	mfsd8_danre major facilitator superfamily domain-containing protein 8 os=danio rerio gn=mfsd8 pe=2 sv=1	2
GROS_g01376		1
GROS_g01377		1
GROS_g01379	col36_caeel cuticle collagen 36 os=caenorhabditis elegans gn=col-36 pe=3 sv=1	1
GROS_g01391	hsp90_brupa heat shock protein 90 os=brugia pahangi gn=hsp90 pe=2 sv=2	1
GROS_g01432		1
GROS_g01443	gbr1_caeel gamma-aminobutyric acid receptor exp-1 os=caenorhabditis elegans gn=exp-1 pe=1 sv=1	1
GROS_g01454	gun2_bacsu endoglucanase os=bacillus subtilis (strain 168) gn=egls pe=1 sv=1	1
GROS_g01518	nfu1_human nfu1 iron-sulfur cluster scaffold mitochondrial os=homo sapiens gn=nfu1 pe=1 sv=2	1
GROS_g01528	acha7_macmu neuronal acetylcholine receptor subunit alpha-7 os=macaca mulatta gn=chrna7 pe=2 sv=1	1
GROS_g01546		1
GROS_g01561		1
GROS_g01587	neur_drome protein neuralized os=drosophila melanogaster gn=neur pe=1 sv=2	1
GROS_g01618		1
GROS_g01632		1
GROS_g01765		2
GROS_g01776	pcdgc_pantr protocadherin gamma-a12 os=pan troglodytes gn=pcdhga12 pe=3 sv=1	1
GROS_g01798	pigb_xenla gpi mannosyltransferase 3 os=xenopus laevis gn=pigb pe=2 sv=1	1
GROS_g01815		1
GROS_g01917		1
GROS_g01952		1
GROS_g01957		1
GROS_g01992	kdm4_caeel lysine-specific demethylase 4 os=caenorhabditis elegans gn=jmjd-2 pe=3 sv=2	1
GROS_g02023		1

GROS_g02044		1
GROS_g02046		1
GROS_g02061	tsn1_mouse tetraspanin-1 os=mus musculus gn=tspan1 pe=2 sv=1	1
GROS_g02122		1
GROS_g02225	calr_oncvo calreticulin os=onchocerca volvulus gn=crt-1 pe=2 sv=2	1
GROS_g02229	se1l1_rat protein sel-1 homolog 1 os=rattus norvegicus gn=sel1l pe=2 sv=2	1
GROS_g02233		1
GROS_g02261	col90_caeel cuticle collagen 90 os=caenorhabditis elegans gn=col-90 pe=3 sv=1	1
GROS_g02276		1
GROS_g02323		1
GROS_g02341	s39a1_danre zinc transporter zip1 os=danio rerio gn=slc39a1 pe=2 sv=1	1
GROS_g02355		1
GROS_g02358		1
GROS_g02365	hibch_danre 3-hydroxyisobutyryl- mitochondrial os=danio rerio gn=hibch pe=2 sv=1	1
GROS_g02381		2
GROS_g02396		1
GROS_g02446		1
GROS_g02458		2
GROS_g02490	gpx3_caeel glutathione peroxidase 3 os=caenorhabditis elegans gn=gpx-3 pe=3 sv=1	1
GROS_g02579	t106b_danre transmembrane protein 106b os=danio rerio gn=tmem106b pe=3 sv=1	1
GROS_g02610		1
GROS_g02747		1
GROS_g02752		1
GROS_g02826		1
GROS_g02865	cks1_caebr cyclin-dependent kinases regulatory subunit os=caenorhabditis briggsae gn=cks-1 pe=3 sv=1	1
GROS_g02899	gcy10_caeel receptor-type guanylate cyclase gcy-10 os=caenorhabditis elegans gn=odr-1 pe=1 sv=2	2

GROS_g02909	tyr3_caeel tyrosinase-like protein tyr-3 os=caenorhabditis elegans gn=tyr-3 pe=3 sv=5	1
GROS_g02936		1
GROS_g02968	mup4_caeel transmembrane matrix receptor mup-4 os=caenorhabditis elegans gn=mup-4 pe=1 sv=4	1
GROS_g03108	lgc50_caeel ligand-gated ion channel 50 os=caenorhabditis elegans gn=lgc-50 pe=3 sv=3	2
GROS_g03131		1
GROS_g03153	padc1_mouse protease-associated domain-containing protein 1 os=mus musculus gn=pradc1 pe=1 sv=1	2
GROS_g03181		1
GROS_g03186		2
GROS_g03236		2
GROS_g03242		1
GROS_g03254	met13_bovin methyltransferase-like protein 13 os=bos taurus gn=mettl13 pe=2 sv=1	1
GROS_g03316	syem_chick probable glutamate--trna mitochondrial os=gallus gallus gn=ears2 pe=2 sv=1	1
GROS_g03442	rrp1_caeel ribosomal rna processing protein 1 homolog os=caenorhabditis elegans gn= pe=3 sv=2	1
GROS_g03450		1
GROS_g03459		1
GROS_g03474		1
GROS_g03483		1
GROS_g03503		1
GROS_g03514	trea_rabit trehalase os=oryctolagus cuniculus gn=treh pe=1 sv=1	1
GROS_g03525		1
GROS_g03621		1
GROS_g03625		1
GROS_g03633		1
GROS_g03640		1
GROS_g03641		1
GROS_g03662	nhr1_caeel nuclear hormone receptor family member nhr-1 os=caenorhabditis elegans gn=nhr-1 pe=2 sv=4	1

GROS_g03841	ars_hempu arylsulfatase os=hemicentrotus pulcherrimus pe=1 sv=1	1
GROS_g03853		1
GROS_g03872		1
GROS_g03931	notc1_mouse neurogenic locus notch homolog protein 1 os=mus musculus gn=notch1 pe=1 sv=3	1
GROS_g03990	let4_caeel leucine-rich repeat-containing protein let-4 os=caenorhabditis elegans gn=let-4 pe=2 sv=2	1
GROS_g04009	nlgn1_caeel neuroligin-1 os=caenorhabditis elegans gn=nlg-1 pe=1 sv=1	1
GROS_g04094	twk7_caeel t family of potassium channels protein 7 os=caenorhabditis elegans gn=twk-7 pe=3 sv=3	1
GROS_g04149		1
GROS_g04208	ach8_caeel neuronal acetylcholine receptor subunit eat-2 os=caenorhabditis elegans gn=eat-2 pe=2 sv=1	1
GROS_g04267		1
GROS_g04280		1
GROS_g04287	bub1_caeel mitotic checkpoint serine threonine-protein kinase bub1 os=caenorhabditis elegans gn=bub-1 pe=1 sv=1	2
GROS_g04357		1
GROS_g04359		1
GROS_g04362		1
GROS_g04396	dpog1_drome dna polymerase subunit gamma- mitochondrial os=drosophila melanogaster gn=tam pe=1 sv=2	1
GROS_g04473		1
GROS_g04500		3
GROS_g04677	gunz_dicd3 endoglucanase z os=dickeya dadantii (strain 3937) gn=celz pe=1 sv=2	1
GROS_g04702	nucg_caeel endonuclease mitochondrial os=caenorhabditis elegans gn=cps-6 pe=1 sv=1	1
GROS_g04728	s39a1_danre zinc transporter zip1 os=danio rerio gn=slc39a1 pe=2 sv=1	1
GROS_g04827	trnt1_human cca trna nucleotidyltransferase mitochondrial os=homo sapiens gn=trnt1 pe=1 sv=2	1
GROS_g04885	stan_drome protocadherin-like wing polarity protein stan os=drosophila melanogaster gn=stan pe=1 sv=4	1

GROS_g04940		1
GROS_g04943		1
GROS_g04978		1
GROS_g04990		1
GROS_g05045	nas35_caeel zinc metalloproteinase dpy-31 os=caenorhabditis elegans gn=dpy-31 pe=1 sv=2	2
GROS_g05049	sdk2_chick protein sidekick-2 os=gallus gallus gn=sdk2 pe=1 sv=2	1
GROS_g05064		1
GROS_g05126	pptc7_xenla protein phosphatase ptc7 homolog os=xenopus laevis gn=pptc7 pe=2 sv=1	1
GROS_g05186		1
GROS_g05191		1
GROS_g05336	alat2_xenla alanine aminotransferase 2 os=xenopus laevis gn=gpt2 pe=2 sv=1	1
GROS_g05526		1
GROS_g05583		1
GROS_g05601	ugt60_caeel udp-glucuronosyltransferase ugt-60 os=caenorhabditis elegans gn=ugt-60 pe=3 sv=2	1
GROS_g05680	notc1_danre neurogenic locus notch homolog protein 1 os=danio rerio gn=notch1a pe=2 sv=1	1
GROS_g05682		1
GROS_g05687	glrb4_caeel glycine receptor subunit beta-type 4 os=caenorhabditis elegans gn=ggr-1 pe=3 sv=2	2
GROS_g05703		1
GROS_g05738		2
GROS_g05756		1
GROS_g05758	cp4cu_blage cytochrome p450 4c21 os=blattella germanica gn=cyp4c21 pe=2 sv=1	1
GROS_g05773		1
GROS_g05781	wrt6_caeel warthog protein 6 os=caenorhabditis elegans gn=wrt-6 pe=1 sv=2	2
GROS_g05848		1
GROS_g05878	ach10_chick neuronal acetylcholine receptor subunit alpha-10 os=gallus gallus gn=chrna10 pe=3 sv=1	1
GROS_g05879		1

GROS_g05931	edem3_xenla er degradation-enhancing alpha-mannosidase-like protein 3 os=xenopus laevis gn=edem3 pe=2 sv=2	1
GROS_g05955		1
GROS_g05987		1
GROS_g06052		1
GROS_g06088	alxa_caeel apoptosis-linked gene 2-interacting protein x 1 os=caenorhabditis elegans gn=alx-1 pe=2 sv=3	1
GROS_g06161		1
GROS_g06224	mgt4a_rat alpha- -mannosyl-glycoprotein 4-beta-n-acetylglucosaminyltransferase a os=rattus norvegicus gn=mgat4a pe=2 sv=1	1
GROS_g06240		2
GROS_g06242		1
GROS_g06243	lon3_caeel cuticle collagen lon-3 os=caenorhabditis elegans gn=lon-3 pe=1 sv=1	1
GROS_g06268		1
GROS_g06308	txd12_human thioredoxin domain-containing protein 12 os=homo sapiens gn=txndc12 pe=1 sv=1	1
GROS_g06311		2
GROS_g06322		1
GROS_g06355	vps55_caeel vacuolar protein sorting-associated protein 55 homolog os=caenorhabditis elegans gn= pe=3 sv=3	1
GROS_g06362	ch3l1_rat chitinase-3-like protein 1 os=rattus norvegicus gn=chi3l1 pe=2 sv=3	1
GROS_g06590		1
GROS_g06597		1
GROS_g06655		1
GROS_g06693		1
GROS_g06749		1
GROS_g06834	tyr1_caeel tyrosinase-like protein tyr-1 os=caenorhabditis elegans gn=tyr-1 pe=3 sv=2	1
GROS_g06835	tyr1_caeel tyrosinase-like protein tyr-1 os=caenorhabditis elegans gn=tyr-1 pe=3 sv=2	1
GROS_g06848		1
GROS_g06869		1

GROS_g06970	ift80_human intraflagellar transport protein 80 homolog os=homo sapiens gn=ift80 pe=1 sv=3	2
GROS_g06992		1
GROS_g07110		1
GROS_g07117		1
GROS_g07168		1
GROS_g07197		1
GROS_g07218	ttr5_caeel transthyretin-like protein 5 os=caenorhabditis elegans gn=ttr-5 pe=3 sv=1	1
GROS_g07312		1
GROS_g07318		1
GROS_g07336	co145_caeel cuticle collagen 145 os=caenorhabditis elegans gn=col-145 pe=3 sv=2	2
GROS_g07338	gun_closa endoglucanase os=clostridium saccharobutylicum gn=egla pe=3 sv=1	2
GROS_g07341		2
GROS_g07343	ima3_caeel importin subunit alpha-3 os=caenorhabditis elegans gn=ima-3 pe=1 sv=2	1
GROS_g07346	cp2m1_oncmy cytochrome p450 2m1 os=oncorhynchus mykiss gn=cyp2m1 pe=1 sv=1	1
GROS_g07381		1
GROS_g07482	tbcd9_human tbc1 domain family member 9 os=homo sapiens gn=tbc1d9 pe=2 sv=2	1
GROS_g07492		1
GROS_g07518	cup4_caebr acetylcholine receptor-like protein cup-4 os=caenorhabditis briggsae gn=cup-4 pe=3 sv=2	1
GROS_g07532	ymp8_caeel uncharacterized glycosyltransferase os=caenorhabditis elegans gn= pe=3 sv=3	1
GROS_g07574		1
GROS_g07595		1
GROS_g07653		2
GROS_g07665		1
GROS_g07678		1
GROS_g07692	ctns_caeel cystinosin homolog os=caenorhabditis elegans gn=ctns-1 pe=3 sv=2	1

GROS_g07745		1
GROS_g07764	glrb4_caeel glycine receptor subunit beta-type 4 os=caenorhabditis elegans gn=ggr-1 pe=3 sv=2	1
GROS_g07872		1
GROS_g07943		1
GROS_g07976		1
GROS_g07987	clc2_caeel clc-like protein 2 os=caenorhabditis elegans gn=clc-2 pe=2 sv=2	1
GROS_g08021		1
GROS_g08166	cle1_gloro clavata3 esr -related protein 1 os=globodera rostochiensis gn=cle-1 pe=2 sv=1	1
GROS_g08196	amdl_caeel probable peptidyl-glycine alpha-amidating monooxygenase pamn-1 os=caenorhabditis elegans gn=pamn-1 pe=1 sv=2	1
GROS_g08202	est1_caebr gut esterase 1 os=caenorhabditis briggsae gn=ges-1 pe=2 sv=1	1
GROS_g08205		1
GROS_g08276	ptp4_caeel tyrosine-protein phosphatase 4 os=caenorhabditis elegans gn=ptp-4 pe=2 sv=3	1
GROS_g08284		1
GROS_g08339		1
GROS_g08342		1
GROS_g08356		1
GROS_g08384		1
GROS_g08408	rft2_salsa riboflavin transporter 2 os=salmo salar gn=rft2 pe=2 sv=1	1
GROS_g08526		1
GROS_g08527		1
GROS_g08540	tens_chick tensin os=gallus gallus gn=tns pe=1 sv=2	1
GROS_g08554		2
GROS_g08559	sptc1_caeel serine palmitoyltransferase 1 os=caenorhabditis elegans gn=sptl-1 pe=3 sv=1	1
GROS_g08584		1
GROS_g08587		1

GROS_g08588	lam2_caeel laminin-like protein lam-2 os=caenorhabditis elegans gn=lam-2 pe=1 sv=3	1
GROS_g08589	parl_ponab presenilins-associated rhomboid-like mitochondrial os=pongo abelii gn=parl pe=2 sv=1	1
GROS_g08608	ost3_caeel probable dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 3 os=caenorhabditis elegans gn= pe=3 sv=2	1
GROS_g08642		1
GROS_g08712		1
GROS_g08729		1
GROS_g08879	rl40_brarp ubiquitin-60s ribosomal protein l40 os=brassica rapa pekinensis pe=2 sv=2	1
GROS_g08888		1
GROS_g08985	dus2l_mouse trna-dihydrouridine synthase	1
GROS_g09000		1
GROS_g09208	gde_horse glycogen debranching enzyme os=equus caballus gn=agl pe=2 sv=1	1
GROS_g09301		2
GROS_g09312		1
GROS_g09362		1
GROS_g09395		1
GROS_g09415	nas5_caeel zinc metalloproteinase nas-5 os=caenorhabditis elegans gn=nas-5 pe=3 sv=2	2
GROS_g09435		1
GROS_g09453	lgc50_caeel ligand-gated ion channel 50 os=caenorhabditis elegans gn=lgc-50 pe=3 sv=3	1
GROS_g09510		2
GROS_g09554		3
GROS_g09568		1
GROS_g09592		1
GROS_g09598		1
GROS_g09654		1
GROS_g09663		2
GROS_g09689		1

GROS_g09695	t120b_danre transmembrane protein 120b os=danio rerio gn=tmem120b pe=2 sv=1	1
GROS_g09703		1
GROS_g09730	pdi2_caeel protein disulfide-isomerase 2 os=caenorhabditis elegans gn=pdi-2 pe=1 sv=1	1
GROS_g09735	csca_ecolx sucrose-6-phosphate hydrolase os=escherichia coli gn=csca pe=3 sv=1	1
GROS_g09746	nas6_caeel zinc metalloproteinase nas-6 os=caenorhabditis elegans gn=nas-6 pe=2 sv=2	1
GROS_g09772		1
GROS_g09780		1
GROS_g09781		1
GROS_g09858	sem1a_caeel semaphorin-1a os=caenorhabditis elegans gn=smp-1 pe=2 sv=2	1
GROS_g09868	ileuc_mouse leukocyte elastase inhibitor c os=mus musculus gn=serpinb1c pe=2 sv=1	3
GROS_g09903		2
GROS_g10056		1
GROS_g10075		1
GROS_g10091		1
GROS_g10104		1
GROS_g10174		1
GROS_g10199		1
GROS_g10203		2
GROS_g10209	snf3_caeel sodium- and chloride-dependent betaine transporter os=caenorhabditis elegans gn=snf-3 pe=1 sv=1	1
GROS_g10476		1
GROS_g10491		1
GROS_g10494	echa_human trifunctional enzyme subunit mitochondrial os=homo sapiens gn=hadha pe=1 sv=2	1
GROS_g10613		1
GROS_g10646		1
GROS_g10692		1
GROS_g10717		3
GROS_g10718		1

GROS_g10725		1
GROS_g10726		1
GROS_g10749	sqt1_caeel cuticle collagen sqt-1 os=caenorhabditis elegans gn=sqt-1 pe=3 sv=2	1
GROS_g10832		1
GROS_g10845		2
GROS_g10887		1
GROS_g10946		1
GROS_g10954	chit1_human chitotriosidase-1 os=homo sapiens gn=chit1 pe=1 sv=1	1
GROS_g10977		1
GROS_g11015		1
GROS_g11016		1
GROS_g11019		1
GROS_g11094	tyr1_caeel tyrosinase-like protein tyr-1 os=caenorhabditis elegans gn=tyr-1 pe=3 sv=2	1
GROS_g11113	co145_caebr cuticle collagen 145 os=caenorhabditis briggsae gn=col-145 pe=3 sv=1	1
GROS_g11124		1
GROS_g11198	cup4_caebr acetylcholine receptor-like protein cup-4 os=caenorhabditis briggsae gn=cup-4 pe=3 sv=2	1
GROS_g11219	lip1_caeel lipase os=caenorhabditis elegans gn= pe=1 sv=1	3
GROS_g11291		1
GROS_g11313		1
GROS_g11346	est4a_human carboxylesterase 4a os=homo sapiens gn=ces4a pe=2 sv=2	1
GROS_g11403		1
GROS_g11410	casp1_canfa caspase-1 os=canis familiaris gn=casp1 pe=2 sv=1	1
GROS_g11524	catz_rat cathepsin z os=rattus norvegicus gn=ctsz pe=1 sv=2	1
GROS_g11533		1
GROS_g11540		1
GROS_g11594		2
GROS_g11637		1
GROS_g11715		1
GROS_g11717		1

GROS_g11737	co9a1_mouse collagen alpha-1 chain os=mus musculus gn=col9a1 pe=2 sv=2	2
GROS_g11775		1
GROS_g11776		1
GROS_g11822	dpy5_caeel cuticle collagen dpy-5 os=caenorhabditis elegans gn=dpy-5 pe=3 sv=1	1
GROS_g11876	skpo1_caeel peroxidase skpo-1 os=caenorhabditis elegans gn=skpo-1 pe=2 sv=1	1
GROS_g11885		1
GROS_g11888		1
GROS_g11889		1
GROS_g11906		1
GROS_g11913		1
GROS_g11929		2
GROS_g11934		1
GROS_g11995	brsk2_human serine threonine-protein kinase brsk2 os=homo sapiens gn=brsk2 pe=1 sv=3	1
GROS_g12006		1
GROS_g12069	ppce_pig prolyl endopeptidase os=sus scrofa gn=prep pe=1 sv=1	1
GROS_g12132		2
GROS_g12148	cubn_rat cubilin os=rattus norvegicus gn=cubn pe=1 sv=2	1
GROS_g12281		1
GROS_g12337	lon3_caeel cuticle collagen lon-3 os=caenorhabditis elegans gn=lon-3 pe=1 sv=1	1
GROS_g12505		1
GROS_g12507		1
GROS_g12544		1
GROS_g12557		1
GROS_g12625		1
GROS_g12711		1
GROS_g12752	gcy7_caeel receptor-type guanylate cyclase gcy-7 os=caenorhabditis elegans gn=gcy-7 pe=2 sv=1	1
GROS_g12862		1
GROS_g12892		1
GROS_g12950		1

GROS_g12961	hsp10_caeel heat shock protein hsp- os=caenorhabditis elegans gn=hsp- pe=3 sv=1	1
GROS_g12986		1
GROS_g12993		1
GROS_g13015		1
GROS_g13045		1
GROS_g13138		1
GROS_g13258		1
GROS_g13285	glcm4_caeel glucosylceramidase 4 os=caenorhabditis elegans gn=gba-4 pe=3 sv=2	2
GROS_g13292		2
GROS_g13350	s29a3_bovin equilibrative nucleoside transporter 3 os=bos taurus gn=slc29a3 pe=2 sv=1	1
GROS_g13555	dpm3_caeel probable dolichol-phosphate mannosyltransferase subunit 3 os=caenorhabditis elegans gn=dpm-3 pe=3 sv=1	1
GROS_g13651	ttr46_caeel transthyretin-like protein 46 os=caenorhabditis elegans gn=ttr-46 pe=3 sv=1	2
GROS_g13743		1
GROS_g13767		5
GROS_g13832		1
GROS_g14178		1
GROS_g14179		1
GROS_g14180	ranb9_xenla ran-binding protein 9 os=xenopus laevis gn=ranbp9 pe=2 sv=1	1
GROS_g14202	far1_oncgu fatty-acid and retinol-binding protein 1 os=onchocerca gutturosa gn=far-1 pe=1 sv=2	1
GROS_g14292		1

7.11 Proteins with predicted hypodermis upstream motif also found in surface protein extraction mass spectrometry data

Gene number	Annotation	Number of motifs	Signal peptide
GROS_g00232	nuclease domain-containing protein 1	1	n
GROS_g00261	proteasome subunit beta type-3	1	n
GROS_g00537	probable trans-2-enoyl- reductase mitochondrial	1	n
GROS_g00815	Concanavalin A like galectin	3	n
GROS_g00816	32 kda beta-galactoside-binding lectin	3	n
GROS_g00873	probable elongation factor 1-gamma	1	n
GROS_g01062	cap-gly domain-containing linker protein 1	1	n
GROS_g01153	32 kda beta-galactoside-binding lectin lec-3	1	n
GROS_g01225	pyruvate dehydrogenase e1 component subunit mitochondrial	1	n
GROS_g01251	glutathione synthetase	1	n
GROS_g01337	polyadenylate-binding protein 2	1	n
GROS_g01391	heat shock protein 90	1	y
GROS_g01580	camp-dependent protein kinase regulatory subunit	2	n
GROS_g01765	hypothetical protein	2	y
GROS_g01860	nuclear ribonucleoprotein h	1	n
GROS_g01926	ATP synthase F1 complex epsilon subunit	1	n
GROS_g01996	32 kda beta-galactoside-binding lectin	1	n
GROS_g02187	fatty acid-binding protein homolog 9	2	n
GROS_g02355	hypothetical protein	1	y

GROS_g02490	glutathione peroxidase 3	1	y
GROS_g02810	glutaredoxin-1	1	n
GROS_g02839	60s ribosomal protein l28	1	n
GROS_g02876	sodium potassium-transporting atpase subunit alpha	1	n
GROS_g02877	t-complex protein 1 subunit theta	1	n
GROS_g02901	nadh dehydrogenase	2	n
GROS_g02946	succinate-semialdehyde mitochondrial	1	n
GROS_g02968	transmembrane matrix receptor mup-4	1	y
GROS_g03171	adp-ribosylation factor 1	1	n
GROS_g03318	26s protease regulatory subunit 7	1	n
GROS_g03389	enolase	1	n
GROS_g03390	enolase	1	n
GROS_g03409	proteasome subunit alpha type-5	1	n
GROS_g03558	60s ribosomal protein l17	2	n
GROS_g03730	atp-dependent 6-phosphofructokinase 1	1	n
GROS_g03733	apoptotic chromatin condensation inducer in the nucleus	1	n
GROS_g03987	actin-depolymerizing factor isoform c	1	n
GROS_g04143	hypothetical protein	1	n
GROS_g04188	60s ribosomal protein l38	1	n
GROS_g04192	argonaute protein wago-1	1	n
GROS_g04677	endoglucanase z	1	y
GROS_g05006	40s ribosomal protein s3	1	n
GROS_g05049	protein sidekick-2	1	n
GROS_g05052	probable fumarate mitochondrial	2	n
GROS_g05053	probable fumarate mitochondrial	1	n
GROS_g05336	alanine aminotransferase 2	1	y

GROS_g05438	26s proteasome non-atpase regulatory subunit 14	1	y
GROS_g05537	calponin homolog	1	n
GROS_g05587	pre-mrna-processing factor 19	1	n
GROS_g05729	neuronal cell adhesion molecule	2	n
GROS_g06088	apoptosis-linked gene 2-interacting protein x 1	1	n
GROS_g06143	carbonic anhydrase 1	1	n
GROS_g06257	presequence mitochondrial	1	n
GROS_g06434	levanase	1	n
GROS_g06693	Tissue inhibitor of metalloproteinase	1	y
GROS_g06704	actin-depolymerizing factor isoforms a b	1	n
GROS_g06898	probable malate mitochondrial	1	n
GROS_g07118	ras-related protein rab-7a	1	y
GROS_g07278	co-chaperone protein daf-41	1	n
GROS_g07617	twitchin	1	n
GROS_g07796	mesocentin	1	n
GROS_g07983	60s ribosomal protein l30	1	n
GROS_g08255	actin-interacting protein 1	1	n
GROS_g08685	titin	1	n
GROS_g08707	gut esterase 1	1	n
GROS_g08952	EF-hand domain pair	1	n
GROS_g08969	glutathione synthetase	1	n
GROS_g09082	glycogen synthase kinase-3 beta	1	n
GROS_g09208	glycogen debranching enzyme	1	n
GROS_g09397	eukaryotic translation initiation factor 5a-1	1	n
GROS_g09517	40s ribosomal protein s10b	1	n

GROS_g09623	4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate synthase	1	n
GROS_g09823	ras suppressor protein 1	1	n
GROS_g09828	beta-ureidopropionase	1	n
GROS_g09998	60s acidic ribosomal protein p2	1	n
GROS_g10139	polyadenylate-binding protein 4	1	y
GROS_g10336	hypothetical protein	1	n
GROS_g10401	peptidyl-prolyl cis-trans isomerase 3	1	n
GROS_g10494	trifunctional enzyme subunit mitochondrial	1	y
GROS_g10844	proteasome subunit alpha type-6	1	n
GROS_g11082	integrin-linked protein kinase homolog pat-4	1	n
GROS_g11414	myosin-2	1	n
GROS_g11624	atp synthase subunit mitochondrial	1	n
GROS_g11682	ribulose-phosphate 3-epimerase	1	n
GROS_g11744	nadh dehydrogenase	1	n
GROS_g11801	coronin-like protein cor-1	1	n
GROS_g11841	rab gdp dissociation inhibitor alpha	2	n
GROS_g11879	spliceosome rna helicase ddx39b homolog	1	n
GROS_g12066	cuticle collagen bli-1	1	n
GROS_g12131	myosin regulatory light chain 1	1	n
GROS_g12635	probable atp-citrate synthase	1	n
GROS_g12961	heat shock protein hsp	1	n
GROS_g14202	fatty-acid and retinol-binding protein 1	1	y

7.12 Buffer recipes

10x TBE; add 108 g Tris and 55 g boric acid to 900 mL dH_2O, mix until dissolved. Add 40 mL 0.5 M Na_2EDTA (pH 8.0). Adjust volume to 1 L. Mix well.

1x PBS; add 8 g NaCl, 0.2 g KCl, 1.44 g Na_2HPO_4, 0.24 g KH_2PO_4. Mix well and adjust the pH to 7.4 with concentrated HCl. Adjust to 1 L

2x Sample buffer; 2 mL Tris (1 M, pH 6.8), 4.6 mL 50% glycerol, 1.6 mL SDS solution (10%), 0.4 mL bromophenol blue (0.5%), 0.4 mL β-mercaptoethanol. Store at -80^{o}C.

Cell resuspension buffer for recombinant protein collection; 10 mM imidazole, 50 mM Tris pH 7.5, 300 mM NaCl.

Nickel column wash buffer; 40 mM imidazole, 50 mM Tris pH 7.5, 300 mM NaCl.

Protein storage buffer; 50 mM HEPES (pH 7.8), 0.1 M NaCl, 10% glycerol

TEV cleavage buffer; 150 mM KCl, 20 mM Tris-HCl (pH 7.7), 3 mM $MgCl_2$, 0.5 mM EDTA (pH 8.0), 0.1% TWEEN, 1 mM DTT

M9 buffer; add 3 g KH_2PO_4, 6 g Na_2HPO_4 and 5 g NaCl to 800 mL dH_2O, mix well. Add 1 mL 1 M $MgSO_4$ and adjust to 1 L with dH_2O.

***N. benthamiana* infiltration buffer**; 100 mL dH_2O, 1 ml of 1M MES, 1 mL of 1M $MgCl_2$, 250 µL of 0.1M acetosyringone.

20x SSC; 1 M NaCl, 0.3 M $Na_3C_6H_5O_7$, pH 7.2.

Malaeic acid buffer; 0.1 M maleic acid, 0.15 M NaCl adjusted to pH 7.5 with concentrated NaOH.

Hybridisation buffer; 5 mL deionized formamide, 2 mL 20x SSC, 1 mL 10% Boehringer blocking reagent in maleic acid buffer, 1 mL 20% SDS, 100 µL 100x Denhardt's solution, 100 µL 0.1 M EDTA pH 8.0, 200 µL fish (salmon) sperm DNA (10 mg/ml), 62.5 µL yeast tRNA (500 U/mL), DEPC treated H_2O to 10 mL.