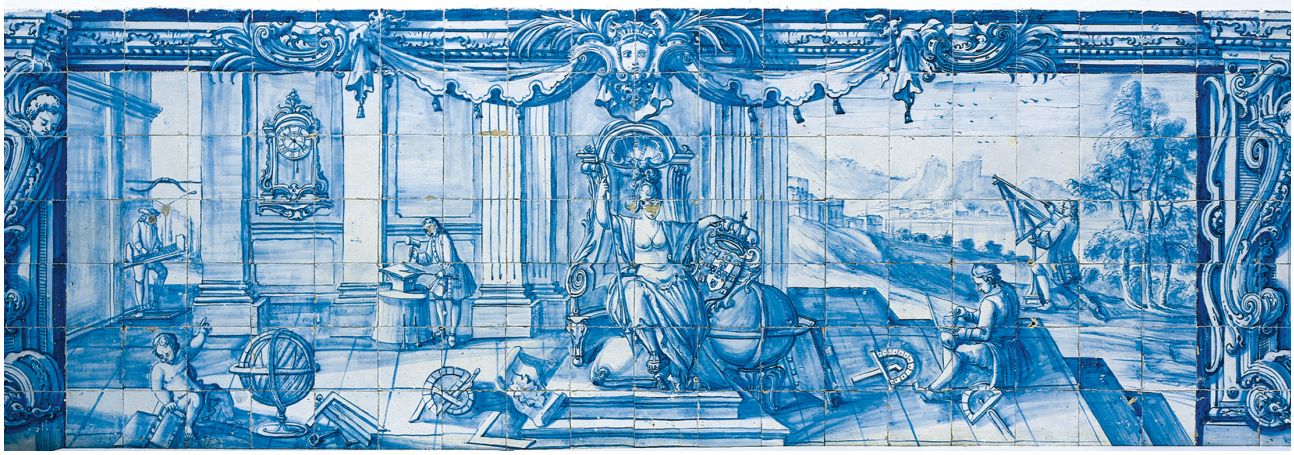
FUNCTIONAL CHARACTERIZATION OF PUTATIVE EFFECTORS IN PINEWOOD NEMATODE, *BURSAPHELENCHUS XYLOPHILUS*

Maria Margarida Saial Santos Guiomar Espada

Tese apresentada à Universidade de Évora
para obtenção do Grau de Doutor em Biologia

ORIENTADORES: *Professor Manuel Galvão de Melo e Mota*
Professor John T. Jones

ÉVORA, MARÇO 2018



INSTITUTO DE INVESTIGAÇÃO E FORMAÇÃO AVANÇADA

FUNCTIONAL CHARACTERIZATION OF PUTATIVE EFFECTORS IN PINEWOOD NEMATODE, *BURSAPHELENCHUS XYLOPHILUS*

Maria Margarida Saial Santos Guiomar Espada

Tese apresentada à Universidade de Évora
para obtenção do Grau de Doutor em Biologia

ORIENTADORES: *Professor Manuel Galvão de Melo e Mota*
Professor John T. Jones

ÉVORA, MARÇO 2018

To my daughter Luísa

ACKNOWLEGMENTS

First and foremost, I wish to express my sincere gratitude to my supervisors, Manuel Mota and John Jones that supported me during the past four years, for all your help, guidance, enthusiasm and expertise. I am deeply grateful for the opportunity of being part of both teams, that allowed me to have a very good scientific training and experience.

I would like to thank Fundação para a Ciência e a Tecnologia (Portugal) for the financial support through the PhD scholarship (SFRH/BD/84541/2012).

I would like to thank Sebastian Eves-van den Akker for all his support in bioinformatics during this work. It was a pleasure to work with you and thank you for pushing me to go further.

I would like to thank to my colleagues and friends at NemaLab and University of Évora, Cláudia Vicente, Patrick Materatski, Clarisse Brígido, Paulo Vieira, Pedro Barbosa, Patrícia Avó, Ana Cláudia Silva and Jordana Branco, Francisco Nascimento, Carlos Gutiérrez e Inácia Ferreira. Thank you for your scientific support, encouragement, friendship, and good times throughout this journey.

I would like to thank everyone at nematology lab at The James Hutton Institute and friends in Dundee, Daniel, Gonçalo, Mei, Camille, Elena and Sophie. Thank you so much for the support and good times in Scotland.

I would like to thank everyone at The Baum lab in Iowa State University and Taisei Kikuchi group, at University of Miyazaki, who I have worked with.

I would like to thank to my closest friends and family, that encouraged and helped me over this journey. To P. Miguel, thank you for your understanding, support and continuous encouragement.

The present work was financially supported by FCT – Fundação para a Ciência e a Tecnologia, IP– by the grant SFRH/BD/84541/2012, supported by POPH/FSE through the “Programa Operacional Humano” and the “Fundo Social Europeu”.

THESIS AND OTHER SCIENTIFIC PUBLICATIONS

The present work resume the following publications:

Espada, M., Jones, J.T. and Mota, M. (2016) Characterization of glutathione S-transferases from the pinewood nematode, *Bursaphelenchus xylophilus*. *Nematology* 18(6), 697-709.

Espada, **M.**, Silva, A.C., Eves-van den Akker, S., Cock,P.J.A., Mota, M. and Jones, J.T. (2016) Identification and characterization of parasitism genes from the pinewood nematode *Bursaphelenchus xylophilus* reveals a multi-layered detoxification strategy. *Molecular Plant Pathology* 17, 286-295.

Tsai, I.J., Tanaka, R., Kanzaki, N., Akiba, M., Yokoi, T., **Espada, M.**, Jones, J.T. and Kikuchi, T. (2016) Transcriptional and morphological changes in the transition from mycetophagous to phytophagous phase in the plant parasitic nematode *Bursaphelenchus xylophilus*. *Molecular Plant Pathology* 17, 77-83.

Other publications:

Nascimento, F.X., **Espada, M**., Barbosa, P., Rossi, M.J., Vicente, C.S. and Mota, M. (2016) Non-specific transient mutualism between the plant parasitic nematode, *Bursaphelenchus xylophilus*, and the opportunistic bacterium *Serratia quinivorans* BXF1, a plant-growth promoting pine endophyte with antagonistic effects. *Environmental Microbiology* 18(12), 5265-5276.

Valadas, V., **Espada, M**., Nobre, T., Mota, M. and Arnholdt-Schmitt, B. (2015) AOX in parasitic nematodes: a matter of lifestyle? In Alternative Respiratory Pathways in Higher Plants (Gupta,K.J., Mur, L.A.J. and Neelwarne, B., eds), pp. 315-318. Wiley-Blackwell.

Mallez, S., Castagnone, C., **Espada, M**., Vieira, P., Eisenback, J.D., Harrell, M., Mota, M., Aikawa, T., Akiba, M., Kosaka, H., Castagnone-Sereno, P. and Guillemaud, T. (2015) Worldwide invasion routes of the pinewood nematode: What can we infer from population genetics analyses? *Biological Invasions* 17, 1199-1213.

Vieira, P., Castagnone, C., Mallez, S., **Espada**, **M.**, Navas, A., Mota, M. and Castagnone-Sereno, P. (2014) Sequence variability of the MspI satellite DNA family of the pinewood nematode *Bursaphelenchus xylophilus* at different geographic scales. *Molecular Phylogenetics and Evolution* 70, 120-129.

Vicente, C.S.L., Nascimento, F. X., **Espada, M**., Barbosa, P., Hasegawa, K., Mota, M. and Oliveira, S. (2013) Characterization of bacterial communities associated with the pine sawyer beetle *Monochamus galloprovincialis*, the insect vector of the pinewood nematode *Bursaphelenchus xylophilus*. *FEMS Microbiology Letters*, 347, 130–139.

Mallez S., Castagnone C., **Espada M**., Vieira P., Eisenback J.D., *et al*. (2013) First insights into the genetic diversity of the pinewood nematode in its native area using new polymorphic microsatellite loci. *PLoS ONE* 8(3), e59165.

Nascimento, F.X., Vicente, C.L.S., Barbosa, P., **Espada, M**., Glick, B., Mota, M. and Oliveira, S. (2013) Evidence for the involvement of ACC deaminase from *Pseudomonas putida* UW4 in the biocontrol of pine wilt disease caused by *Bursaphelenchus xylophilus*. *BioControl,* 58(3), 427-433.

ABSTRACT

Plant-parasitic nematodes cause huge economic and ecological losses in agriculture and forestry ecosystems worldwide. The migratory endoparasitic nematode *Bursaphelenchus xylophilus* is unique in having phytoparasitic and mycetophagous phases in its life cycle. During the phytoparasitic stage, the nematode migrates within pine trees feeding on the contents of cortex and xylem parenchymal cells. Interactions of the nematode with the plant host are mediated by effectors - secreted proteins originating from the pharyngeal gland cells. The main objective of this work was to identify and characterise nematode effector genes that play key roles in parasitism.

Next-generation sequencing and bioinformatic analysis were used as a neutral approach to identify novel effectors. The transcriptome of *B. xylophilus* has been sequenced to compare gene expression in the mycetophagous and plant-parasitic stages to identify genes and morphological adaptions involved in plant parasitism. Additionally, transcripts from pharyngeal gland cells were sequenced. Analysis of the data revealed new parasitism-related proteins. A promoter DNA sequence motif was identified that is associated with expression in the pharyngeal gland cells, and was used to predict further effector sequences. A panel of 118 predicted effector genes with a signal peptide, at least one occurrence of the motif and that are upregulated *in planta* were identified.

Functional data suggest that effectors are one key part of a multi-layered detoxification strategy deployed by *B. xylophilus* in order to protect itself from host defence responses. *B. xylophilus* secretes detoxification enzymes into the host, while simultaneously upregulating other detoxification enzymes within its digestive system. We showed that one of these enzymes – a glutathione S-transferase - has a protective activity against defence compounds produced by the host. These data represent the most comprehensive analysis of novel effectors from this nematode to date.

Caracterização funcional de putativos *effectors* do nemátode da madeira do pinheiro, *Bursaphelenchus xylophilus*

RESUMO

Os nemátodes fitoparasitas representam enormes perdas económicas e ecológicas na Agricultura e em sistemas florestais. O nemátode endoparasita migratório *Bursaphelenchus xylophilus* tem a característica única de ser fitófago e micófago, no seu ciclo de vida. Durante a fase parasítica, o nemátode migra dentro dos pinheiros, alimentando-se dos conteúdos das células do parenquima presentes no córtex e xilema. As interações do nemátode com a planta hospedeira são mediadas por *effectors* – proteínas secretadas com origem nas glândulas esofágicas. O objectivo principal deste trabalho foi a identificação e caracterização dos *effectors* do nemátode que têm um papel preponderante no parasitismo.

Sequenciação de última geração e análises bioinformáticas foram usadas como uma abordagem neutra para identificar novos *effectors*. O transcriptoma do *B. xylophilus* foi sequenciado e a expressão dos genes foi comparada entre as fases micófaga e parasítica, para identificar os genes e adaptações morfológicas envolvidas no parasitismo. Adicionalmente, foram sequenciados os transcritos originários das glândulas esofágicas. A análise dos dados revelou novas proteínas de parasitismo. Foi identificado um *motif* de DNA no promotor, cuja sequência está associada à expressão nas glândulas esofágicas e que pode ser utilizado para predizer novos *effectors*. Foi identificado um conjunto de 118 novos genes *effectors* com sinal peptídeo, com pelo menos uma ocorrência do *motif* na região promotora e altamente expressos na planta.

Os dados funcionais sugerem que os *effectors* são parte importante na estratégica de destoxificação, a diferentes níveis, estabelecida pelo nemátode para se proteger das respostas de defesa do hospedeiro. *B. xylophilus* secreta enzimas de destoxificação dentro do hospedeiro, enquanto simultaneamente expressa outras destas enzimas no seu sistema digestivo. Aqui demonstramos que uma destas enzimas - glutationa S-transferase – tem uma actividade protectora contra compostos de defesa produzidos pelo hospedeiro. Estes dados representam a análise mais completa de novos *effectors* deste nemátode encontrados até ao presente.

LIST OF ABBREVIATIONS

ANOVA	Analysis of variance
BLAST	Basic local alignment search tool
bp	Base pair
BUX	Bursaphelenchus xylophilus
cDNA	Complement dna
CYP	Cytochrome
CWDE	Cell wall degrading enzyme
DE	Differentially expressed
DG	Dorsal gland cell
DIG	Digoxigenin
DPI	Days post infection
DNA	Deoxyribonucleic acid
DOG	Dorsal gland cell motif
ENA	European nucleotide number
ETI	Effector triggered immunity
FAR	Fatty acid and retinol-binding protein
FDR	False discovery rate
FF	Fungal feeder
FMO	Flavin monooxygenase
FMRF	Phe-Met-Arg-Phe peptides (neuropeptide)
FPKM	Fragments per kilobase per million
GC	Gland cell
GO	Gene ontology
GST	Glutathione S-transferase
GH	Glycoside hydrolase
HYP	Hypervariable
HGT	Horizontal gene transfer
IPTG	Isopropyl β-D-1-thiogalactopyranoside
JP	Japanese
mRNA	Messenger RNA
NR	Non-redundant
OD	Absorbance
PAMPs	Pathogen associated molecular patterns
PCN	Potato cyst nematode
PCR	Polymerase chain reaction
PDA	Potato dextrose agar
PEL	Pectate lyase
PPN	Plant parasitic nematode

PT	Portuguese
PTI	Pattern triggered immunity
PWD	Pine wilt disease
PWN	Pinewood nematode
qPCR	Quantitative PCR
Rcr3	Papain-like cysteine protease from tomato
RKN	Root knot nematode
RNA	Ribonucleic acid
RNAi	RNA interference
RNAseq	RNA sequencing
ROS	Reactive oxygen species
RPKM	Reads per kilobase per million
rRNA	Ribosomal RNA
RT-PCR	Reverse transcriptase-pcr
SIGNALP	Signal peptide server for prediction of presence and location of SP
SP	Signal peptide
SPRYSEC	Secreted SPRY-domain containing protein
SNP	Single nucleotide polymorphism
SVG	Subventral gland cell
TMHMM	Prediction of transmembrane helices in proteins
TTR	Transthyretin-like protein
UDP	Uridine 5'-diphospho-glucuronosyltransferase
UGT	UDP-glucuronosyltransferase
VAP	Venom allergen protein

TABLE OF CONTENTS

CHAPTER I

GENERAL INTRODUCTION

1. NEMATODES

Nematodes belong to the Phylum Nematoda and are roundworms with a non-segmented, elongated body contained within a resistant cuticle. Nematodes are widely dispersed throughout the world and are thought to represent up to four in five animals on the planet. The Nematoda is one of the largest animal Phylum in terms of the number of described species, with more than 25000 species catalogued to date (Williamson and Kumar, 2006; Perry and Moens, 2011; Blaxter and Koutsovoulos, 2015). Nematodes are present in every ecological niche, from marine to soil environments and can become adapted to stressful conditions (such as water or oxygen stress) (Bongers and Ferris, 1999; Perry and Moens, 2011). Most nematodes are free-living and some of these are used as useful bioindicators in environmental monitoring (Bongers and Ferris, 1999). One free living species, *Caenorhabditis elegans,* has been widely used as a model organism for genetics and developmental biology and is now the most intensively studied and best understood organism on the planet (wormbook.org, Corsi *et al.*, 2015). However, many nematodes have evolved to become successful parasites of both animals and plants and have severe effects on mankind (Perry and Moens, 2011; Blaxter and Koutsovoulos, 2015). Up to one quarter of the world's population is thought to be infected with a parasitic nematode infection and some of these, such as *Brugia malayi* (the infective agent of filariasis) and *Ascaris* cause debilitating disease, particularly in developing nations.

1.1 Plant-parasitic nematodes

1.1.1 Ecological and economical importance

More than 4100 species of plant-parasitic nematodes (PPN) have been described which attack a wide range of plant hosts across the world causing enormous economic and ecological damage. Annual crop and forestry losses due to PPN are estimated at more than $US80 billion per year (Jones *et al*., 2013). All major crop species are impacted by PPN with many having more than one plant-parasitic species that can affect their production. Some PPN species are quarantine organisms and the trade of plants, seeds and/or plant-related products are subject to strict controls in order to prevent further spread of the pathogens. The impact of PPN is severely underestimated in developing countries as these organisms are small, frequently soil dwelling and cause non-specific symptoms (Williamson and Kumar, 2006; Perry and Moens, 2011; Jones *et al*., 2013).

1.1.2 Adaptations for parasitism

Although several different classifications have been proposed for the Nematode, Van Megen *et al* (2009) subdivide the Phylum into 12 clades (Figure 1). Four of these include plant-parasitic nematodes: clade 1 (Triplonchida), clade 2 (Dorylaimida), clade 10 (Aphelenchoididae) and clade 12 (Tylenchida) (Jones *et al.,* 2013; Kikuchi *et al*., 2017), indicating that like parasitism of

animals, the ability to infect plants has evolved independently in Nematodes on multiple occasions (Blaxter and Koutsovoulos, 2015; Bird *et al*., 2015). Plant-parasitic species have co-evolved with plants and have developed a huge diversity of different parasitic lifestyles with a variety of methods feeding, surviving and reproducing.

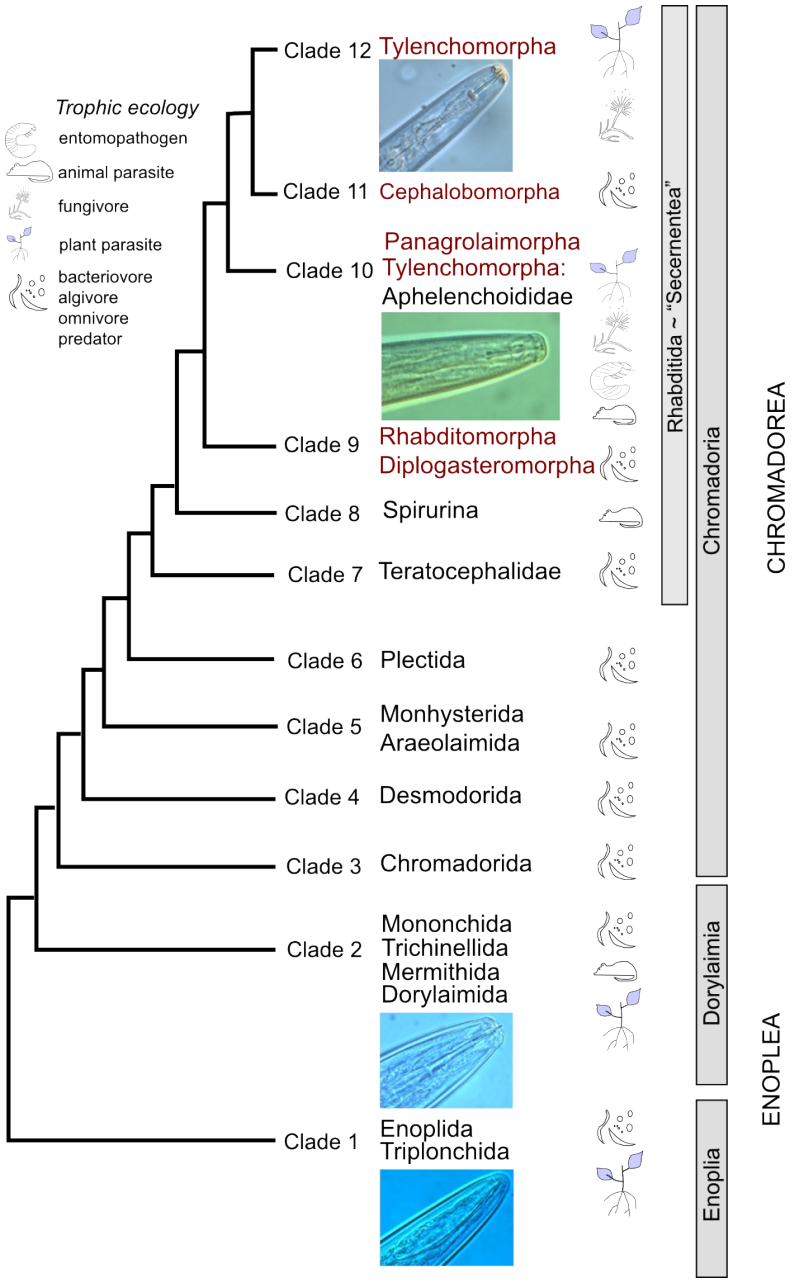


Figure 1- Phylogeny of the Phylum Nematoda. Van Megen (Van Megen *et al*., 2009) classified the nematodes into twelve clades, where four of the clades contain plant-parasitic nematodes. The symbols represent the trophic ecology of each group. Plant-parasitic nematodes are represented by a plant symbol (as indicated in the legend, top left) and have different stylet forms, represented by the small images in the figure. (From Jones *et al*., 2013).

The emergence of genomics and transcriptomics has allowed insight into the different genomic features that are associated with parasitism (Blaxter and Koutsovoulos, 2015). At the time of writing, one hundred nematode genomes have been sequenced, most of which are parasites (described in WormbaseParasite; Howe *et al*., 2017). The genomes of parasitic nematodes sequenced to date, other than that of *Globodera pallida*, are smaller than that of the free-living *C. elegans*, the first nematode genome to be sequenced (*C. elegans* consortium, 1998). This reflects a reduction in the number of genes present in parasitic nematodes compared to the free-living species Kikuchi *et al*., (2017). Gene regulation mechanisms are present that influence the gene expression allowing the nematode to adapt to several different environments and hosts (Kikuchi *et al*., 2017). Comparative genomics have allowed the discovery of new proteins in some parasitic nematodes acquired from other microorganisms such as bacteria and fungi by horizontal transfer (HGT) events. These include the cell wall degrading enzymes in PPN (Blaxter and Koutsovoulos, 2015).

Although the parasitism has evolved independently in the various clades, the presence of a stylet, or its functional equivalent (Figure 1, 2), is common all PPNs. This feeding apparatus has a dual role in parasitism: it is used to disrupt the cell wall during feeding and migration and is used to deliver some of the parasitism proteins – effectors – into the host (Bird *et al*., 2015). There are three different types of stylets in PPNs (Figure 1): The stomatostylet is present in clades 10 and 12; the odontostyle is present in clade 2 and the onchiostyle is present in clade 1 (Kikuchi *et al*., 2017). The stylet is connected to the pharyngeal gland cells (GC) – two subventral and one dorsal – which produce proteins important in several aspects of the parasitic process. The size of the GCs changes throughout the life cycle – with the dorsal GC increasing in size in later stages (Haegeman *et al*., 2012; Carletti *et al*., 2013). These structures are of major importance in parasitism and are obvious targets of study on PPNs (Figure 2).

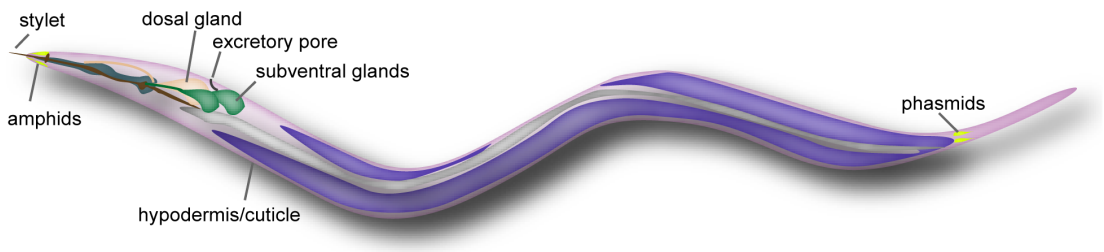


Figure 2 – Schematic representation of a plant-parasitic nematode. The secretory organs are identified in the nematode body. (From Shinya *et al*., 2013).

1.1.2.1 Parasitic lifestyle strategies

Nematodes have a range of different parasitism lifestyles and can be biotrophic or necrotrophic pathogens. The biotrophic PPNs are among the most important ones in agriculture and cause the most severe damage to crops. These nematodes include root-knot (RKN) and cyst nematodes (CN) (genus *Meloidogyne spp*., *Globodera spp. and Heterodera spp.*), infect the roots of a wide range of hosts, and manipulate the host to form complex feeding structures (Jones *et al*., 2013). Once the feeding structure is formed the nematodes change their body configuration and became sedentary. The feeding structures are kept alive for the duration of the nematode life cycle. By contrast, necrotrophic PPNs migrate and reproduce on the plant causing death of the host as they feed. Examples of these nematodes include *Bursaphelenchus xylophilus, Pratylenchus spp*., *Trichodorus spp*., *Ditylenchus spp*., and *Radophulus spp*. (Perry and Moens, 2011).

Plant-parasitic nematodes can be classified as ectoparasites or endoparasites (Figure 3). Endoparasites include migratory and sedentary nematodes, depending on their capacity to migrate inside of the host or form feeding structures, as described above. Ectoparasites, including Longidorid species like the dagger and needle nematodes feed on the roots of the plant, but never enter the host.

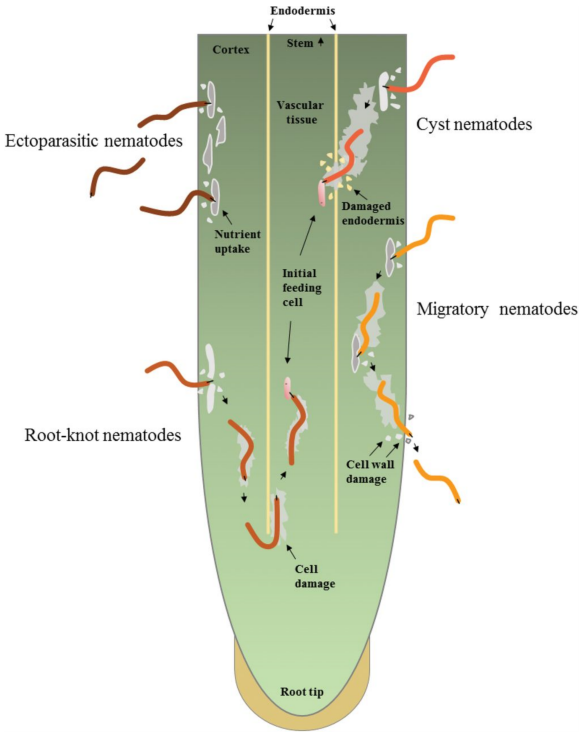


Figure 3 – Illustration on the different parasitism life styles of the plant-parasitic nematodes in the root system: endo and ectoparasites. Cyst and root-knot nematodes are sedentary and form feeding sites inside the roots. Their juveniles forms live in the soil before entering the root system. Migratory nematodes can enter the host and feed on the cortex cells (endoparasites) or just feed from the cell contents without entering the root (ectoparasites). (From: Holbein *et al*., 2016)

Although most PPN infest root systems, species belonging to Aphelenchoididae and Anguinidae can infect the upper part of the plant. This study will focus in one of these nematodes, the migratory endoparasitic *B. xylophilus*, which has one of the most complex life cycles of the migratory parasitic nematodes (Moens and Perry, 2009).

1.2. The pinewood nematode, *Bursaphelenchus xylophilus*

The pinewood nematode (PWN), *B. xylophilus* (Steiner & Buhrer) Nickle (1934) (Figure 4) is a migratory endoparasitic nematode that belongs to the genus *Bursaphelenchus,* previously established by Fuchs in 1937 and is associated with conifer trees, especially those from the *Pinus* genus. Most *Bursaphelenchus* species are mycophagous and feed on fungi that colonise dead plant tissue. These nematodes are usually vectored by an insect species and are therefore designated as phoretic nematodes. Only two species are parasitic – *B. xylophilus* and *B. cocophilus* – the pinewood nematode and the red ring nematode, respectively. (Kanzaki, 2008; Vicente *et al*., 2012; Futai, 2013). The taxonomy classification of the genus *Bursaphelenchus* is the following (Kanzaki, 2008):

Kingdom Animalia
Phylum Nematoda
Class Secernentea,
Order Aphelenchida
Suborder Aphelenchina
Superfamily Aphelenchoidoidea
Family Parasitaphelenchidae,
Subfamily Bursaphelenchinae
Genus *Bursaphelenchus*

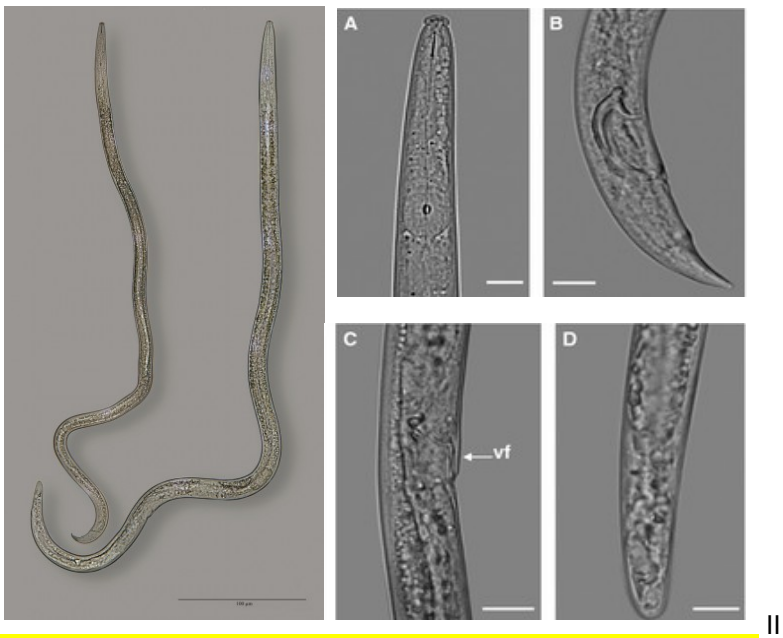


Figure 4 – The pinewood nematode, *Bursaphelenchus xylophilus*. I) Male (on the left) and female (on the right) (courtesy, JD Eisenback). II) Some morphological features that can distinguish this nematode from it close-related species in the genus *Bursaphelenchus*. A- anterior region with wide medium bulb; B- male tail detail with characteristic spicule; C- vulval region (in detail the vulva flap, vf); D- female round tail (From Inácio *et al*., 2015).

1.2.1 Distribution, ecological and economical importance worldwide

Bursaphelenchus xylophilus (Figure 4) is native to North America. Native pine species have evolved tolerance or resistance to *B. xylophilus* and it is therefore not considered as a pathogenic species in this region. At the beginning of the twentieth century the nematode was introduced into the Far East, in Japan, and during the following decades spread out to China (in 1982), Taiwan (1985) and Korea (1988) (Figure 5) (reviewed in Vicente *et al*., 2012). In these regions the trees present have had no previous exposure to the nematode and are extremely susceptible to damage. However, it was only in the 1960s that the nematode was discovered to be the causal agent of the wilting problems observed in conifer forests – Pine wilt disease (PWD) (Kanzaki, 2008; Futai, 2013). *Bursaphelenchus xylophilus* is now recognised as one of the most damaging threats to conifer forests worldwide (Mamiya, 1983; Webster and Mota, 2008; Vicente *et al*., 2012). Over 600 thousand cubic metres of forest timber are destroyed with a value estimated at 10 million US Dollars each year (Jones *et al*., 2008). Damage is caused to forests worldwide at both economic and ecological levels.

In Europe, where forests cover 44% of the European area (including Russian Federation), some conifers species are highly susceptible to PWN, specifically *Pinus pinaster* and *P. nigra* in central and southern areas and *P. sylvestris* in Northern area, which are important forestry species, for woody goods. In the affected regions, the infected trees can die less than one year after infection (depending on the environmental conditions) and the damage caused is extensive. These regions may therefore take decades to recover (Vicente *et al*., 2012). PWN was first reported in Europe in 1999, after its detection in Portugal (Setúbal Peninsula) (Mota *et al*., 1999) in *P. pinaster* trees (Figure 5) and more recently in *P. nigra* (Inácio *et al*., 2015). Over the next decade, the nematode spread to the centre of the country and Madeira island (Fonseca *et al*., 2012). More recently, the nematode was detected in Spain (Robertson *et al*., 2011). The introduction and spread of this nematode has a direct impact on forest natural resources and wood industry, and an indirect effect on the local economy due to European restrictions in the circulation of wood products from affected areas. (EPPO; Webster and Mota, 2008; Vicente *et al*., 2012; Rodrigues *et al*., 2015).

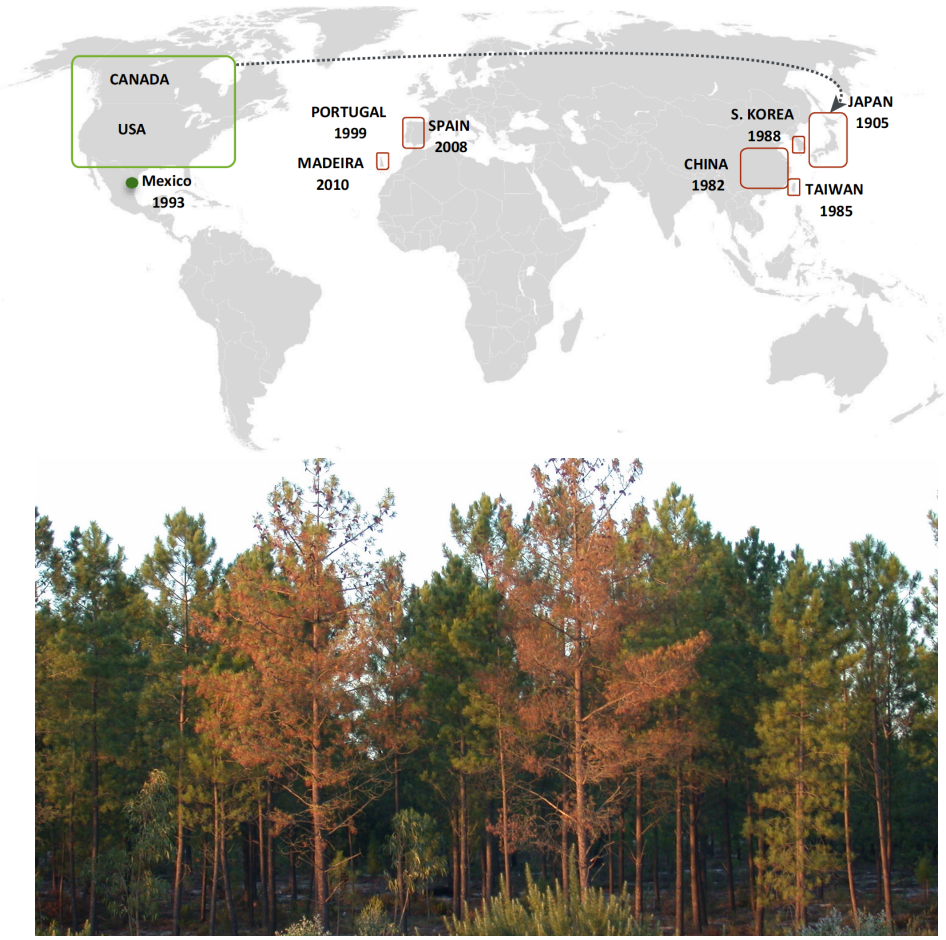


Figure 5 – Pine wilt disease dispersal worldwide (top) and disease symptoms on maritime pine trees (*P. pinaster*) in Setúbal Peninsula, 30km SE of Lisbon (Portugal) (bottom). Wilting of the trees causes the browning/reddening of the needles is the most characteristic symptom of the disease caused by the pinewood nematode, PWN, *Bursaphelenchus xylophilus*. (Courtesy, M. Mota) (Map adapted from Zhao *et al*., 2014).

1.2.2 Biology – life cycle

The PWN life cycle has several unique features including the ability to feed on plant cells and endophytic fungi (Jones *et al*., 2013) and the fact that its life cycle is one that involves a vector to spread the species (Figure 6). The nematode feeds and reproduces in the stems of the tree but has not been detected in the needles of the tree or in the soil beneath the tree. The nematode is transmitted to a new host when an infected adult insect bark beetle (*Monochamus species*) feeds on the top of a healthy pine tree. The nematode enters through the feeding wounds directly into the cortex. Once inside, all propagative stages of the nematode are worm-like, mobile and parasitic, except for the egg stage and are found within host tissues (Figure 7) (Mamiya, 2012; Vicente *et al*., 2012; Futai, 2013).

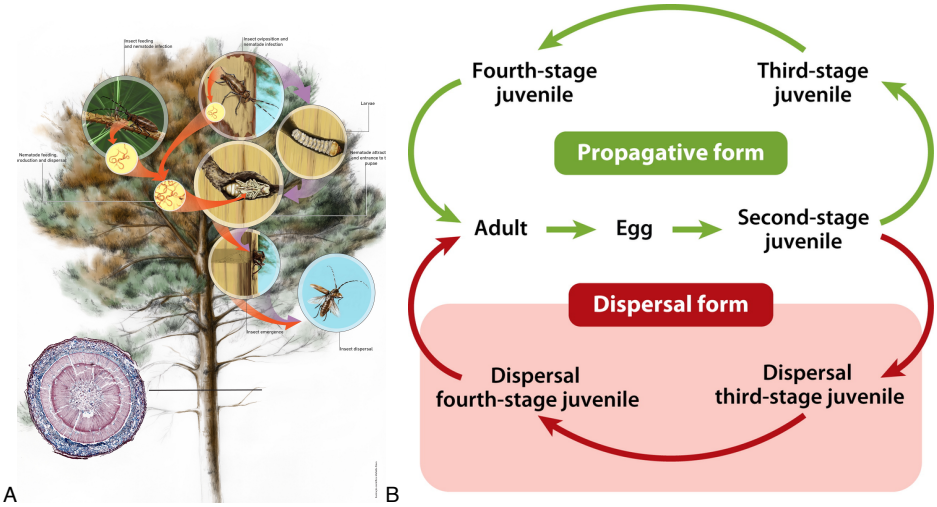


Figure 6 – The life cycle of *Bursaphelenchus xylophilus*. A and B) The dispersal and propagative forms of the life cycle. A is also a PWD representation. (From Nascimento *et al.*, 2015 (A) and Futai, 2013 (B)).

The nematode feeds on the parenchyma cells of the ray canals and resin ducts in the cortex and in the xylem (Figures 6, 7). The nematode uses the tracheids, which are non-living cell canals, to migrate inside the tree and continues reproducing. At a later stage of infection, the nematode population increases and the nematodes are detected in cortex, phloem, cambium and xylem where they destroy the parenchyma cells of the radial and axial resin canals. At this point, there is disruption of the water flow in the plant, blocking of the vascular system (embolisms) and cavitation leading to wilting symptoms and tree death (Mamiya, 2012; Futai, 2013). Once the tree host is dying, the population of endophytic fungi increases and the nematode switches to its fungal feeding strategy. As the abundance of fungi declines, the nematode enters its survival stage – a specialized survival and dispersal Dauer stage that is transmitted by insects to other hosts. The Dauer stage is attracted to the pupal chambers of beetles and attaches itself to the adult beetle as it emerges. When the adult insects emerge from the bark sufficient nematodes are carried in its tracheids to infect a new healthy host (Moens and Perry, 2009; Jones *et al.*, 2008; Vicente *et al.*, 2012).

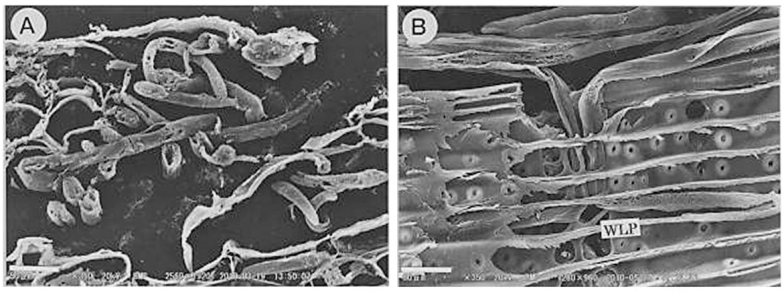


Figure 7 – Localisation of the pinewood nematode in the Japanese red pine tissues (*Pinus densiflora* Siebold & Zucc.) (scanning electron micrographs of the pine). A: disrupted parenchyma cells of the cortex after 24 hours post infection in the pine shoots (5-month old). B: nematode in axial resin and ray canals. (From Mamiya, 2012)

1.2.3 The genome of *Bursaphelenchus xylophilus*

All *Bursaphelenchus* genus species investigated to date have approximately the same genome size (Kikuchi *et al.,* 2017). The full genome sequence of *B. xylophilus* (Kikuchi *et al*., 2011) predicts more than 18074 protein coding genes, has the same karyotype as *C. elegans* and the highest G+C content of all PPNs examined to date (Kikuchi *et al*., 2017). The first version of the genome is annotated and available online at GeneDB (http://www.genedb.org/Homepage/Bxylophilus; Logan-Kumpler *et al*., 2012). Analysis of the genome revealed the largest number of digestive proteases among all PPNs. In addition, PWN has expanded families of lysosome and cytochrome P450 pathway-related enzymes, which may reflect the parasitic life style and the environment in which the nematode lives. A family of glycosyl hydrolase (GH) 45, beta-1,4-endoglucanases, acquired by horizontal gene transfer from fungi is also present and is unique among PPNs (Kikuchi *et al*., 2004; Kikuchi *et al*., 2011). A more recent intra-species diversity study, on six Japanese *B. xylophilus* strains, revealed a high level of genetic diversity with 4% of the genome having genetic variations such as single nucleotide polymorphisms (SNPs) or indels (insertions/deletions). These variations may be associated with the multiples introductions of this species into new geographic areas (invasive species) or may reflect the pathogenic or ecological traits of the nematode (Palomares-Rius *et al*., 2015).

2. NEMATODE EFFECTORS

2.1 Plant-nematode interactions

2.1.1 Definitions of effectors

The interactions of plant pathogens, including plant-parasitic nematodes, with their hosts are mediated by effectors: secreted proteins produced by the pathogen that modify the host to the benefit of the pathogen. Effectors are proteins that have key roles in parasitism. Effectors have also been given more narrow definitions and Bird *et al.*, (2015) defined them as “any pathogen molecules that suppress host defences or manipulate the host to allow provision of food to the pathogen”. Many of the effectors of plant-parasitic nematodes are produced in a set of GC from where they are secreted into the host through the stylet (Haegeman *et al*., 2012; Mantelin *et al*., 2017). However, other nematode secretory tissues can also be a source of effectors. For example, the HYP (hyper-variable apoplastic) effectors of potato cyst nematodes are secreted into the apoplast from the amphids (Eves-van den Akker *et al*., 2014a) and a series of antioxidant proteins are secreted to the parasite surface from the hypodermis (*e.g.* Robertson *et al.,* 2000). The study of effector biology has been revolutionized by the availability of genome and transcriptome sequences of plant-parasitic nematodes. Effectors from plant-parasitic nematodes have been identified that degrade the host cell wall to allow infection (*e.g.*

Haegeman *et al*., 2012*)*, that suppress defence responses of the host (*e.g.,* Jaouannet *et al*., 2012) and that protect the pathogen from host derived toxins (*e.g.* Robsertson *et al*., 2000) (Figure 8). However, in spite of this progress there are still major gaps in our knowledge of plant-nematode interactions. Much of the progress described in the literature relates to the interactions of the sedentary endoparasitic nematodes with their hosts (Gheysen and Mitchum, 2011; Haegeman *et al*., 2012; Quentin *et al*., 2013; Mantelin *et al*., 2017). By contrast, much less is known about the processes targeted by migratory endoparasitic nematodes in general, and PWN in particular.

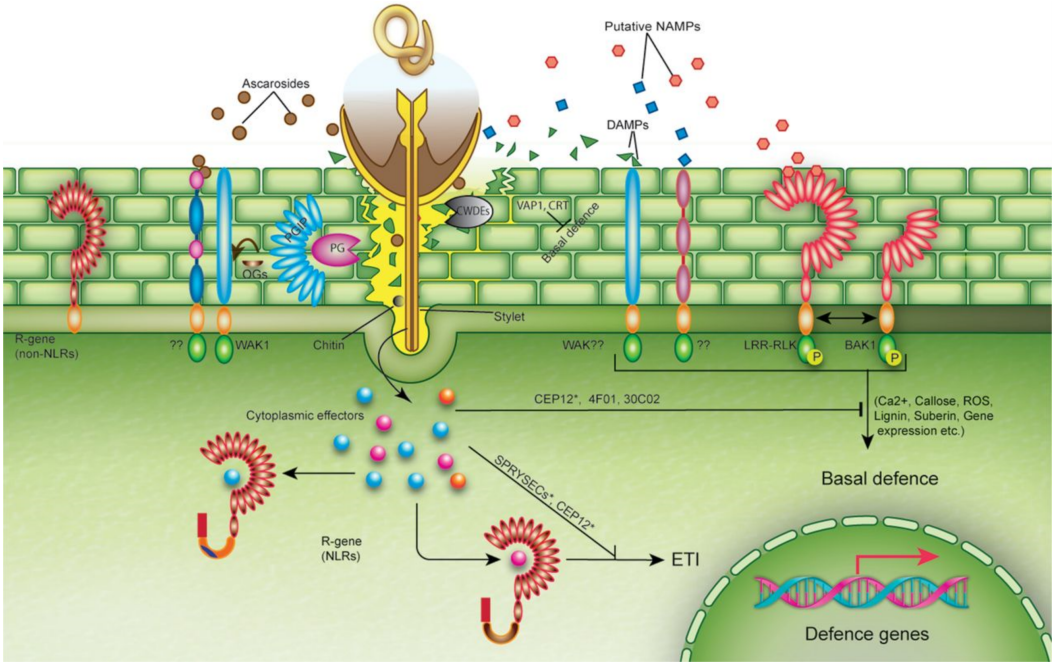


Figure 8 – Plant-parasitic nematodes interaction with the plant hosts. Representation of the different nematode effectors involved in parasitism and its localization in host cell. (From Holbein *et al*., 2016).

2.1.2 Invasion and migration of the host tissues

The first barrier that the plant-parasitic nematode faces is the plant cell wall (Figure 8). To be a successful parasite, the nematode needs to break down this complex and well-organized structure formed mainly from pectin, hemicellulose and cellulose, but which also contains other compounds (polymers) including arabinose, xylan and various proteins. The degradation of the plant cell wall is achieved using a combination of mechanical and enzymatic actions of the stylet and various cell wall degrading enzymes. The first enzyme, and therefore first effector to be reported in PPNs (Smant *et al*., 1998), was a beta-1,4-endoglucanase from CN. Since then, many other enzymes and cell wall modifying proteins have been described in different PPNs including other cellulases, pectate lyases (Popeijus *et al*., 2000; Huang *et al*., 2003; Vanholme *et al*., 2007; Wieczorek *et al*., 2014) that degrade pectin and arabinases that hydrolise pectin, expansins and xylanases (Mitreva-Dautova *et al*., 2006). Polygalacturonase which hydrolyses xylan, has been described from several different nematodes (Jaubert *et al*., 2002). Expansins,

which can disrupt non covalent bonds between cellulose microfibilis were first described in *G. rostochiensis* and represented the first report of these proteins in the animal kingdom (Qin *et al*., 2004). Each of these cell wall modifying proteins is thought to have been acquired by horizontal gene transfer from bacteria (reviewed by Haegeman *et al.,* 2011a). Interestingly, the cellulase present in *B. xylophilus* is from GH45 and has sequence similarity with sequences described in fungi (Kikuchi *et al*., 2004). This is in contrast with the GH family 5 cellulases present in other PPN and suggests that an independent HGT event has occurred within the *Bursaphelenchus* clade.

2.1.3 Protection of the nematode against plant defences

Once the nematode is inside the host, the environment can be very unfavorable. In order for the nematode to survive it needs to protect against host defence responses, and in the case of biotrophic nematodes, it will need to protect the feeding structure (Smant and Jones, 2011). In response to infection, the plant releases compounds including terpenoids, isoflavonoids and reactive oxygen species (ROS) that need to be neutralized by the nematode (Smant and Jones, 2011). A variety of different antioxidant or detoxification proteins are known to be secreted into the host, including a Glutathione S- transferase (GST) by *M. incognita* (Dubreuil *et al*., 2007). Several enzymes that can metabolise ROS have been identified on the surface of PPN, including peroxiredoxins (Dubreuil *et al*., 2011; Robertson *et al.,* 2000); superoxide dismutase and glutathione peroxidase (Jones *et al.,* 2004). A fatty acid and retinol-binding protein (FAR) was localised in the surface cuticle (hypodermis) of *G. pallida* and was shown to suppress lipoxygenase mediated breakdown of fatty acid, a key component of the Jasmonic acid signaling pathway (Prior *et al.,* 2001). A similar protein has also been identified in the migratory nematodes *Aphelenchoides besseyi* (Cheng *et al*., 2013), and *Pratylenchus penetrans* (Vieira *et al*., 2017).

2.1.4 Suppression of host defences

Biotrophic nematodes, such as RKN and PCN, secrete several proteins that act has suppressors of host defences. The venom-allergen protein from *G. rostochiensis* (Gp-VAP-1) targets a plant Rcr3 that plays a key role in apoplastic defence signaling (Lozano-Torres *et al*., 2012). Other enzymes such as annexins, that bind calcium and are associated with abiotic stresses were identified in *H. avenae* (Patel *et al*., 2010) and it was hypothesised that these can target host oxireductases to interfere with host defences. Several studies suggested that nematode chorismate mutase, another sequence acquired by horizontal gene transfer from bacteria, has a role in manipulating plant defences. The plant innate immune response is thought to have two layers. In the first of these, conserved pathogen molecules (Pathogen Associated Molecular Patterns - PAMPs) are recognized by pattern recognition receptors leading to induction of pattern triggerd immunity (PTI). Pathogens deploy effectors that suppress PTI and these are recognized in the second layer of defences, termed effector

triggered immunity (ETI). Induction of ETI is often associated with a strong, localized cell death response termed the hypersensitive response (HR) (Smant and Jones, 2011; Mantelin *et al*., 2017). Nematode effectors that suppress each of these defence responses have been identified. For example, a secreted calreticulin from *M. incognita* has been shown to suppress PTI while several secreted SPRY domain proteins (SPRYSECs) from cyst nematodes have been shown to suppress cell death associated with ETI (*e.g* Mei *et al.,* 2015). Suppression of host defences is therefore a key function of nematode effectors.

2.2 Identification of candidate effectors

2.2.1 Bioinformatic analysis

The analysis of genomic and transcriptomic data provide a huge opportunity to identify new effector genes involved in parasitism. The availability of nematode genomes started in the early 1990s when the genome of *Caenorhabditis elegans* was sequenced (*C. elegans* sequencing consortium, 1998). The first PPN genome to be fully sequenced was that of *Meloidogyne incognita* (Abad *et al*., 2008), followed by the closely-related species *M. hapla* (Opperman *et al*., 2008). Since then, an increasing number of both endo- and ectoparasitic genomes and transcriptomes have been published and are now available in platforms such as Wormbase parasite (Howe *et al*., 2017) and GeneDB (Logan-Klumpler *et al*., 2012) (Table 1).

To identify effector candidate genes, an *in silico* pipeline can be developed that is based on the fact that an effector is a secreted protein, with the presence of a signal peptide, that lacks transmembrane domain, and which is upregulated at a specific stage of the parasite life stage. Once a list of candidate effectors has been identified these can be functionally annoted based on a search for sequence similarity to other known protein domains in the databases, comparisons with sequences in other parasitic nematodes and predicted spatial localization of the protein (Reid and Jones, 2014). The most commonly used tools are SignalP (Bendtsen *et al*., 2004), TMHMM (Krogh *et al*., 2001) and Blast. This computational biology analysis can be performed using either command line tools or more user-friendly tools, such as the Galaxy platform (Cock and Pritchard, 2014).

Table 1 – List of plant-parasitic nematodes with their genome and/or transcriptomic data sequenced and available.

Nematode species	Genome	Transcriptome
Migratory		
Bursaphelenchus xylophilus	Yes (Kikuchi *et al*., 2011)	Mycophagous and parasitic life-stages (Espada *et al*., 2016; Tsai *et al.,* 2016); GC (*in preparation*)
Aphelenchoides besseyi		Mixed-stage life stages (Wang *et al*., 2014)
Pratylenchus coffeae	Yes (Burke *et al.,* 2015)	Mixed stages; EST; 454 (Haegeman *et al*., 2011b)
Pratylenchus thornei		Wheat plant; carrot discs; 454 (Nicol *et al*., 2012)
Pratylenchus zeae		Vermiform juveniles and adults; sugar-cane (Fosu-Nyarko *et al*., 2015)
Pratylenchus penetrans	Genome skimming (Denver *et al*., 2016)	Mixed parasitic life stages (3 and 7 DPI) (Vieira *et al*., 2015)
Pratylenchus neglectus	Genome skimming (Denver *et al*., 2016)	
Radopholus similis		Mixed life stages; EST (Jacob *et al*., 2008)
Rotylenchus reniformis		J2, sedentary stages (Eves-van den Akker *et al.,* 2016b)
Xiphinema index		Standard and Stressed conditions (S. Eves-van den Akker, pers. commun.)
Xiphinema americanum	Genome skimming (Denver *et al*., 2016)	
Longidorus elongatus		Preplanting (non-feeding) stage (S. Eves-van den Akker,, pers. commun.)
Anguina agrostis	Genome skimming (Denver *et al*., 2016)	
Ditylenchus africanus		EST (Haegeman *et al*., 2009)
Ditylenchus destructor	Zheng *et al*., 2016	Zheng *et al*., 2016
Sedentary		
Meloidogyne incognita	v1 (Abad *et al.,* 2008); v2 (Blanc-Mathieu *et al*., 2017)	Egg, early and late parasitic stages, adult female and male (Dubreuil et al. 2007; Jaouannet et al. 2012; Blanc-Mathieu *et al.,* 2017)
Meloidogyne hapla	Yes (Opperman *et al*., 2008)	
Meloidogyne floridensis	Yes (Lunt *et al.* 2014; Szitenberg *et al.* 2017)	
Meloidogyne enterolobii	Yes (Szitenberg *et al*. 2017)	
Meloidogyne arenaria	Yes (Blanc-Mathieu *et al*., 2017)	Eggs and infective J2 (Blanc-Mathieu *et al.,* 2017)
Meloidogyne javanica	Yes (Blanc-Mathieu *et al*., 2017; Szitenberg *et al*. 2017)	Eggs and infective J2 (Blanc-Mathieu *et al.,* 2017)
Meloidogyne graminicola		J2 (Haegeman *et al*., 2013; Petitot *et al*., 2016)
Globodera pallida	Yes (Cotton *et al.,* 2014)	Egg, J2, 7, 14, 21, 28, 35 DPI; male (Cotton *et al.,* 2014)
Globodera rostochiensis	Yes (Eves-van den Akker *et al*., 2016a)	Egg, J2, 14DPI (Eves-van den Akker *et al*., 2016a)
Globodera ellingtonae	Genome skimming (Denver *et al*., 2016)	Mixed life stages (Phillips *et al*., 2017)
Heterodera avenae		Egg, J1, J2, post-parasitic J2, J3, J4 and adults (Kumar *et al*, 2014; Yang *et al*., 2017)
Nacobbus aberrans		J2, migratory and sedentary adult life stages (Eves-van den Akker *et al*., 2014b)

DPI: days post infection; J: juveniles; v: version.

2.2.2 The pharyngeal gland cells

Pathogen effectors modulate plant defence responses, facilitate infection and initiate or maintain feeding sites (Gheysen and Mitchum, 2011). Although some effectors originate from amphids (e.g. HYP effectors in *G. pallida*, Eves-van den Akker *et al*., 2014a), and hypodermis (e.g. Jones *et al*., 2004), the majority of effectors are expressed in the GC. Tylenchida and

Aphelenchida nematodes have three gland cells – two subventral and one dorsal (Figure 9). In RKN and CN, the subventral GC open into the oesophageal lumen inside the median bulb and the dorsal cell is connected through a valve at the base of the stylet (Haegeman *et al*., 2012). In Aphelenchida, the three cells are bulbiform and multinucleate and the cells contain abundant secretory granules in the lumen and are all connected to the median bulb (Carletti *et al*., 2013). Proteins produced in these parasitism-specialized tissues are introduced into the host through the stylet and vary in morphology during the nematode life cycle. The subventral gland cells are enlarged in juvenile stages (in Aphelenchida) and mainly active in Tylenchida during penetration and migration in the roots. By contrast, the dorsal gland cell is enlarged in the adult parasitic life stage (Aphelenchida) and during the feeding site induction and maintenance (Haegeman *et al*., 2012; Carletti *et al*., 2013). Their adaptation and variation in size throughout the life cycle is indicative of their importance in parasitism.

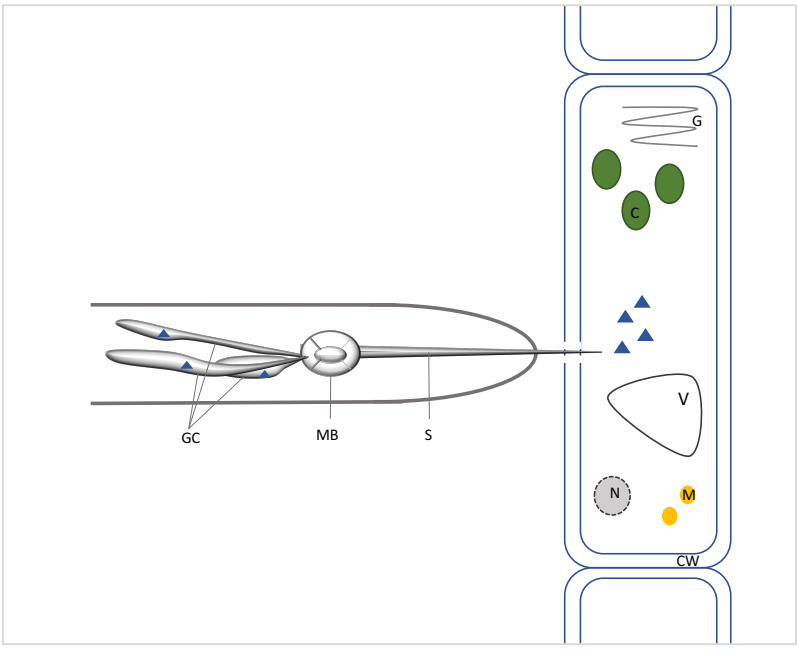


Figure 9 – Representation of the pinewood nematode infection in the pine tree cell, as an example of migratory endoparasitic nematode. Effectors are represented in the blue triangles and are produced in the gland cells and secreted into the plant cell by the stylet. C: chloroplast GC: gland cells; MB: medium bulb; S: stylet; CW: cell wall; M: mitochondrion; V: vacuole; N: nucleus; C: chloroplast; G: golgi.

Various methodologies have been developed to allow the identification of effectors expressed in these specialized cells. Early studies achieved this goal by analysis of Expressed Sequence Tags derived from cDNA from the GC of RKN and CN (e.g. Huang *et al*., 2003). More recently a method for microaspitration of the cytoplasm of gland cells, followed by RNA isolation and sequencing has been described (Maier *et al*., 2013). More recently, Eves-van den Akker *et al*. (2016a) identified a putative promoter motif - the DOGbox (ATGCCA) – associated with genes expressed in dorsal gland cell. The authors hypothesized that given that this non-coding genetic signature unifies many otherwise sequences unrelated effectors, it implies the existence of a

regulator, most likely a protein or a protein complex, that binds to this sequence in order to control expression of downstream genes (Eves-van den Akker *et al*., 2016a, c). Both, gland cell effectorome and transcriptional information will provide a new approach and a suite of tools for the discovery of a more complete set of novel mechanisms underlying parasitism.

2.3 *B. xylophilus* effectorome

2.3.1. Parasitism-related proteins

Until recently, very limited information was available on the effectors used by PWN to infect trees, reflecting a more general paucity of information on tree pathogens. *B. xylophilus* is not directly related to the sedentary PPNs (Figure 1), or other migratory nematodes that have been investigated in detail. The parasitism-related proteins that have been described in PWN include plant-cell walls degrading enzymes (beta-1,4-endoglucanase, GH45 family), pectate lyases (PEL) (Kikuchi *et al*., 2006) and expansins (Kikuchi *et al*., 2009). Enzymes that degrade fungal cell walls such as beta-1,3-endoglucanase (Kikuchi *et al*., 2005) and chitinases (from GH18 family) have also been described, although these are not truly effectors (Shinya *et al*, 2013). Several other proteins have been described and are possibly involved in the protection of the nematode against host defences by helping it to overcome ROS generated by plant cell defences such as peroxiredoxins (Li *et al*., 2016) and catalases (Vicente *et al*., 2015). Other proteins without known function have been identified including venom-allergen proteins (VAP) (Kang *et al*., 2012; Lin *et al*., 2011) and calreticulin (Li *et al.*, 2011). These previous studies were limited to characterization of enzymes used to degrade the plant or fungal cell walls or attempts to identify PWN orthologues of effectors from root-knot or cyst nematodes.

2.3.2. Functional studies

The functional studies on parasitism-related proteins help to understand the role of the effector genes in the parasitism process. This is achieved by silencing the gene or overexpression and evaluating the effect on the plant host and on the nematode development, survival and capacity of parasitism. One of the most promising methods for gene silencing is by using RNA interference (RNAi). Park *et al*. (Park *et al*., 2008) tested three different methods to deliver double strand RNA to *B. xylophilus* – microinjection, soaking and electroporation – and in each case the RNAi phenotypes for essential genes was relatively low but consistent (cytochrome C, myosin heavy chain, tropomyosin, heat shock proteins). The propagation and dispersion of the nematode decreased when the beta-1,4- endoglucanase gene expression was silenced by soaking RNAi (Ma *et al*., 2011) and nematode paralysis and uncoordinated movement was the result of soaking in dsRNA (double strand RNA) from the calponin gene (Bx-UNC-87) (Cardoso *et al*., 2015). More recently, Wang *et al*., 2016 described a new silencing vector (pDH-RH) for silencing the expression of genes of interest. The dsRNA vector can express the genes in transformed filamentous fungi via *Agrobacterium* and induce the knockout on the expression

of the target gene. This RNAi was delivered by feeding and tested on the nematode *dpy* genes (Wang *et al*., 2016). However, functional analysis of genes that are important for infection in a tree pathogen remains challenging.

3. OBJECTIVES AND THESIS OUTLINE

This thesis aims to identify new nematode parasitism proteins and understand the molecular mechanisms underlying parasitism. Therefore, this thesis addresses the question: What are the biological processes that allow a migratory endoparasitic nematode to infect trees? To accomplish this goal, genomic and transcriptomic approaches were used as non-biased approaches towards the identification of new parasitism-related elements. In Chapter 2 and 3, we will identify a complete suite of parasitism proteins – effectors – secreted by *B. xylophilus* and determine the differentially expressed genes in parasitic and pre-parasitic stages, using whole-nematode transcriptomic data. In Chapter 4 we will determine the functional role of one identified key parasitism protein in the nematode-host interaction. Together with information on transcriptional upregulated data we will, in Chapter 5, analyse the transcriptomic data from a specialized parasitism tissue and identify regulatory elements associated with expression of key pathogenicity secreted proteins.

4. REFERENCES

Abad, P., Gouzy, J., Aury, J.M., Castagnone-Sereno, P., Danchin, E.G., *et al.* (2008) Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*. *Nature Biotech* 26, 909–15.

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

Bird, D. McK., Jones, J.T., Opperman, C.H., kikuchi, T. and Danchin, E.G.J. (2015) Signatures of adaptation to plant parasitism in nematode genomes. *Parasitology* 142, S71-S84.

Blanc-Mathieu, R., Perfus-Barbeoch, L., Aury, J-M., Da Rocha, M., Gouzy, J., Sallet, E., martin-Jimenez, C., Bailly-Bechet, M., Castagnone-Sereno, P., Flot, J-F, kozlowski, D.K., Cazareth, J., Couloux, A.M., da Silva, C., Guy, J., Kim-Jo, Y., Rancurel, C., Schiex, T., Abad, P., Wincker, P. and Danchin, E.G.J. (2017) Hybridization and polyploidy enable genomic plasticity without sex in the most devastating plant-parasitic nematodes. *PLoS Genet* 13(6): e1006777.

Blaxter, M. and Koutsovoulos, G. (2015) The evolution of parasitism in Nematoda. *Parasitology* 142, S26-39.

Bongers, T. and Ferris, H. (1999) Nematode community structure as a bioindicator in environmental monitoring. *Trends Ecol. Evol.* 1486, 224-228.

Burke, M., Scholl, E.H., Bird, D.M., Schaff, J.E., Colman, S.D., Crowell, R., Diener; S., Gordon, O., Graham, S., Wang, X., Windham, E., Wright, G. M. and Opperman, C.H. (2015) The plant parasite *Pratylenchus coffeae* carries a minimal nematode genome. *Nematology* 17, 621–37.

Cardoso, J.M.S., Fonseca, L., Gomes, P., Egas, C. and Abrantes, I. (2015) Molecular characterization and functional analysis of a calponin gene from the pinewood nematode. *Forest Pathol.* 45, 467–473.

Carletti, B., Paoli, F., Isidoro, N. and Roversi, P.F. (2013) Ultrastructure of the anterior alimentary tract of *Bursaphelenchus mucronatus* Mamiya *et* Enda (Nematoda Aphelenchoididae). *Redia,* XCVI, 69-77.

Cheng, X., Xiang, Y., Xie, H., Xu, C-L., Xie, T-F., Zhang, C., Li, Y. (2013) Molecular Characterization and Functions of Fatty Acid and Retinoid Binding Protein Gene (Ab-far-1) in *Aphelenchoides besseyi*. *PLoS One* 8(6): e66011.

C. elegans Sequencing Consortium (1998) Genome sequence of the nematode *Caenorhabditis elegans*. A platform for investigating biology. *Science* 282: 2012-2018.

Cock, P.J.A. and Pritchard, L. (2014) Galaxy as a platform for identifying candidate pathogen effectors. In Plant-pathogen interactions: methods and protocols. Methods in molecular biology, volume 1127 (Birch, P. Jones, J.T. and Bos, J., eds.) pp. 3-16. Springer, New York.

Corsi, A.K., Whighman, B., Chalfie, M. (2015) A transparent window into biology: a primer on *Caenorhabditis elegans*. *Genetics* 200(2), 387-407.

Cotton, J.A., Lilley, C.J., Jones, L.M., Kikuchi, T., Reid, A.J., Thorpe, P., Tsai, I.J., Beasley, H., Blok, V., Cock, P.J.A., Eves-van den Akker, S., Holroyd, N., Hunt, M., Mantelin, S., Naghra, H., Pain, A., Palomares-Rius, J.E., Zarowiecki, M., Berriman, M., Jones, J.T. and 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 Biol*. 15, R43.

Denver, D.R., Brown, A.M.V., Howe, D.K., Peetz, A.B. and Zasada, I.A. (2016) Genome Skimming: a rapid approach to gaining diverse biological insights into multicellular pathogens. PLoS Pathog 12(8), e1005713.

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 Phytol.* 176(2), 426-36.

Dubreuil, G., Deleury, E., Magliano, M., Jaouannet, M., Abad, P., Rosso, M.N. (2011) Peroxiredoxins from the plant parasitic root-knot nematode, *Meloidogyne incognita*, are required for successful development within the host. *Int J Parasitol*. 41, 385-96.

EPPO quarantine pest. *Bursaphelenchus xylophilus*. Data Sheets on quarantine pests.

Espada, M., Silva, A.C., Eves-van den Akker, S., Cock, P.J.A., Mota, M. and Jones, J.T. (2016) Identification and characterization of parasitism genes from the pinewood nematode *Bursaphelenchus xylophilus* reveals a multi-layered detoxification strategy. *Mol. Plant Pathol.* 17, 286-295.

Eves-van den Akker, S., Lilley, C.J., Jones, J.T. and Urwin, P.E.. (2014a) Identification and Characterisation of a Hyper-Variable Apoplastic Effector Gene Family of the Potato Cyst Nematodes. *PLoS Pathog* 10(9), e1004391.

Eves-van den Akker, S., Lilley, C., Danchin, E., Rancurel, C., Cock, P.J.A., Urwin, P.E. and Jones, J.T. (2014b) The transcriptome of *Nacobbus aberrans* reveals insights into the evolution of sedentary endoparasitism in plant-parasitic nematodes. *Genome Biol. Evol.* 6, 2181–94.

Eves-van den Akker, S., Laetsch, D. R., Thorpe, P., Lilley, C.J., Danchin, E.G.J., da Rocha, M., Rancurel, C., Holroyd, N.E., Cotton, J.A., Szitenberg, A., Grenier, E., Montarry, J., Mimee, B., Duceppe, M., Boyes, I., Marvin, J.M.C., Jones, L.M., Yusup, H.B., Lafond-Lapalme, J., Esquibet, M., Sabeh,M., Rott, M., Overmars, H., Finkers-Tomczak, A., Smant, G., Koutsovoulos, G., Blok, V., Mantelin, S., Cock, P.J.A., Phillips, W., Henrissat, B., Urwin, P.E., Blaxter, M. and Jones, J.T. (2016a) 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. and Urwin, P.E. (2016b) Functional C-terminally encodedpeptide (CEP) plant hormone domains evolved de novo in the plant parasite *Rotylenchulus reniformis*. *Mol Plant Pathol* 17, 1265–75.

Eves-van den Akker, S. and Birch, P. R. (2016c). Opening the Effector Protein Toolbox for Plant–Parasitic Cyst Nematode Interactions. *Mol. Plant* 9, 1451-1453.

Fonseca, L., Cardoso, J.M.S., Lopes, A., Pestana, M., Abreu, F., Nunes, N., Mota, M. and Abrantes, I. (2012) The pinewood nematode, *Bursaphelenchus xylophilus*, in Madeira Island. *Helminthologia* 49, 96-103.

Fosu-Nyarko, J., Tan, J.C.H., Gill, R., Agrez, V.G., Rao, U. and Jones, M.G.K. (2015) *De novo* analysis of the transcriptome of *Pratylenchus zeae* to identify transcripts for proteins required for structural integrity, sensation, locomotion and parasitism. *Mol. Plant Pathol*. 1784, 532-552.

Futai, K. (2013) Pine wood nematode, *Bursaphelenchus xylophilus*. *Annu Rev Phytopathol*. 51, 5.1-5.23.

Gheysen, G. and Mitchum, M.G. (2011) How nematodes manipulate plant development pathways for infection. *Curr opin plant biol* 14, 415–421.

Haegeman, A., Jacob, J., Vanholme, B., Kyndt, T., Makedonka, M. and Gheysen, G. (2009) Expressed sequence tags of the peanut pod nematode *Ditylenchus africanus*: the first transcriptome analysis of an Anguinid nematode. *Mol. Biochem. Parasitol*. 167(1), 32-40.

Haegeman, A., Jones, J.T. and Danchin, E. (2011a) Horizontal gene transfer in nematodes: a catalyst for plant parasitism? *Mol Plant Microbe Interact.* 24, 879-887.

Haegeman, A., Joseph, S. and Gheysen, L. (2011b) Analysis of the transcriptome of the root lesion nematode *Pratylenchus coffeae* generated by 454 sequencing technology. *Mol Biochem Parasitol*. 178, 7-14.

Haegeman, A., Mantelin, S., Jones, J.T. and Gheysen, G. (2012) Functional roles of effectors of plant-parasitic nematodes. *Gene* 492, 19-31.

Haegeman, A., bauters, L., Kyndt, T., Rahman, M.M. and Gheysen, G. (2013) Identification of candidate effector genes in the transcriptome of the rice root knot nematode *Meloidogyne graminicola*. *Mol. Plant. Pathol.* 14(4), 379-390.

Holbein, J., Grundler, F.M. and Siddique, S. (2016) Plant basal resistance to nematodes: an update. *J Exp Bot.* 67(7), 2049-61.

Howe, K.L., Bolt, B.J., Shafie, M., Kersey, P. and Berriman, M. (2017) WormBase ParaSite − a comprehensive resource for helminth genomics. *Mol Biochem Parasitol*. 215, 2-10.

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

Inácio, M.L., Nóbrega, F., Vieira, P., Bonifácio, L., Naves, P., Sousa, E. and Mota, M. (2015) First detection of *Bursaphelenchus xylophilus* associated with *Pinus nigra* in Portugal and in Europe. *For. Pathol*. 45, 235-238.

Jacob, J., Mitreva, M., Vanholme, B. and Gheysen, G. (2008) Exploring the transcriptome of the burrowing nematode *Radopolus similis*. *Mol. Genet. Genomics* 280, 1-17.

Jaouannet, M., Perfus-Barbeoch, L., Deleury, E., Magliano, M., Engler, G., Vieira, P., Danchin, E.G.J., Da Rocha, M., Coquillard, P., Abad, P. and Rosso, M.N. (2012) A root-knot nematode-secreted protein is injected into giant cells and targeted to the nuclei. *New. Phytol*.194, 924-931.

Jaubert S., Laffaire, J.B., Abad, P. and Rosso, M.N. (2002) A polygalacturonase of animal origin isolated from the root-knot nematode *Meloidogyne incognita*. *FEBS Lett.*, 522, 109-12.

Jones, J.T., Reavy, B., Smant, G. and Prior, A.E. (2004) Glutathione peroxidases of the potato cyst nematode *Globodera rostochiensis*. *Gene* 7(324), 47-54.

Jones, J.T., Li. H., Moens, M., Mota, M. and Kikuchi, T. (2008) *Bursaphelenchus xylophilus*: opportunities in comparative genomics and molecular host-parasite interactions. *Mol Plant Pathol* 9, 357-368.

Jones, J.T., Haegeman, A., Danchin, E.G.J., Gaur, H.S., Helder, J., Jones, M.J.K., Kikuchi, T., Manzanilla-López, R., Palomares-Rius, J.E., Wesemael, W.M.L. and Perry, R.N. (2013) Top 10 plant-parasitic nematodes in molecular plant pathology. *Mol. Plant Pathol.* 14, 946-61.

Kang, J.S., Koh, Y.H, Moon, Y.S and Lee, S.H. (2012) Molecular properties of a venom allergen-like protein suggest a parasitic function in the pinewood nematode *Bursaphelenchus xylophilus* . *Int. J. Parasitol.* 42, 63-70.

Kanzaki, N. (2008) Taxonomy and systematics of the nematode genus *Bursaphelenchus* (Nematoda:Parasitaphelenchidae). In Pine wilt disease (Zhao, G., Futai, K. and Takeuchi, Y. eds.)., pp. 44-66. Springer, Japan.

Kikuchi, T., Jones, J.T., Aikawa, T., Kosaka, H. and Ogura, N. (2004) A family of GHF45 cellulases from the pine wood nematode *Bursaphelenchus xylophilus*. *FEBS Lett.* 572, 201-205.

Kikuchi, T., Shibuya, H. and Jones, J.T. (2005). Molecular and biochemical characterization of an Endo-ß-1,3-glucanase from the pine wood nematode *Bursaphelenchus xylophilus. Biochem. J.* 389, 117-125.

Kikuchi, T., Shibuya, H., Aikawa, T. and Jones, J.T. (2006) Cloning and characterization of pectate lyases secreted by the pine wood nematode *Bursaphelenchus xylophilus*. *Mol. Plant Microbe In*. 19, 280– 287.

Kikuchi, T., Li, H., Karim, N., Kennedy, M.W., Moens, M. and Jones, J.T. (2009) Identification of putative expansin-like genes from the pinewood nematode, *Bursaphelenchus xylophilus*, and evolution of the expansin gene family within the Nematoda. *Nematology* 11, 355-364.

Kikuchi, T., Cotton, J.A., Dalzell, J.J., Hasegawa, K., Kanzaki, N., McVeigh, P., Takanashi, T., Tsai, I.J., Assefa, S.A., Cock, P.J.A., Otto, T.D., Hunt, M., Reid, A.J., Sanchez-Flores, A., Tsuchihara, K., Yokoi, T., Larsson, M.C., Miwa, J., Maule, A.G., Sahashi, N., Jones, J.T. and Berriman, M. (2011) Genomic insights into the origin of parasitism in the emerging plant pathogen *Bursaphelenchus xylophilus*. *PLoS Pathogens* 7, e1002219.

Kikuchi, T., Eves-van den Akker, S. and Jones, J.T. (2017) Genome evolution of plant-parasitic nematodes. *Annu. Rev. Phytopathol*. 55, 333-354.

Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E.L.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567-580.

Kumar, M., Gantasala, N.P., Roychowdhury, T., Thakur, P.K., Banakar, P., Shukla, R.N., Jones, M.G.K. and Rao, U. (2014) De novo transcriptome sequencing and analysis of the cereal cyst nematode, *Heterodera avenae*. *PLoS ONE* 9(5): e96311.

Li, X., Zhuo, L., Luo, M., Sun, L. and Liao, S. (2011) Molecular cloning and characterization of a calreticulin cDNA from the pinewood nematode *Bursaphelenchus xylophilus*. Exp. *Parasitol.* 128, 121–126.

Li, Z., Zheng, Q. and Zhou, X. (2016) A 2-Cys peroxiredoxin in response to oxidative stress in the pine wood nematode, *Bursaphelenchus xylophilus*. *Sci. Rep.* 6, 27438.

Lin, S., Jian, H., Zhao, H., Yang, D. and Liu, Q. (2011) Cloning and characterization of a venom allergen-like protein gene cluster from the pinewood nematode *Bursaphelenchus xylophilus*. *Exp. Parasitol*.127, 440–447.

Logan-Klumpler, F.J., De Silva, N., Boehme, U., Rogers, M.B., Velarde, G., McQuillan, J.A., Carver, T., Aslett, M., Olsen, C., Subramanian, S., Phan, I., Farris, C., Mitra, S., Ramasamy, G., Wang, H., Tivey, A., Jackson, A., Houston, R., Parkhill, J., Holden, M., Harb, O.S., Brunk, B.P., Myler, P.J., Roos, D., Carrington, M., Smith, D.F., Hertz-Fowler, C. and Berriman, M. (2012) GeneDB--an annotation database for pathogens. *Nucleic acids res 40, D98-108.*

Lozano-Torres, J.L., Wilbers, R.H.P., Gawronski, P., Boshoven, J.C., Finkers-Tomczak, A., Cordewener, J.H.G., America, A.H.P., Overmars, H.A., Van,'t Klooster, J.W., Baranowski, L., Sobczak, M., Ilyas, M., van der Hoorn, R.A.L., Schots, A., de Wit, P.J.G.M., Bakker, J., Goversea, A. and Smant, G. (2012) Dual disease resistance mediated by the immune receptor Cf-2 in tomato requires a common virulence target of a fungus and a nematode. *Proc. Natl. Acad. Sci. USA* 109(25), 10119-24.

Lunt, D.H., Kumar, S., Koutsovoulos, G. and Blaxter, M.L. (2014) The complex hybrid origins of the root knot nematodes revealed through comparative genomics. *Peer J* 2, e356.

Ma, H.B., Lu, Q., Liang, J. and Zhang, X.Y. (2011) Functional analysis of the cellulose gene of the pine wood nematode, *Bursaphelenchus xylophilus*, using RNA interference. *Genet. Mol. Res.* 10 (3), 1931-1941.

Mamiya, Y. (1983) Pathology of the pine wilt disease caused by *Bursaphelenchus xylophilus*. *Annu. Rev. Phytopathol*. 21, 201-220.

Mamiya, Y. (2012). Scanning electron microscopy of pine seedling wood tissues sections inoculated with the pinewood nematode, *Burspahelenchus xylophilus*, previously prepared for light microscopy. *J. Nematol* 44, 255-259.

Mantelin, S., Thorpe, P. and Jones, J.T. (2017). Translational biology of nematode effectors. Or, to put it another way, functional analysis of effectors – what's the point? *Nematology* 19, 251-261.

Maier, T.R., Hewezi, T., Peng, J.Q. and Baum, T.J. (2013) Isolation of whole esophageal gland cells from plant-parasitic nematodes for transcriptome analyses and effector identification. *Mol Plant Microbe In.* 26, 31-35.

Mei, Y., Thorpe, P., Guzha, A., Haegeman, A., Blok, V.C., MacKenzie, K., Gheysen, G., Jones, J.T and Mantelin, S. (2015) Only a small subset of the SPRY domain gene family in *Globodera pallida* is likely to encode effectors, two of which suppress host defences induced by the potato resistance gene *Gpa2*. *Nematology* 17, 409-424.

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

Moens, M. and Perry, R. (2009) Migratory plant endoparasitic nematodes: a group rich in contrasts and divergence. *Ann. Rev. Phytophathol*. 47, 313-32.

Mota, M., Braasch, H., Bravo, M.A., Penas, A.C., Burgermeister, W., Metge, K. and Sousa, E. (1999) First report of *Bursaphelenchus xylophilus* in Portugal and in Europe. *Nematology* 1, 727-734.

Nascimento, F.X., Hasegawa, K., Mota, M. and Vicente, C.S. (2015) Bacterial role in pine wilt disease development – review and future perspectives. *Environ. Microbiol. Rep.* 7(1), 51–63.

Nicol, P., Gill, R., Fosu-Nyarko, J. and Jones, M.G. (2012) *De novo* analysis and functional classification of the transcriptome of the root lesion nematode, *Pratylenchus thornei*, after 454 GS FLX sequencing. *Int. J. Parasitol*. 42(3), 225-237.

Opperman, C.H., Bird, D.M., Williamson, V.M., Rokhsar, D.S., Burke, M., Cohn, J., Cromer, J., Diener, S., Gajan, J., Graham, S., Houfek, T.D., Liu, Q., Mitros, T., Schaff, J., Schaffer, R., Scholl, E., Sosinski, B.R., Thomas, V.P. and Windham, E. (2008) Sequence and genetic map *Meloidogyne hapla*: a compact nematode genome for parasitism. *Proc. Natl. Acad. Sci. U.S.A.* 105(39), 14802-14807.

Palomares-Rius, J., Tsai, I.J., Karim, N., Akiba, M., Kato, T., Maruyama, H., Takeuchi, Y. and Kikuchi, T. (2015) Genome-wide variation in the pinewood nematode *Bursaphelenchus xylophilus* and its relationship with pathogenic traits. *BMC Genomics* 16,845.

Park, J., Lee, K.Y., Lee, S., Oh, W., Jeong, P., Woo, T., Kim, C., Paik, Y. and Koo, H. (2008) The efficiency of RNA interference in *Bursaphelenchus xylophilus*. *Mol. Cells* 26, 81-86.

Patel, N., Hamamouch, N., Li, C., Hewezi, T., Hussey, R.S., Baum, T.J.,Mitchum, M.G., and Davis, E.L. (2010) A nematode effector protein similar to annexins in host plants. *J. Exp. Bot.* 61, 235–248.

Perry, R. and Moens, M. (2011) Introduction to plant-parasitic nematodes; modes of parasitism. In Genomics and molecular of plant-nematode interactions (Jones, J.T., Gheysen, G., Fenoll, C., eds). Springer Science+Business Media.

Petitot, A.-S., Dereeper, A., Agbessi, M., Da Silva, C., Guy, J., Ardisson, M. and Fernandez, D. (2016) Dual RNA-seq reveals *Meloidogyne graminicola* transcriptome and candidate effectors during the interaction with rice plants. *Mol. Plant Pathol.* 17, 860–874.

Phillips, W.S., Eves-van den Akker, S. and Zasada, I. (2017) Draft transcriptome of *Globodera ellingtonae*. *J Nematol* 49(2), 129-130.

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

Prior, A.E., Jones, J.T., Blok, V.C., Beauchamp, J., McDermott, L., Cooper, A. and Kennedy, M.W. (2001) A surface-associated retinol- and fatty acid-binding protein (GpLBP20) from the potato cyst nematode *Globodera pallida* – lipid binding activities, structural analysis and expression pattern. *Biochem. J.* 356, 387-394.

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

Quentin, M., Abad, P. and Favery, B. (2013) Plant parasitic nematode effectors target host defense and nuclear functions to establish feeding cells. *Front. Plant Sci*. 4, 53.

Reid, A.J. and Jones, J.T. (2014) Bioinformatics analysis of expression data to identify effector candidates. In Plant-pathogen interactions: methods and protocols, methods in molecular biology (Birch , P., Jones, J.T. and Bos, J., eds.) pp. 17-27.Springer, New York.

Robertson, L., Robertson, W.M., Sobczak, M., Bakker, J., Tetaud, E., Arinagayayam, M.R., Ferguson, M.A.J., Fairlamb, A.H. and Jones, J.T. (2000) Cloning, expression and functional characterisation of a thioredoxin peroxidase from the potato cyst nematode *Globodera rostochiensis*. *Mol. Biochem. Parasitol.* 111, 41-49.

Robertson, L., Arcos, S.C., Escuer, M., Merino, R.S., Esparrago, G., Abelleira, A. and Navas, A. (2011) Incidence of the pinewood nematode *Bursaphelenchus xylophlius* Steiner &Buhrer, 1934 (Nickle, 1970) in Spain. *Nematology* 13(6), 755-757.

Rodrigues, J.M., Sousa, E. and Abrantes, I. (2015) Pine wilt disease: historical overview. In Pine wilt disease in Europe: biological interactions and integrated management. (Sousa, E., Vale, F. and Abrantes, I., eds.). pp 13-32. FNAPF (Federação Nacional de Asssociações de Proprietários Florestais).

Shinya, R., Morisaka, H., Kikuchi, T., Takeuchi, Y., Ueda, M., and 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(6), e67377.

Smant, G., Stokkermans, J.P., Yan, Y., de Boer, J.M., Baum, T.J., Wang, X., Hussey, R.S., Gommers, F.J., Henrissat, B., Davis, E.L., Helder, J., Schots, A. and Bakker, J. (1998) Endogenous cellulases in animals: isolation of beta-1, 4-endoglucanase genes from two species of plant-parasitic cyst nematodes. *Proc. Natl. Acad. Sci. USA* 95(9), 4906-4911.

Smant, G. and Jones, J.T. (2011) Suppression of plant defences by nematodes. In Genomics and Molecular genetics of plant-nematode interactions. (Jones, J.T., Gheysen, G. and Fenoll, C., eds). pp. 273-286. SpringerScience+Business Media.

Szitenberg, A., Salazar-Jaramillo, L., Blok, V.C., Laetsch, D.R., Joseph, S., Williamson, V.M., Blaxter, M.L. and Lunt, D.H. (2017) Comparative genomics of apomictic root-knot nematodes: hybridization, ploidy, and dynamic genome change. *BioRxiv*. https://doi.org/10.1101/136085.

Tsai, I.J., Tanaka, R., Kanzaki, N., Akiba, M., Yokoi, T., Espada, M., Jones, J.T. and Kikuchi, T. (2016) Transcriptional and morphological changes in the transition from mycetophagous to phytophagous phase in the plant-parasitic nematode *Bursaphelenchus xylophilus*. *Mol. Plant Pathol.* 17(1), 77-83.

Vanholme, B., Wouter, V., Vanhouteghem, K., de Meutter, J., Cannoot, B. and Gheysen, G. (2007) Molecular characterization and functional importance of pectate lyase secreted by cyst nematode *Heterodera schachtii*. *Mol. Plant Pathol*. 8(3), 267-278.

Van Megen, H., Van Den Elsen, S., Holterman, M., Karssen, G., Mooyman, P., Bongers, T., Holovachov, O., Bakker, J. and Helder, J. (2009) A phylogenetic tree of nematodes based on about 1200 full-length small subunit ribosomal DNA sequences. *Nematology* 11(6), 927-950.

Vicente, C., Espada, M., Vieira, P. and Mota., M. (2012). Pine wilt disease: a threat to European forestry. *Eur. J. Plant Pathol*. 133, 89-99.

Vicente, C.S.L., Ikuyo, Y., Shinya, R., Mota, M. and Hasegawa, K. (2015) Catalases induction in high virulence pinewood nematode *Bursaphelenchus xylophilus* under hydrogen peroxide-induced stress. *PLoS ONE* 10(4), e0123839.

Vieira, P., Eves-van den Akker, S., Verma, R.,Wantoch, S., Eisenback, J.D. and 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(12), e0144674.

Vieira, P., Kamo, K. and Eisenback. J.D. (2017). Characterization and silencing of the fatty acid- and retinol binding. Pp-far-1 gene in *Pratylenchus penetrans*. *Plant Pathol*. 66, 1214-1224.

Wang, F., Li, D., Wang, Z., Dong, A., Liu, L., Wang, B., Chen, Q. and Liu, X. (2014) Transcriptomic analysis of the rice white tip nematode, *Aphelenchoides besseyi* (Nematoda: Aphelenchoididae). *PLoS ONE* 9(3), e91591.

Wang, M., Wang, D., Zhang, X. , Wang, X., Liu, W., Hou, X., Huang, X., Xie, B. and Cheng, X. (2016) Double-stranded RNA-mediated interference of dumpy genes in *Bursaphelenchus xylophilus* by feeding on filamentous fungal transformants. *Int. J. Parasitol.* 46, 351-60.

Webster, J. and Mota, M. (2008) Pine wilt disease: global issues, trade and economic impact. In Pinewilt disease: a worldwide threat to forest ecosystems (Mota, M. and Vieira, P., eds) pp1-3. Springer.

Wieczorek, K., El Ashry, A., Quentin, E.N., Grundler, F.M.W., Favery, B., Seifert, G.J. and Bohlmann, H. (2014) A distinct role of pectate lyases in the formation of feeding structures induced by cyst and root-knot nematodes. *Mol Plant Microbe Interact*. 27, 901–912.

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

Yang, D., Chen, C., Liu, Q. and Jian, H. 2017. Comparative analysis of pre- and post-parasitic transcriptomes and mining pioneer effectors of *Heterodera avenae*. *Cell Biosci* 7, 11.

Zhao, L., Mota, M. Vieira, P. Butcher, R.A. and Sun, J. (2014) Interspecific communication between pinewood nematode, its insect vector, and associated microbes. *Trends Parasitol.* , 30, 299-308.

Zheng, J., Peng, D., Chen, L., Liu, H., Chen, F., Xu, M., Ju, S., Ruan, L. and Sun, M. (2016) The *Ditylenchus destructor* genome provides new insights into the evolution of plant parasitic nematodes. *Proc. R. Soc. B* 283, 20160942.

CHAPTER II

IDENTIFICATION AND CHARACTERIZATION OF PARASITISM GENES FROM THE PINEWOOD NEMATODE *BURSAPHELENCHUS XYLOPHILUS* REVEALS A MULTI-LAYERED DETOXIFICATION STRATEGY

Data published in:
Margarida Espada, Ana Cláudia Silva, Sebastian Eves-van den Akker, Peter J.A. Cock, Manuel Mota, John T. Jones
Molecular Plant Pathology, 2016, 17(2), 286-95

Summary

The migratory endoparasitic nematode, *Bursaphelenchus xylophilus*, which is the causal agent of pine wilt disease, has phytophagous and mycetophagous phases during its life cycle. This highly unusual feature distinguishes it from other plant-parasitic nematodes and requires profound changes in biology between modes. During the phytophagous stage the nematode migrates within pine trees, feeding on the contents of parenchymal cells. Like other plant pathogens, *B. xylophilus* secretes effectors from pharyngeal gland cells into the host during infection. We provide the first description of changes in the morphology of these gland cells between juvenile and adult life stages. Using a comparative transcriptomic approach and an effector identification pipeline we identify numerous novel parasitism genes which may be important for mediating interactions of *B. xylophilus* with its host. In-depth characterisation of all parasitism genes using *in situ* hybridisation reveals two major categories of detoxification proteins, those specifically expressed in either the pharyngeal gland cells or the digestive system. These data suggest that *B. xylophilus* incorporates effectors in a multilayer detoxification strategy in order to protect itself from host defence responses during phytophagy.

Introduction

The pinewood nematode (PWN), *Bursaphelenchus xylophilus*, is a migratory plant endoparasitic nematode and is the causal agent of Pine Wilt Disease (PWD). The PWD complex includes the pathogenic agent, its insect vector (cerambycid beetles of the genus *Monochamus*) and the host, which can be one of several different *Pinus* species. *Bursaphelenchus xylophilus* is native to North America and causes little damage to indigenous tree species. However, it was introduced into China and Japan at the start of the 20th Century and here it has caused significant damage under the appropriate environmental conditions (Jones *et al.,* 2013). The nematode was found in Europe for the first time in 1999 (Mota *et al.,* 1999) and has now been detected in mainland Portugal, Madeira Island and Spain (Mota *et al*, 1999; Robertson *et al*., 2011; Fonseca *et al*., 2012). Pine wood represents a major proportion of the forestry industry and the rapid spread of this disease has become a major problem with the potential to cause significant economic losses and damage to forests on an ecological scale (Mota and Vieira, 2008; Vicente *et al*., 2012a).

The PWN has two different life cycle stages – a phytophagous parasitic stage and a mycetophagous stage. This highly unusual feature distinguishes it from other plant parasitic nematodes (PPN) and enables it to reproduce and survive in the host at the later stages of PWD when healthy plant tissues may be absent but fungi are abundant (Vicente *et al*., 2012a; Jones *et al*., 2013). Like many other nematode species, *B. xylophilus* has four juvenile stages prior to the mature adult and all life stages are vermiform. Nematodes can feed on fungi in dead

or dying trees and as nematode numbers increase, and food becomes scarce, a survival and dispersal stage develops (the *dauer* juvenile) that migrates to beetle pupal chambers. When the adult insect emerges, the *dauer* stage of the nematode enters the tracheid and is transported to a new host. The nematode may be transported to a dead or dying tree colonised with fungi, in which case the mycetophagous cycle described above begins again. Alternatively, the nematode can infect healthy host trees through maturation feeding wounds made by the insect. Once inside the pine cortex the nematode migrates to the xylem resin and ray canals and feeds on parenchyma cells leading to cell death (Mamiya, 2012). The tree releases polyphenolic coumpounds (causing browning of the tissues during infection), terpenoids, reactive oxygen species (ROS) and lipid peroxides during the early stages of infection as part of a strong defence response (Fukuda, 1997). Nematode numbers increase and water transport through the infected tree is compromised leading to wilt and, consequently, to death of the tree (Jones *et al*., 2008; Futai, 2013).

Although a genome sequence has been reported for *B. xylophilus* (Kikuchi *et al.,* 2011) the details of the mechanisms underlying the interaction between this nematode and its host remain unclear. Although peptides and plant hormones have been suggested to play important roles in the interactions between plants and nematodes, some of the most important nematode-derived factors that manipulate the host are effector proteins, many of which are produced in the pharyngeal gland cells and secreted into the host through the stylet. In aphelenchids (Ord. Rhabditida), which include *B. xylophilus*, these glands are composed of two subventral and one dorsal gland cell. Despite the morphological similarity of *B. xylophilus* to other PPNs, it is taxonomically unrelated (van Megen *et al.,* 2009) and has a uniquely complex mode of parasitism.

Effectors have been identified from PPNs, including effectors that induce changes in the host cells, facilitate migration and modulate host defences (reviewed by Haegeman *et al*., 2012; Mitchum *et al.,* 2013). However, the vast majority of these studies have focused on cyst and root-knot nematodes. Previous studies on PWN have often relied on attempting to identify orthologues of cyst nematode or root-knot nematode effectors from Expressed Sequence Tag (EST) and genomic datasets (Kikuchi *et al*., 2011; Yan *et al.,* 2012). This has allowed identification of a range of cell wall degrading enzymes that disrupt the plant and fungal cell wall, such as GH45 cellulases, several pectate lyases, expansins and beta-1,3-endoglucanases (Kikuchi *et al.*, 2004; Kikuchi *et al.*, 2005, Kikuchi *et al.*, 2006; Kikuchi *et al.*, 2009). However, PWN has an entirely different parasitic strategy from cyst nematodes and root knot nematodes, which does not require the nematode to keep host tissues alive for a prolonged period of biotrophy, and is taxonomically unrelated to these nematodes. It is therefore important to consider alternative approaches which do not make *a priori* assumptions about the nature of effector molecules. For example, one study has used proteomic analysis of secreted proteins collected from nematodes stimulated with pine extracts and identified cell

wall degrading enzymes, detoxification enzymes and peptidases amongst the secreted proteins (Shinya *et al*., 2013). In an alternative approach, microarray analysis has been used to identify secreted proteins upregulated during infection (Qiu *et al*., 2013).

Here we describe a differential expression based approach for identification of effectors from PWN. We use RNAseq and bioinformatic analyses to identify a panel of potentially secreted proteins upregulated after infection. Importantly, and in contrast to other studies of this type, we use *in situ* hybridisation to examine spatial expression profiles of candidate effectors and confirm that some are expressed in the pharyngeal gland cells. We show that detoxification proteins are deployed in a two-layer strategy, most likely in order to counter defence responses of the host. In addition, we examine morphological changes in the PWN pharyngeal gland cells across the life cycle and compare this with the development of these structures in cyst and root-knot nematodes.

Results

Characterisation of the pharyngeal gland cells of PWN

Previous studies on effectors of PWN have not attempted to identify the specific gland cells in which different putative effectors are expressed. This is frequently justified on the basis that the pharyngeal gland cells are difficult to distinguish as they are dorsally overlapping and all connect to similar positions in the large median oesophageal bulb (Nickle *et al.,* 1981). To rectify this, and to allow the precise site of expression of effectors to be determined, we first undertook a detailed morphological analysis of the structure of the pharyngeal gland cells in juveniles and adults of *B. xylophilus*. The dorsal and subventral gland cells were readily distinguished in both juveniles and adults (Figure 1). Measurements of the gland cells showed that although there was no significant difference in the size of the subventral gland cells between juveniles and adults, the dorsal gland is significantly larger ($p =< 0.05$) in the adult stage than in the juvenile stages (Figure 1; Table 1).

Table 1 - Measurements of the dorsal and subventral pharyngeal gland cells of *B. xylophilus*, BxPt75OH isolate [in µm and in form: mean ± SD (range)], calculated from ten individuals for each life stage.

	Juveniles	**Adults**
Dorsal gland cell	30.9±4.43 (24-38.2)	66.9±6.48 (53.5-73.8)
Subventral gland cells	57.5±8.62 (41.9-72)	41.5±2.26 (39.2-45.1)

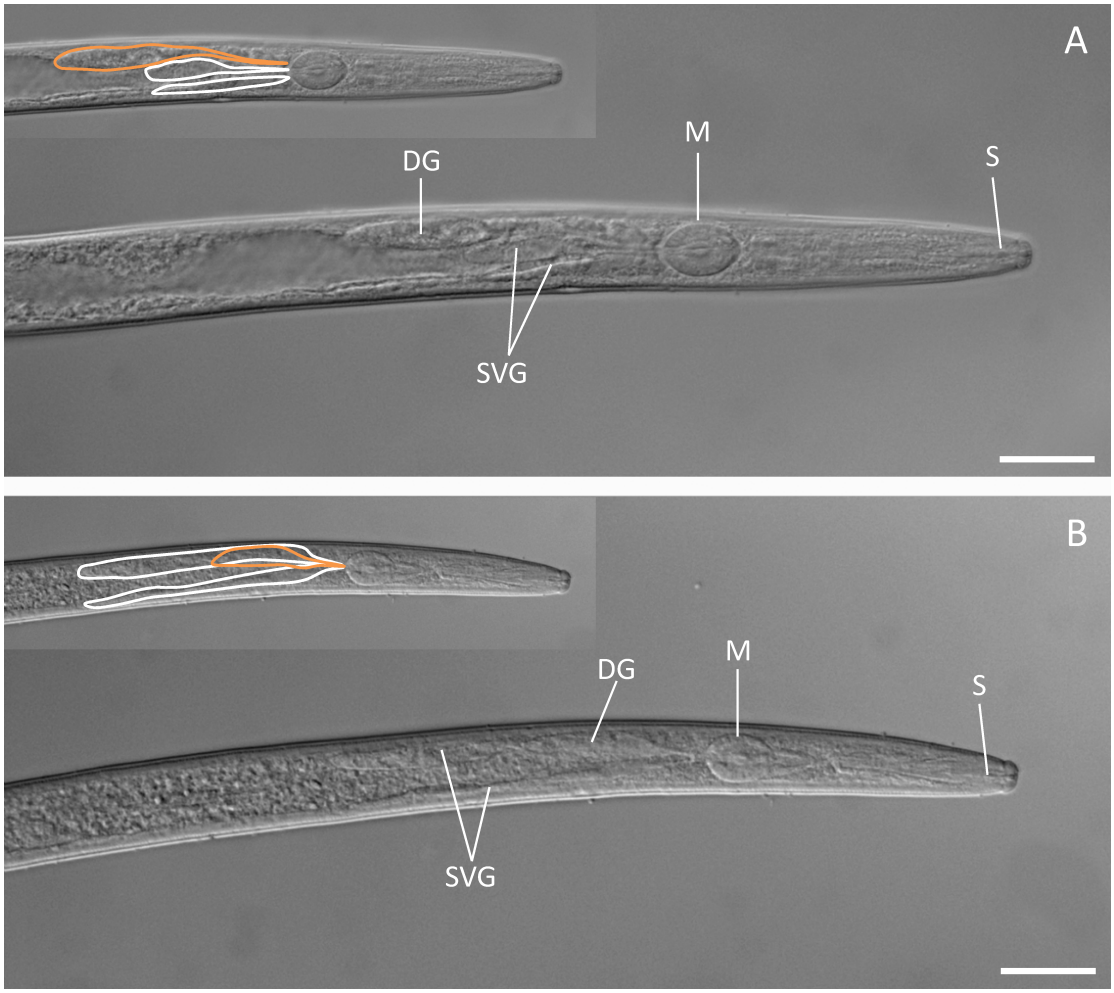


Figure 1- Positions of pharyngeal gland cells in adult (A) and juvenile (B) *B. xylophilus.* M: Median bulb; DG: Dorsal glands; SVG: Subventral glands; S: Stylet. Subventral glands (white) and dorsal gland (orange) are outlined in the duplicate figures below the main panels. (Scale bar = 20µm)

Differential gene expression in mycetophagous and phytophagous stages of B. xylophilus *and identification of candidate effectors*

Differential gene expression analysis showed extensive variation between replicates of some life conditions, in particular the fungal feeding (FF) and 15 days post infection (DPI) samples which failed to cluster in a heat map analysis. This meant that only twenty-nine transcripts were identified as being differentially expressed between the mycetophagous and phytophagous life stages (Supplementary Figure 1). These genes represent a much lower proportion of the *B. xylophilus* genes than expected, given the very different environments that these life stages represent. In spite of this, genes that may have a role in the host-parasite interaction were included in the sequences identified as differentially expressed after infection, including glutathione S-transferase (GST), GH45 cellulases, peptidases and GH16 endoglucanases (Table S1). An alternative differential expression approach was used in parallel. The top 200 sequences upregulated in the parasitic life stage of the nematode were identified. These

sequences included numerous known effectors from this species (*e.g.* cell wall degrading enzymes). The most highly represented Gene Ontology (GO) terms in this set of 200 genes in the molecular function category were hydrolase, oxidoreductase and lyase activity (Supplementary Figure 2). Seventy three of these 200 genes were predicted to have a signal peptide and to lack transmembrane domains. This represents a significant enrichment of potentially secreted proteins compared to the proportion in the whole predicted gene set for this nematode (36.5% versus 12.7%; p = <0.0001; chi-square test analysis). Fewer than half (33) of these 73 potentially secreted proteins gave matches in BLAST searches against the non-redundant (NR) database while the other 40 sequences encoded proteins that gave no matches and were therefore considered pioneers. A subset of 46 putatively secreted proteins were subsequently selected for further analysis (Table 2); these were the most highly upregulated during infection and/or had matches in the database which suggested a potential role in parasitism. These sequences include transcripts encoding several classes of proteases, fatty acid transport proteins, putative V5/TPx1 allergen-like proteins (VAPs), a lysozyme, several enzymes involved in the detoxification of xenobiotic compounds and the most highly expressed pioneer genes (Table 2). The pipeline used to generate this list of candidate effectors is summarised in Figure 2.

Table 2 – List of candidate effector genes categorized by predicted function.

Predicted function	Putative protein domain (GeneDB annotation)
Proteases (10)	Aspartic protease A1 (5) Cysteine proteases C1A (1); C46 (1) Serine-type protease (2) Metallo-type protease M13 (1)
Fatty Acid Metabolism (2)	Fatty acid retinoid binding proteins
Detoxification Of Xenobiotic Compounds (12)	FMO (flavin monooxygenase) (2) UDP-glucuronosyl transferase (2) Multicopper putative acid oxidase (1) Glutathione S-transferase (2) Cytochrome P450 (3) Acid phosphatase (1) Epoxide hydrolase (1)
Unknown Proteins Domain (Pioneers) (16)	None
Protein With Toxin Domain (2)	Metridin-like Shk toxin domain
Allergens (1)	Putative allergen V5/TPx1
Glycosyl Hydrolase Classes (2)	GH29 (alpha-L-fuco domain) GH30 GH2
Lysozyme Activity (1)	Lysozyme 7,8

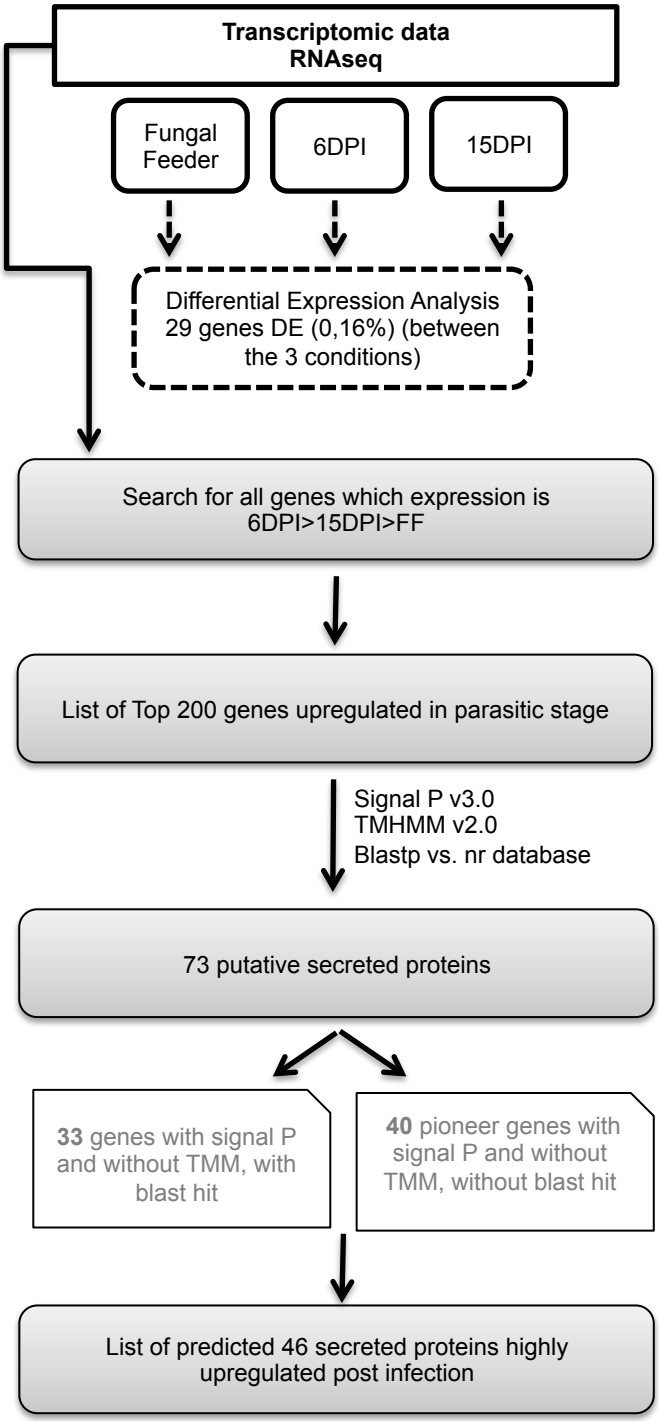


Figure 2 - Bioinformatic pipeline for the identification of candidate effectors from *B. xylophilus*. FF: Fungal feeder; DPI: days post infection.

Localisation and validation of effectors

In situ hybridisation was used to investigate the spatial expression patterns of the 46 putatively secreted proteins in mixed life stage-nematodes. The majority of the genes that gave a signal (18 sequences) were expressed in the intestine (Figure 3) while one gene was expressed in the glandular tissues surrounding the anterior sense organs (Figure 3A) and seventeen genes gave no signal in *in situ* hybridisation reactions (not shown). Ten genes were expressed in the gland cells; four in the dorsal gland cell and six in the subventral gland cells (Figure 4). The gland cell genes were similar in sequence to a putative fatty acid and retinoid binding protein (*BUX.s00422.201*) (Figure 4a), two pioneer genes (*BUX.s00083.48*, *BUX.s01109.178*) (Figure 4b, d), one cytochrome P450 (*BUX.s00116.698*) (Figure 4c), a lysozyme protein (*BUX.s01066.2*) (Figure 4e) and a predicted VAP protein (*BUX.s00116.606*) (Figure 4f) expressed in the subventral gland cells. Genes similar in sequence to two putative GSTs (*BUX.s01254.333*, *BUX.s00647.112*) (Figure 4h, j), one pioneer gene (*BUX.s01144.122*) (Figure 4i) and a peptidase C1A (*BUX.01147.177*) (Figure 4k) were expressed in the dorsal gland cell. No signal was detected using sense probes (*e.g.* Figure 4l, n). The ten gland cell localised sequences represent novel effectors that could be delivered into the host through the stylet during infection.

The expression levels of the ten putative effectors identified as being expressed in the gland cells were validated by semi-quantitative RT-PCR and compared with the results from the normalized expression values obtained by RNAseq (Figure S3). The RT-PCR showed that all the ten putative effector genes were expressed in nematodes after infection of the host. All of them, with the exception of the putative lysozyme (*BUX.s01066.2*) and cytochrome P450 (*BUX.s00116.698*), were also expressed in the fungal feeder condition. These latter two genes were only expressed at 15dpi and 6dpi, respectively.

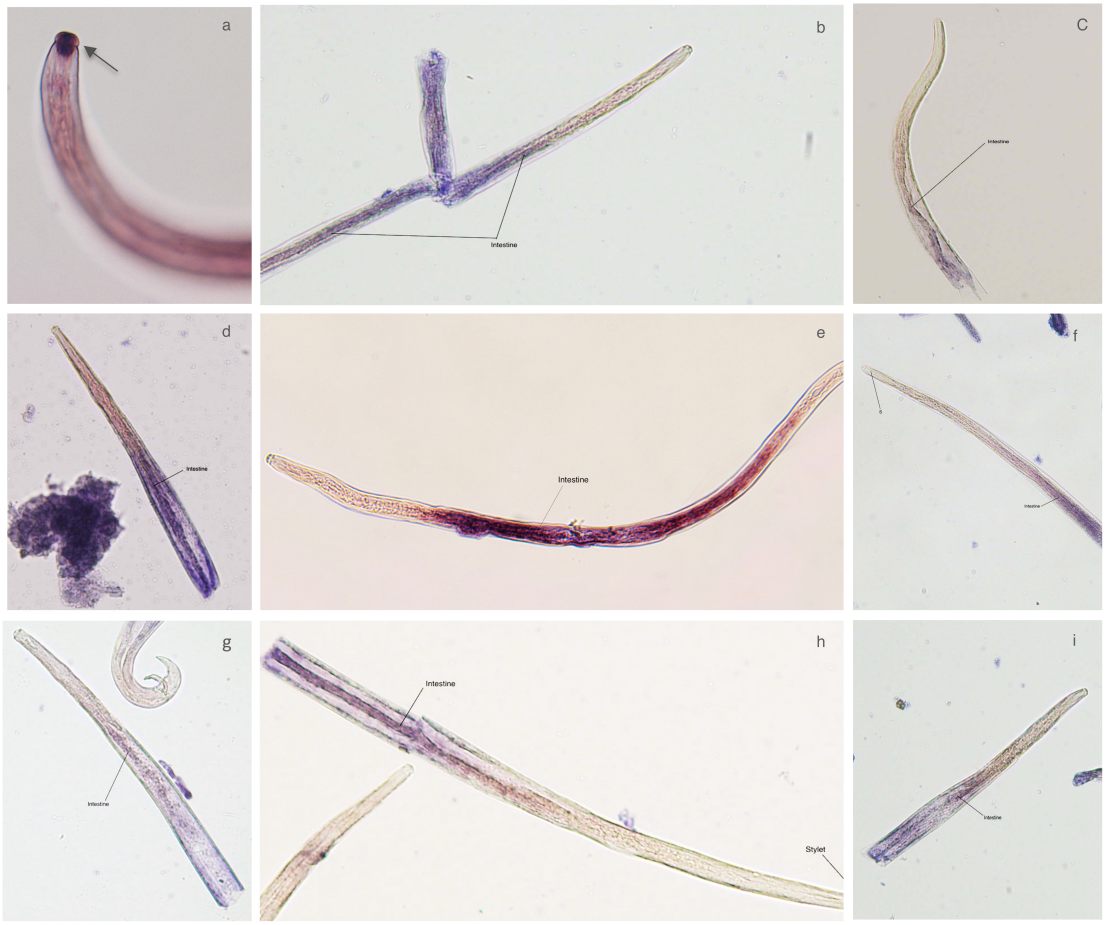


Figure 3 - Localisation of the candidate proteases and detoxification enzymes encoding genes expression in the intestine by *in situ* hybridization, with the exception of putative epoxide hydrolase (a) (*BUX.s00298.34*) that was expressed in the glandular tissues surrounding the anterior sense organs. b, putative multicopper oxidase (*BUX.s01281.17*); c, putative flavin monooxygenase (*BUX.s01337.7*); d, putative peptidase C46 (*BUX.s01109.245*); e, putative UDP-glucuronosyl transferase (UGT) (*BUX.s00422.680*); f, putative CYP33 C-related (*BUX.s01144.121*); g, putative peptidase M13 (*BUX.s01661.67*); h, putative peptidase A1 (*BUX.s00532.10*); i, putative peptidase S28 (*BUX.s01144.130*).

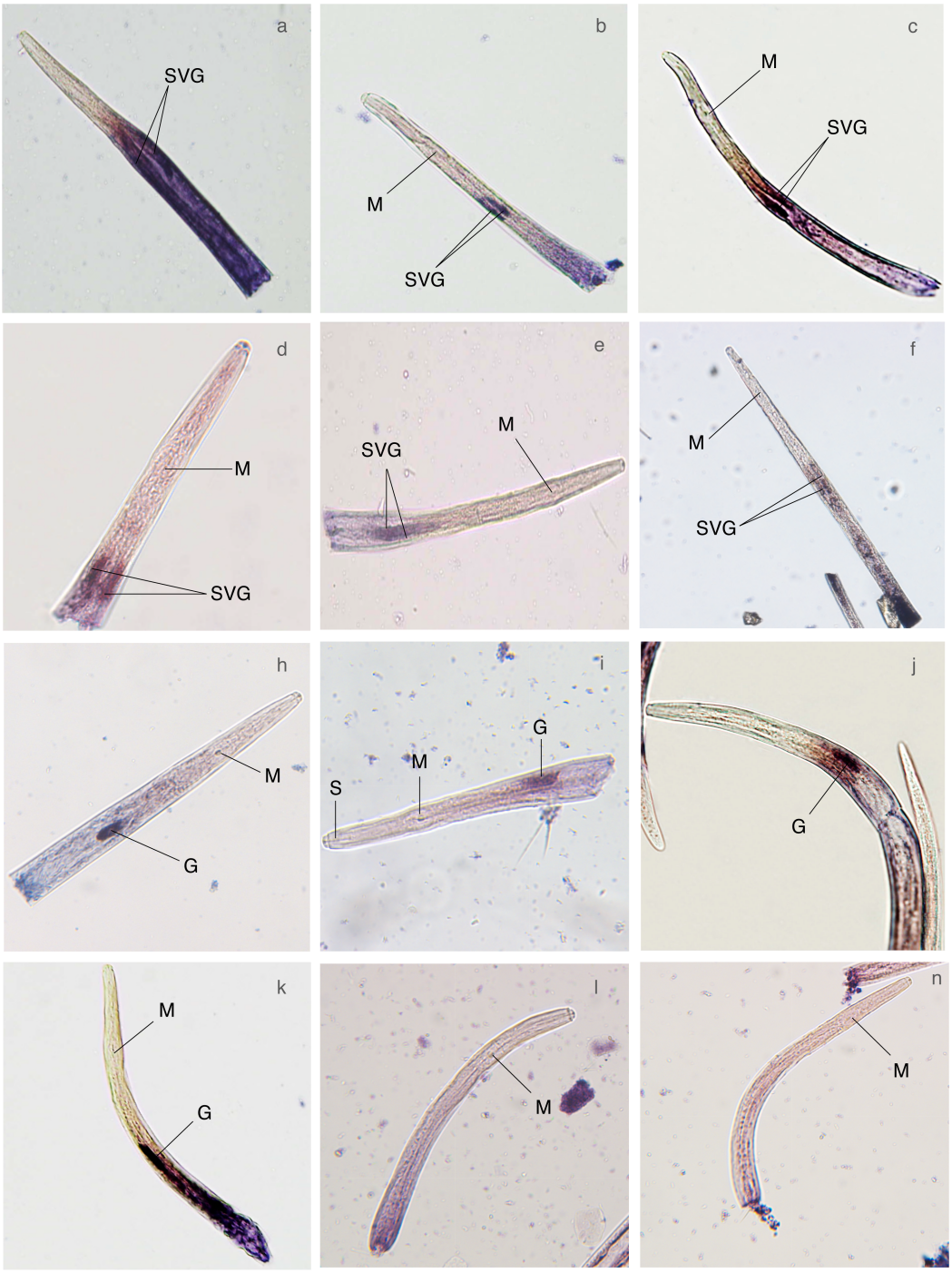


Figure 4 - Localisation of the candidate effectors expression in the pharyngeal gland cells by *in situ* hybridization. a, *BUX.s00422.201*, b, *BUX.s00083.48*, c, *BUX.s00116.698*, d, *BUX.s01109.178*, e, *BUX.s01066.2*, f, *BUX.s00116.606*, h, *BUX.s01254.333*, i, *BUX.s01144.122*, j, *BUX.s00647.112*, k, *BUX.s01147.177*, l and n are control Forward probe. M/MB: Median bulb; G: Dorsal gland cell; SVG: Subventral glands.

Discussion

A range of morphobiometric, ecological and population genetic studies have been carried out on *B. xylophilus* (Moens and Perry, 2009). Other studies have identified host physiological changes that occur upon the infection of the nematode (Fukuda, 1997; Hirao *et al*., 2012; Mamiya, 2012). However, compared to cyst and root-knot nematodes, little information is available on the nature of effectors secreted by PWN or the details of the molecular basis by which it parasitizes plants. *Bursaphelenchus xylophilus* has a unique feeding behaviour, a complex life cycle and infests a narrow host range of pine tree species. These features, coupled with the economic damage that it causes, make further studies on effector biology of *B. xylophilus* a priority.

The pharyngeal gland cells are the source of the majority of nematode effectors (*e.g.* Haegeman *et al*., 2012). Like most tylenchid nematodes (including root-knot and cyst nematodes) and other nematode groups, *B. xylophilus* has two subventral gland cells and one dorsal gland cell (Gheysen and Jones, 2006; Maule and Curtis, 2011; Haegeman *et al*., 2012). In *B. xylophilus*, the three pharyngeal gland cells dorsally overlap the intestine and are connected to similar positions in the large median bulb, which can make them difficult to distinguish (Nickle *et al*., 1981). Despite this, we were able to show that the dorsal gland cell in *B. xylophilus* is larger in the adult stages than in juveniles, as is seen in the sedentary stages of root-knot and cyst nematodes such as *Meloidogyne incognita* and *Heterodera glycines* (Endo, 1987; Hussey and Mims, 1990; Endo 1993). In sedentary nematodes the subventral gland cells decrease in size after the formation of the feeding structure (Maule and Curtis, 2011). By contrast, the subventral gland cells of *B. xylophilus* remain similar in size in juvenile and adult stages, suggesting a prolonged role in parasitism. Consistent with this, the majority of putative effectors identified here were subventral gland expressed. Together our findings align well with a recent study on *B. mucronatus*, a species closely related to *B. xylophilus*, which showed that a larger number of secretory granules are present in the subventral glands during the juvenile stages and in the dorsal gland during the adult stages (Carletti *et al*., 2013).

We generated transcriptomic datasets from mycophagous (pre-invasive of the host) and phytophagous (post-invasion of the host) stages of the nematode. Our first analysis unexpectedly showed extensive variation between replicates of the nematode samples, particularly at the later stages of infection. A similar independent study (T. Kikuchi pers. comm.) has shown that the environmental conditions (*e.g* time of year) experienced by the host have a profound effect on gene expression in parasitic *B. xylophilus* and it is likely that the variability seen here reflects a similar process. In order to collect the relatively large numbers of nematodes required for analysis, samples were collected from many different trees that may have been exposed to different environmental conditions. In spite of these issues we were able to identify a panel of genes that were significantly upregulated after infection and secreted

proteins were enriched in these sequences. Subsequent *in situ* hybridisation experiments identified ten putative effector proteins expressed in the gland cells, validating the approach. A comparison of these secreted proteins with the PWN secretome dataset obtained in a previous study using a proteomic approach (Shinya *et al*., 2013), showed that five of the effectors identified here were also identified in secreted proteins collected from *B. xylophilus* (data not shown). Although there are clearly differences in the results obtained using the two approaches, it is reassuring to see some measure of cross validation between the two studies.

A significant proportion of the sequences upregulated during the transition to parasitism, including some of the identified effectors, are likely to have roles in protecting the nematode from host defence responses. Pine trees respond to nematode infection by releasing a range of defence compounds in the areas surrounding the entry wound including ethylene, terpenoids (alpha and beta-pinene), ROS and lipid peroxides (Fukuda, 1997). Our study revealed that one secreted cytochrome P450 and two secreted GSTs upregulated at the early stages of infection (6dpi) are expressed in the subventral and dorsal gland cells respectively (Figure 5). These two enzymes are major components of the pathway leading to metabolism of xenobiotic compounds in the free-living nematode, *Caenorhabditis elegans* (Lindblom and Dodd, 2006). A secreted GST has also been identified that plays an important role in parasitism of plants by root-knot nematodes, and that most likely protects the nematode against host defences (Lindblom and Dodd, 2006; Dubreuil *et al*., 2007). Our results suggest that GST plays a similar role in *B. xylophilus* parasitism.

Our analysis showed that a range of transcripts encoding other enzymes potentially involved in the detoxification of xenobiotic compounds (including epoxide hydrolase, multicopper oxidase, flavin monooxygenase, UGT and cytochrome P450) are upregulated after infection but are expressed in the intestine (Figure 3). A recent study in *C. elegans* showed that the intestine is the first line of defence against xenobiotic compounds to oxidative-stress and emphasized the importance of phase 2 detoxification enzymes in this process (Crook-McMahon *et al*., 2014). Our data suggest that *B. xylophilus* uses a two-layered approach to protect itself against host-derived xenobiotic compounds. Some enzymes involved in detoxification pathways are secreted into the host representing the first layer, while others are upregulated in the digestive system, which will be exposed to ingested host materials, and represent the second.

The other identified effectors have a range of potential roles in the host-parasite interaction. One effector was similar to secreted venom allergen like proteins (VAPs) from other nematodes and was highly expressed 6 dpi. Three secreted VAPs have previously been characterized from PWN (Lin *et al*., 2011). It has been suggested that one of these (*Bx-vap-1*) is involved in migration of PWN inside the host (Kang *et al*., 2012). More recently, a study of the potato cyst nematode *Globodera rostochiensis* has shown that VAPs from this species are required for suppression of host immunity, possibly through a proteinase inhibition activity (Lozano-Torres

et al., 2014). VAPs are conserved throughout nematodes and are frequently upregulated in parasitic nematodes upon infection. It is therefore possible that VAPs are widely deployed against host defence responses that require the activity of host proteinases.

The *B. xylophilus* genome encodes hundreds of proteinases (Kikuchi *et al*., 2011). Our RNAseq analysis showed that several, including cysteine, metallo, aspartic and serine catalytic classes, are upregulated after infection. The majority of these were expressed in the intestine (Figure 3), consistent with a role in digestion. However, we identified a cysteine proteinase C1A that is expressed in the gland cells and upregulated at the later stage of infection (15dpi). This enzyme could have a role in digesting host tissues during migration or may also target host proteins involved in defence responses, as has been shown in animal parasitic nematodes (Sajid and McKerrow, 2002; Malagón *et al*., 2013). Consistent with this, plants are known to deploy proteinase inhibitors against pathogens (Xia, 2004).

A secreted fatty acid and retinol binding protein (FAR) was identified that is expressed in the subventral gland cells during the infection of the host. Most nematode lipid binding proteins are thought to be important for internal transport of lipids. However, FAR proteins have been identified both cyst (*Globodera pallida*) and root-knot nematodes that bind precursors of lipid-based plant defence signalling compounds important in the jasmonate signalling pathway (Prior *et al.,* 2001; Iberkleid *et al.,* 2013). The role of these pathways in terms of the interaction between *B. xylophilus* and its host remains to be determined.

One effector sequence was similar to lysozymes from a range of nematode species. Nematode lysozymes may have a role in digestion of host proteins and may also be important in protection of nematodes against other pathogens. Several lysozymes with antibacterial activity have been described from *C. elegans* (Boehnisch *et al*., 2011) that are thought to play an important role in defence against pathogenic bacteria. It is known that *B. xylophilus* is associated with a range of bacterial species that may form an important component of the infection process (Vicente *et al*., 2012b). The deployment of lysozyme by *B. xylophilus* may restrict bacterial growth in the regions infected by the nematode, reducing competition for food resources.

Our analysis also identified three pioneer genes expressed in the subventral and dorsal gland cells that are highly upregulated at 6 and 15 dpi. Given the absence of these proteins from other nematodes, they are likely to play key roles in the biology of *B. xylophilus*. Effectors from other nematodes are frequently novel proteins (*e.g.* Gao *et al.,* 2003). Characterising the function of such sequences in detail is likely to be challenging.

In summary, we describe a transcriptomic approach that has allowed identification of ten novel effectors and eighteen proteins from the digestive system of *B. xylophilus*. We also demonstrate that the gland cells of this species, like those of other plant-parasitic nematodes

change in structure during the life cycle. Our data suggest that *B. xylophilus* uses a multi-layered system of enzymatic detoxification to metabolise host derived xenobiotics within the host and in the digestive system.

Experimental procedures

Biological material

The Portuguese isolate of *B. xylophilus,* BxPt75OH, used in this study originated from a symptomatic pine tree in Oliveira do Hospital district, in the central region of mainland Portugal. The nematode was identified to species level (Nickle *et al*., 1981) and cultures were maintained in Erlenmeyer flasks containing *Botrytis cinerea* on barley seeds at 25ºC (Evans, 1970). Nematodes were extracted using the Baermann funnel technique (Southey, 1986) for 24 hours followed by sieving (38µm).

Morphometric studies of the pharyngeal gland cells

Mixed life-stage nematodes were killed by heat (water bath for approximately 15 minutes until the temperature reaches 60ºC) and fixed in 4% formaldehyde and prepared for mounting according to Siddiqi (1964). The nematodes were transferred into lactophenol and incubated for 24 hours at 40ºC. Nematodes were then transferred to a solution of 75% glycerine: 25% lactophenol for approximately 24 hours at 40ºC, until the lactophenol had evaporated and the nematodes were in pure glycerine. The nematodes were then mounted in glycerine surrounded by a ring of paraffin on a glass slide. A coverslip was placed on the top of the paraffin ring and the preparation was heated until the paraffin had melted. The slides were observed under a laser scanning microscope (Zeiss LSM 710) using the DIC (Differential Interference Contrast) method.

Measurements of the dorsal and subventral glands cells were performed from ten individuals for each of the life stages (juveniles and adults) , mounted using an agar pad technique as described by Eisenback (2012). Statistical significance was tested using Mann-Whitney U test analysis (STATISTICA v12.0) (Mann and Whitney, 1947). Images (measurements) were recorded using an Olympus BX50 light microscope and Cell Software (Olympus).

PWN inoculation trials

Two-month old maritime pine trees (*Pinus pinaster*) obtained from a Portuguese nursery were used for inoculation of the PWN isolate. Approximately 2000 mixed life-stage nematodes were cultured on fungi as described above and inoculated into a small wound (5mm) made on the pine stem using a sterilized scalpel. Infections were conducted under controlled conditions (average temperature 23ºC, 50% humidity). A subset of the nematodes prepared for each biological replicate were frozen in liquid nitrogen and stored at -80ºC for RNA extraction as the mycetophagous controls. The inoculated nematodes were collected from the trees, six and

fifteen days post infection. For this, the pine stems were cut and nematodes were collected by the Baermann funnel technique for approximately 2hrs. Nematodes were centrifuged by sucrose flotation (50%), washed three times in 1X Phosphate Buffered Saline (PBS) and frozen in liquid nitrogen.

RNA Extraction and Sequencing

Nematode RNA was extracted from samples corresponding to three different conditions: fungal feeding (pre-inoculation), 6 days post infection (dpi) and 15 dpi. RNA extraction was performed using the GeneJET RNA Purification Kit (Fermentas-ThermoScientific) following the manufacturer's instructions. RNA integrity number was assessed using a Bioanalyser (Agilent Technologies). The samples (two biological replicates for fungal feeding condition and three biological replicates for the other two conditions), each with a RNA Integrity Number - RIN over the value of 7, were used for paired end sequencing at The Genome Analysis Centre (TGAC, UK), on the Illumina HiSeq platform. RNAseq data described in this manuscript are available through ENA under accession number PRJEB9165.

Differential gene expression analysis

Raw RNA reads were trimmed of adapter sequences and low quality bases (phred score < 22) using Trimmomatic v0.32 (Bolger *et al*., 2014) (Table S2). Remaining high quality reads (79%) for each library were mapped back to the reference genome (http://www.genedb.org/Homepage/Bxylophilus) (Kikuchi *et al*., 2011) using Tophat2 (Kim *et al*., 2013). Read counts for each gene were determined using bedtools v2.16.2 and normalised (TMM) using Trinity wrapper scripts (Haas *et al*., 2013) for EdgeR (Robinson *et al*, 2010). Two differential expression analyses were carried out on normalised read counts: 1) Transcripts with a minimum fold change of 4 ($p < 0.001$) between conditions were identified using Trinity wrapper scripts for EdgeR, and clustered based on 20% tree height. 2) All genes were ranked by the ratio of their average normalized expression during all *in planta* stages (6 dpi + 15dpi) compared to fungal feeding. The top two hundred most differentially regulated genes were selected for further analyses. Potentially secreted protein sequences were identified using a workflow within a local installation of Galaxy on the basis of the presence of an N-terminal signal peptide (predicted by SignalP 3.0; Bendtsen *et al.,* 2004) and the absence of a transmembrane domain (predicted by TMHMM 2.0; Krogh *et al.,* 2001) (Cock and Pritchard, 2014). A BLASTp search (using Galaxy version 0.1.01) was performed against the non-redundant (NR) database (cutoff value of 1e-03), for all candidates, in order to predict their functions based on sequence similarity. Putative protein domain description is based on the annotation of the *B. xylophilus* genome (version 1.2) available on Gene DB (http://www.genedb.org/Homepage/Bxylophilus).

In situ *hybridisation*

In situ hybridisation using digoxigenin labelled probes was performed in order to determine the spatial expression patterns of candidate effectors based on the protocol described by de Boer *et al*, 1998. For each candidate gene a fragment of approximately 200 base pairs was amplified from the coding region and used as template for synthesis of both sense and antisense probes. The primers used for these reactions are shown in Table S3.

Validation of the expression profiles of candidate effectors

The expression profiles of the genes identified as expressed in the gland cells were validated by semi-quantitative PCR as described in Chen *et al.,* 2005. Actin was used as a control for all reactions (Table S3). Expression levels of each gene relative to the actin control were determined in the three different conditions (FF, 6 and 15 DPI), using cDNA synthesised from total RNA as a template and after 30-35 cycles. The results were analysed by electrophoresis in agarose gels. The qualitative results were compared to the predicted expression values obtained by RNAseq data.

References

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

Boehnisch, C., Wong, D., Habig, M., Isermann, K., Michiels, N.K., Roeder, T., May, R.C. and Schulenburg, H. (2011) Protist-type lysozymes of the nematode *Caenorhabditis elegans* contribute to resistance against pathogenic *Bacillus thuringiensis*. *PLoS One* 6, e24619.

Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, btu170.

Carletti, B., Paoli, F., Isidoro, N. and Roversi, P.F. (2013) Ultrastructure of the anterior alimentary tract of *Bursaphelenchus mucronatus* Mamiya *et* Enda (Nematoda Aphelenchoididae). *Redia,* XCVI, 69-77

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

Crook-McMahon, H., Oláhová, M., Button, E.L., Winter, J.J. and Veal, E.A. (2014) Genome-wide screening identifies new genes required for stress-induced phase 2 detoxification gene expression in animals. *BMC Biol* 12, 64.

Cock, P.J.A. and Pritchard, L. (2014) Galaxy as a platform for identifying candidate pathogen effectors. In Plant-Pathogen Interactions: Methods and Protocols, 2nd Edition. Methods in Molecular Biology Volume 1127 (Birch, P., Jones, J. and Bos, J., eds), pp. 3-15. Springer, New York.

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

Dubreuil, M., Magliano, E., Deleury, P., Abad, P. and Rosso, M.N. (2007) Transcriptome analysis of root-knot nematode functions induced in the early stages of parasitism. *New Phytol.* 176, 426–436.

Eisenback, J. (2012) A technique for making high-resolution megapixel mosaic photomicrographs of nematodes. *J. Nematol* 44, 260–263.

Endo, B.Y. (1987) Ultrastructure of esophageal gland secretory granules in juveniles of

Heterodera glycines. J. Nematol 19, 469-483.

Endo, B.Y. (1993) Ultrastructure of subventral gland secretory granules in parasitic juveniles of the soybean cyst nematode, *Heterodera glycines. J. Helminth. Soc. Washington* 60, 22-34.

Evans, A.A.F. (1970) Mass culture of mycophagous nematodes. *J. Nematol.* 2, 99-100.

Fonseca, L., Cardoso, J.M.S., Lopes, A., Pestana, M., Abreu,F., Nunes, N., Mota, M. and Abrantes, I. (2012) The pinewood nematode, *Bursaphelenchus xylophilus*, in Madeira Island. *Helminthologia* 49, 96 – 103.

Fukuda, K. (1997) Physiological process of the symptom development and resistance mechanism in pine wilt disease. *J. For. Res*. 2, 171-181.

Futai, K. (2013) Pine Wood Nematode, *Bursaphelenchus xylophilus. Annu. Rev. Phytopathol.* 51, 61–83.

Gheysen, G. and Jones, J.T. (2006) Molecular aspects of plant-nematode interactions. In Plant Nematology (Perry, R.N. and Moens, M., eds), pp. 234-254. Wallingford, UK.

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

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N. and Regev, A. (2013) *De novo* transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. *Nat Protoc*. 8, 1494-1512.

Haegeman, A., Mantelin, S., Jones, J.T. and Gheysen, G. (2012) Functional roles of effectors of plant-parasitic nematodes. *Gene* 492, 19-31.

Hirao, T., Fukatsu' E. and Watanabe, A. (2012) Characterization of resistance to pine wood nematode infection in *Pinus thunbergii* using suppression subtractive hybridization. *BMC Plant Biol* 12:13.

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

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. *PLoS ONE* 8, e64586.

Jones, J.T., Li, H., Moens, M., Mota, M. and Kikuchi, T. (2008) *Bursaphelenchus xylophilus*: opportunities in comparative genomics and molecular host–parasite interactions. *Mol. Plant Pathol.* 9, 357–368.

Jones, J.T., Haegeman, A., Danchin, E.G.J., Gaur, H.S., Helder, J., Jones, M.J.K., Kikuchi, T., Manzanilla-López, R., Palomares-Rius, J.E., Wesemael, W.M.L., and Perry, R.N. (2013) Top 10 plant-parasitic nematodes in molecular plant pathology. *Mol. Plant Pathol.* 14, 946-61.

Kang, J.S., Koh, Y.H., Moon, Y.S. and Lee, S.H. (2012) Molecular properties of a venom allergen-like protein suggest a parasitic function in the pinewood nematode *Bursaphelenchus xylophilus*. *Int. J. Parasitol*. 42, 63–70.

Kikuchi, T., Jones, J.T., Aikawa, T., Kosaka, H. and Ogura, N. (2004) A family of GHF45 cellulases from the pine wood nematode *Bursaphelenchus xylophilus*. *FEBS Lett.* 572, 201–205.

Kikuchi, T., Shibuya, H. and Jones, J.T. (2005) Molecular and biochemical characterization of an endo-beta-1,3-glucanase from the pine wood nematode *Bursaphelenchus xylophilus*. *Biochem. J*. 389, 117–125.

Kikuchi, T., Shibuya, H., Aikawa, T. and Jones, J.T. (2006) Cloning and characterization of pectate lyases secreted by the pine wood nematode *Bursaphelenchus xylophilus*. *Mol. Plant Microbe In*. 19, 280– 287.

Kikuchi, T., Li, H., Karim, N., Kennedy, M.W., Moens, M. and Jones, J.T. (2009) Identification of putative expansin-like genes from the pine wood nematode, *Bursaphelenchus xylophilus*, and evolution of the expansin gene family within the Nematoda. *Nematology* 11, 355-364

Kikuchi, T., Cotton, J.A., Dalzell, J.J., Hasegawa, K., Kanzaki, N., McVeigh, P., Takanashi, T., Tsai, I.J., Assefa, S.A., Cock, P.J.A., Otto, T.D., Hunt, M., Reid, A.J., Sanchez-Flores, A., Tsuchihara, K., Yokoi, T., Larsson, M.C., Miwa, J., Maule, A.G., Sahashi, N., Jones, J.T. and Berriman, M. (2011) Genomic insights into the origin of parasitism in the emerging plant pathogen *Bursaphelenchus xylophilus*. *PLoS Pathog.* 7, e1002219.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, R36.

Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E.L.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567-580.

Lin, S., Jian, H., Zhao, H., Yang, D. and Liu, Q. (2011) Cloning and characterization of a venom allergen-like protein gene cluster from the pinewood nematode *Bursaphelenchus xylophilus*. *Exp. Parasitol*.127, 440–447.

Lindblom, T.H. and Dodd, A.K. (2006) Xenobiotic detoxification in the nematode *Caenorhabditis elegans. J. Exp. Zool.* 305, 720-730.

Lozano-Torres, J.L., Wilbers, R.H.P., Warmerdam, S., Finkers-Tomczak, A., Diaz-Granados, A., van Schaik, C.C., Helder, J., Bakker, J., Goverse, A., Schots, A. and Smant, G. (2014) Apoplastic venom allergen-like proteins of cyst nematodes modulate the activation of basal plant innate immunity by cell surface receptors. *PLoS Pathog* 10, e1004569.

Malagón, D., Benítez, R., Kašný, M. and Adroher,F.J. (2013) Peptidases in parasitic nematodes: a review. In Parasites: Ecology, Diseases and Management (Erzinger, G.S., ed), pp.61-102, Nova Science Publishers, Inc.

Mamiya, Y. (2012) Scanning electron microscopy of Pine seedling wood tissue sections inoculated with the Pinewood nematode *Bursaphelenchus xylophilus* previously prepared for light microscopy. *J. Nematol.* 44, 255–259.

Mann, H. B. and Whitney, D. R. (1947) On a test of whether one of two random variables is stochastically larger than the other*. Ann. Math. Stat.*18, 50–60.

Maule, A.G. and Curtis, R. (2011) Parallels between plant and animal parasitic nematode. In Genomics and Molecular Genetics of Plant-Nematode Interactions (Jones, J., Gheysen, G. and Fenoll, C., eds), pp. 221-251., Springer Science+Business Media B.V.

Mitchum, M.G., Hussey, R.S., Baum, T.J., Wang, X., Elling, A.A., Wubben, M. and Davis, E.L. (2013) Nematode effector proteins: an emerging paradigm of parasitism. *New Phytol*.199, 879–894.

Moens, M. and Perry, R.N. (2009) Migratory plant endoparasitic nematodes: a group rich in contrasts and divergence. *Annu. Rev. Phytopathol.* 47, 313–32.

Mota, M.M., Braasch, H., Bravo, M.A., Penas, A.C., Burgermeister, W., Metge, K. and Sousa, E. (1999) First report of *Bursaphelenchus xylophilus* in Portugal and in Europe. *Nematology* 1, 727–734.

Mota, M. and Vieira, P. (2008) Pine wilt disease: a worldwide threat to forest ecosystems. Springer.

Nickle, W. R., Golden, A. M., Mamiya, Y. and Wergin, W. P. (1981) On the taxonomy and morphology of the pine wood nematode, *Bursaphelenchus xylophilus* (Steiner & Buhrer 1934) Nickle 1970. *J. Nematol.* 13, 385–392.

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

Qiu, X., Wu, X., Huang, L., Tian, M. and Ye, J. (2013) Specifically expressed genes of the nematode *Bursaphelenchus xylophilus* involved with early interactions with pine trees. *PLoS One* 8, e78063.

Robertson, L., Arcos, S.C, Escuer, M., Merino, R.S., Esparrago, G. and Abelleira, A. and Navas, A. (2011) Incidence of the pinewood nematode *Bursaphelenchus xylophlius* Steiner & Buhrer, 1934 (Nickle, 1970) in Spain. *Nematology* 13, 755-757.

Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) EdgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26*,* 139–140*.*

Sajid, M. and McKerrow, J.H. (2002) Cysteine proteases of parasitic organisms. *Mol. Biochem. Parasit.* 120, 1-21.

Shinya, R., Morisaka, H., Kikuchi, T., Takeuchi, Y., Ueda, M. and 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, e67377.

Siddiqi, M.R. (1964) Studies on *Discolaimus* spp. (Nematoda: Dorylaimidae) from India. *Z. Zool. Syst. Evol*. 2, 174-184.

Southey, J.F. (1986) Laboratory methods for work with plant and soil nematodes. Ministry of Agriculture Fisheries and Food, HMSO. London, UK.

Van Megen, H., van den Elsen, S., Holterman, M., Karssen, G., Mooyman, P., Bongers, T., Holovachov, O., Bakker, J. and Helder, J. (2009) A phylogenetic tree of nematodes based on about 1200 full-length small subunit ribosomal DNA sequences. *Nematology* 11, 927-950.

Vicente, C.S.L., Espada, M., Vieira, P. and Mota, M. (2012a) Pine wilt disease: a threat to European forestry. *Eur. J. Plant Pathol.* 133, 89-99.

Vicente, C.S.L., Nascimento, F., Espada, M., Barbosa, P., Mota, M., Glick, B.R., and Oliveira, S. (2012b) Characterization of bacteria associated with pinewood nematode *Bursaphelenchus xylophilus*. *PLoS One* 7, e46661.

Xia, Y. (2004) Proteases in pathogens and plant defence. *Cell. Microbiol.* 6, 905-913.

Yan, X., Cheng, X., Wang, Y., Luo, J., Mao, Z., Ferris, V. R. and Xie, B. (2012) Comparative transcriptomics of two pathogenic pinewood nematodes yields insights into parasitic adaptation to life on pine hosts. *Gene* 505, 81-90.

Supporting information

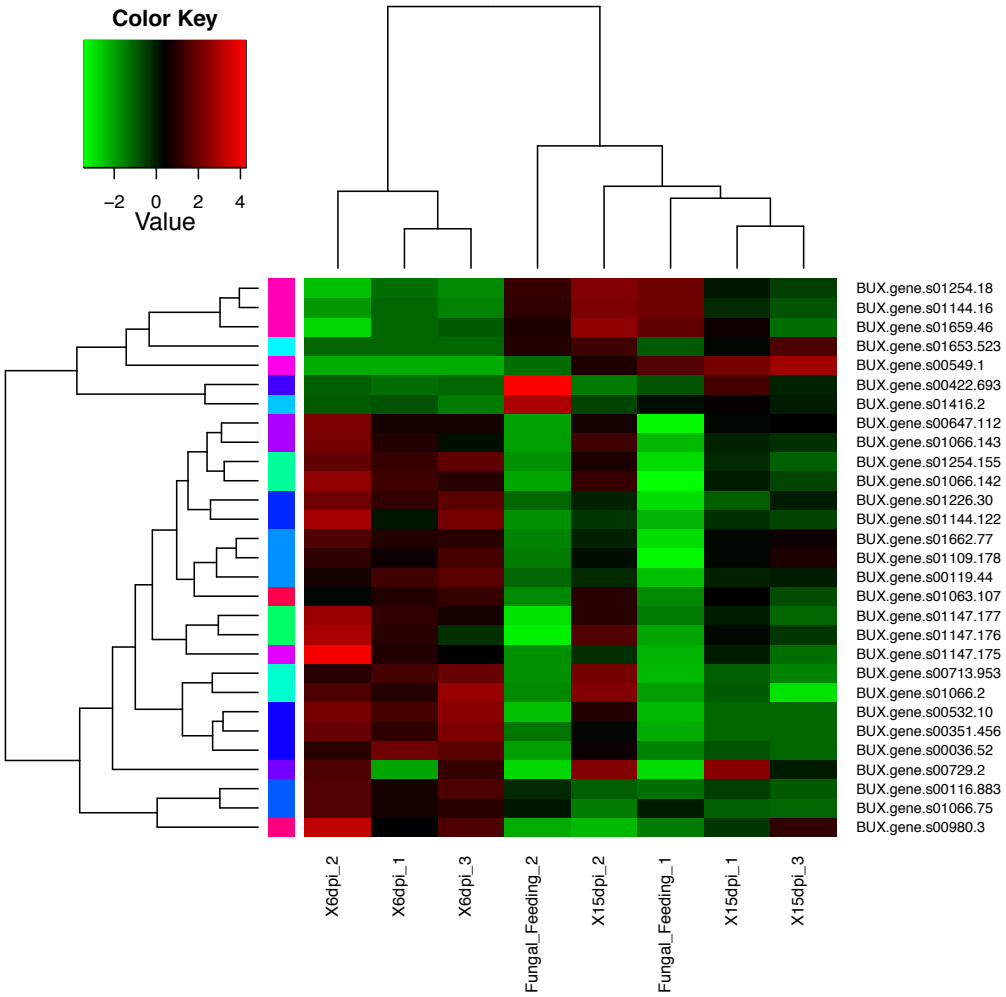


Figure S1 - Differential expression analysis of the transcripts. The heatmap resulting from the RNAseq analysis, using eight samples in three different conditions – pre-invasive/mycetophagous (Fungal Feeding, FF) and post invasive/phytophagous (6 and 15 days post-infection, dpi).

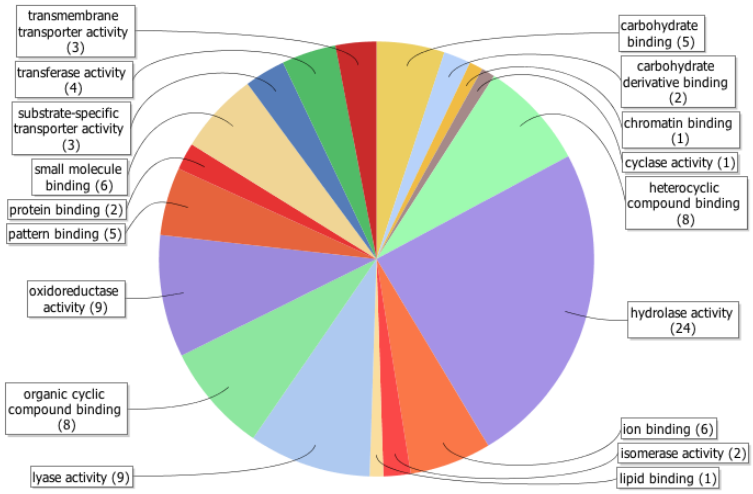


Figure S2 - Analysis of the most represented molecular function (level 3) in the Top 200 set of up regulated genes obtained by a bioinformatics pipeline.

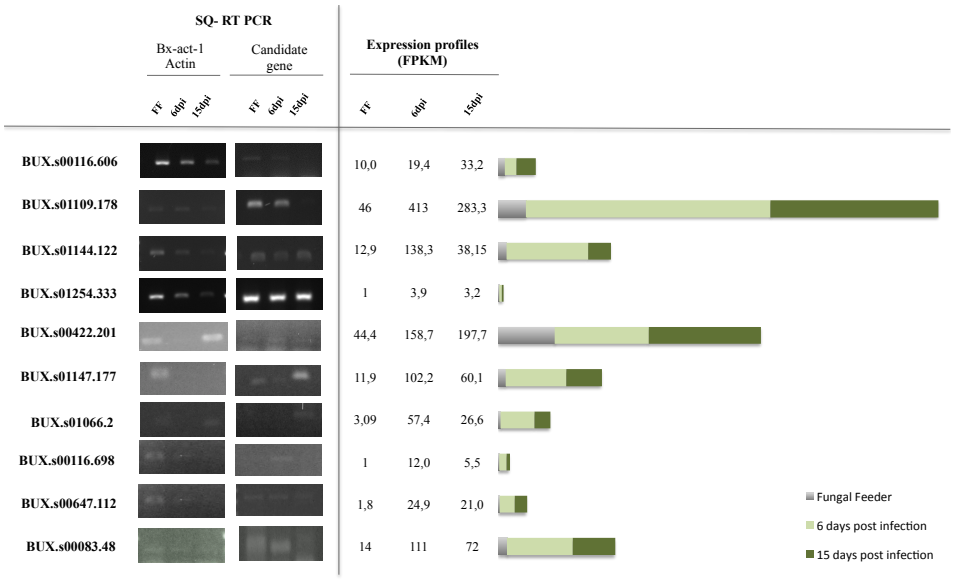


Figure S3 - Validation of the expression of the secreted effectors by semi-quantitative-polimerase chain reaction using the actin as housekeeping gene and the primers described in Supplementary Table 3. The results were analysed by gel electrophoresis and for each candidate the results of both actin and the candidate gene were presented. On the right, the bar chart represents the normalized expression values (fragments per kilobase of exon per million mapped fragments, FPKM) predicted by RNAseq for each candidate gene.

Table S1 – List of the 29 differentially expressed transcripts between mycetophagous and phytophagous stages. Detailed description of the 29 transcripts includes the presence or absence of putative signal peptide, the putative protein domain (according to Gene DB annotation of version 1.2 of the genome; available at http://www.genedb.org/Homepage/Bxylophilus), the top match and e-value of the BLASTp analysis against the non-redundant (NR) database (cut-off value, 1e-03), and the normalized expression profile in the three different conditions [fungal feeding (FF) nematodes and nematodes 6 and 15 days post-infection (dpi)]. The normalized expression values are in fragments per kilobase of exon per million mapped fragments (FPKM).

Gene code (v1.2)	Signal Peptide (Y/N)	Putative protein domain (Gene DB annotation)	Putative BLASTp match	BLAST p-value	FF_1	FF_2	6dpi_1	6dpi_2	6dpi_3	15dpi_1	15dpi_2	15dpi_3
BUX.gene.s00422.693	N	Peptidase aspartic	Hypothetical protein CAEBREN_22199 [Caenorhabditis brenneri]	3e-26	5,35	214,73	3,78	4,5	4,35	29,14	3,07	9,64
BUX.gene.s00729.2	Y	Peptidase A1, aspartic	No match		0,65	0,78	2,09	36,15	28,97	66,66	63,8	11,51
BUX.gene.s01066.143	Y	Glycoside hydrolase, family 16	Beta-1,3-endoglucanase [*Bursaphelenchus xylophilus*]	1e-93	11,79	15,64	122,75	282,7	76,81	62,56	169,8	52,98
BUX.gene.s00647.112	Y	Glutathione S-Transferase	Glutathione S-transferase [*H. saltator*]; glutathione S-transferase-1 [*M. incognita*]	4e-17	0,65	3,14	24,32	73,04	25,51	18,32	24,78	20,16
BUX.gene.s01147.175	Y	Peptidase C1A	Cathepsin [Pseudaletia unipuncta granulovirus]	4e-133	0	0,38	7,99	83,5	5,47	3,93	2,76	1,01
BUX.gene.s00549.1	N	WD40 repeat	No match		12,89	0,84	0	0	0	18,57	6,78	28,29
BUX.gene.s01144.16	N	Cytochrome c oxidase subunit IV	Cytochrome c oxidase subunit IV [Necator americanus]	2e-83	879,61	468,41	98,15	56,99	70,54	180,68	1008,84	117,71
BUX.gene.s01659.46	N	Unknown domain	No match		489,9	243,07	62,03	18,28	70,26	207,8	788,26	57,53
BUX.gene.s01254.18	Y	Putative sugar binding, galectin, carbohydrate reco domain	No match		271,17	159,26	27,75	11,41	21,25	70,65	341,92	46,69
BUX.gene.s00980.3	N	Unknown domain	No match		1,63	0,68	8,5	68,15	21,12	4,44	0,44	15,34
BUX.gene.s01063.107	Y	Unknown domain	No match		15,28	14,72	95,54	61,26	122	67,27	102,17	30,06
BUX.gene.s01147.176	Y	Proteinase inhibitor I29, cathepsin propeptide	Cathepsin [Pseudaletia unipuncta granulovirus]	3e-135	6,79	2,77	66,39	257,44	25,84	39,75	99,58	24,48
BUX.gene.s01147.177	Y	Proteinase inhibitor I29, cathepsin propeptide	Cathepsin [Pseudaletia unipuncta granulovirus]	1e-135	18,68	5,59	116,51	332,42	87,39	54,1	104,6	23,9
BUX.gene.s01066.142	Y	Concanavalin A-like lectin/glucanase	Beta-1,3-endoglucanase [Bursaphelenchus xylophilus]	2e-45	2,09	7,25	84,84	189,99	65,55	31,73	76,54	20,13
BUX.gene.s01254.155	N	Unknown domain	No match		1,76	4,91	46,6	68,59	67,51	15,92	34,81	8,6
BUX.gene.s01066.2	Y	Glycoside Hydrolase, Lisozyme protein	Hypothetical protein LOAG_03788 [Loa loa]; putative lysozyme [Caenorhabditis brenneri]	1e-10	2,59	3,5	26,39	40,06	89,35	6,49	74,48	0,82
BUX.gene.s00713.953	Y	Peptidase aspartic	No match		13,97	22,01	207,66	149,43	303,25	35,55	336,62	26,21
BUX.gene.s01653.523	N	Cation-transporting ATPase 13A1	Putative cation-transporting atpase [Ascaris suum]	0,0	0,15	3,13	0	0	0	1,7	4,59	5,06
BUX.gene.s01416.2	N	Unknown domain	No match		3,3	27,82	1	0,88	0,39	4,11	1,36	2,73
BUX.gene.s00119.44	Y	Glycoside hydrolase, family 45	Beta-1,4-endoglucanase [Bursaphelenchus xylophilus]	3e-161	44,67	112,37	620,04	413,45	860,2	230	207,88	248,15
BUX.gene.s01109.178	Y	Unknown domain	No match		19,71	72,69	298,52	440,9	527,42	253,73	235,07	361,2
BUX.gene.s01662.77	Y	Unknown domain	No match		5,46	15,89	92,04	138,59	96,97	58,58	42,33	73,61
BUX.gene.s01066.75	N	Ryanodine receptor	C. briggsae CBR-UNC-68 protein	3e-11	2,91	3,13	5,18	10,97	7,08	0,9	0,47	0,77
BUX.gene.s00116.883	N	Unknown domain	No match		0,62	2,32	5,3	11,4	10,2	1,74	0,91	0,98
BUX.gene.s01144.122	Y	Unknown domain	No match		10,02	14,36	53,92	375,84	222,77	41,95	39,05	33,44
BUX.gene.s01226.30	N	Carboxylesterase, type B	Gut esterase 1 [Toxocara canis]	2e-63	3,43	14,17	70,7	136,39	107,88	15,2	29,78	31,96
BUX.gene.s00036.52	Y	Unknown domain	No match		0,69	0,22	18,96	8,79	14,72	1,72	6,57	1,26
BUX.gene.s00351.456	N	Peptidase aspartic	aspartic protease [Steinernema carpocapsae]	2e-06	2,82	5,92	37,02	64,54	82,07	6,88	20,46	7,11
BUX.gene.s00532.10	N	Peptidase A1	LPXTG-domain-containing protein cell wall anchor domain [Bacillus cereus]	7e-08	8,75	7,89	135,22	223,15	272,05	22,1	95,76	21,12

Table S2 - Summary of RNAseq data.

File name	Sample description	Pairs of reads	Trimmed for mapping
804_LIB5929_LDI4875_CGATGT_L007	Fungal Feeder (pre-invasive)	21 150 271	15 944 177
804_LIB5930_LDI4876_TGACCA_L007	Fungal Feeder (pre-invasive)	34 234 980	27 264 218
804_LIB5931_LDI4877_ACAGTG_L007	6 days post infection	21 283 531	17 109 668
804_LIB5932_LDI4878_CAGATC_L007	6 days post infection	21 387 006	17 026 799
804_LIB5933_LDI4879_AGTCAA_L007	6 days post infection	19 925 627	15 861 485
804_LIB5934_LDI4880_ATGTCA_L007	15days post infection	20 659 630	16 265 316
804_LIB5935_LDI4881_CCGTCC_L007	15days post infection	27 818 544	21 833 984
804_LIB5936_LDI4882_GTGAAA_L007	15days post infection	21 960 106	17 309 392
	TOTAL	188 419 695	148 615 039

Table S3 - List of pair of primers used for amplification of probes for *in situ* hybridisation. Gene model according to Kikuchi *et al*. (2011) and sequences available at http://www.genedb.org/Homepage/Bxylophilus. bp, base pair.

Gene model	Forward and reverse primers (5´>3')	Expected PCR product size (bp)
BUX.s00139.64	ATGGCTATGAAGATCTTCGTCTGCG	257
	TCCATCGAACGTTTCTGGGA	
BUX.s01518.15	TCGCTTCAACGATAGCTGAG	240
	CCATTACATTCGGCGTCGGG	
BUX.s00116.606	TTGTCTTGGCCTTGTGTTGTGCT	257
	TGCCAACGTAATCAGATGTC	
BUX.s00713.164	TTCTGCTTTTTTGTGGACTTGGC	247
	GCTCAGAAACGAATCTGCGT	
BUX.s00083.48	GAACCCGATCATAACGAGGA	213
	GTCAATGTCGTCGGGAACTT	
BUX.s01144.121	GACGGTGATTTATGGCGAGT	258
	CAAGTCATCCCACTCATGAA	
BUX.s01144.130	ACCCGTGACAAAATTGAAGC	160
	CGTCCGTCGGTCTATGTTTT	
BUX.s01281.17	TGCAAAATGCAGAGCAAATC	167
	GCTCATATGGCCGTAGTGGT	
BUX.s01662.75	AGAATGTCTGCCGCTTCTGT	169
	GCTCAAATTGGCCTTCGATA	
BUX.s01109.133	AAAGTCCCCGGATCACCTAC	225
	CGACCAGCAACATGATCTGG	
BUX.s01092.144	GACCTGCTGCAAATATCGCT	262
	GCTCTCCCCGTTTACCATCT	
BUX.s01066.2	AAGGAGCGGTGGATAAGGTC	225
	TGAATCTGGCGTTGTTGGTC	
BUX.s01144.188	ATTATCGCGCAGGATTCAAC	234
	TAAACACGCCATCGAAATCA	
BUX.s00036.52	TAGCGGCTACCTTGGAGAAG	228
	CCGATGGTTGTTTGCAGTGA	
BUX.s00422.202	CCGTTGTCTGGTTGGGTTTT	197
	GTTTCGTATTGGGCGTGGTT	
BUX.s00422.201	TTTACGATGGGCTGACGACT	236
	CTCTCCCAAGGCTCTTTGGA	
BUX.s01661.67	CTGATCCACAACTCCCAGGT	220
	TGTGGGAGCATAGACGGATC	
BUX.s00729.2	TTCAATTCGGCGCTTATACC	240
	TGACCAACTCAAGCTCATCG	
BUX.s00647.68	AGCCCTAAACTTCAAGGCCT	226
	ATTGGGCTTTCTTGGCTTCC	
BUX.01109.245	GCGTTGATGCCATCTCTTAC	247
	TGCTCTTCTTTCTGCTTCAGC	
BUX.s00139.163	TCACTCAAACATGGCCAACG	197
	CATTGCGTTCTTGACCCCAA	
BUX.s01109.570	TTGTACTGGCGATTGGAGGT	249
	ATCGAGTGAATCCAAAGCGC	
BUX.s01259.69	CTACCGCTGACAAGTGCTCT	290
	TCGGACACTTCTCCGCTTGA	

BUX.s00116.581	TCTTAATTGGACTGTTCGTTTTGGC	150
	CTGAGAGCTCTCATTGGTCG	
BUX.s01337.7	ATGGTGAGGCAACCTCAACG	100
	CTTTATCCTCCGTTCCTCGA	
BUX.s01063.107	TTACTAGTGAGTCATGGCGAGG	229
	TTCAAGTGGCGGAGGATGTA	
BUX.00351.387	TGCCTTATGGCTACTCTCAC	240
	GCTGGGGTAAAATAGGATGC	
BUX.s00116.597	GCGCAGCTGTCGGAAATTAC	280
	TATGATCCGGAGCCGGAGCC	
BUX.s00117.41	CCAACGCCTTGACTGAGTTC	240
	CGCCGGAAACTGTCCCGTAA	
BUX.s00351.404	AACGCTCGGACAGTCTGAAT	214
	CGATCAGAGTCGACAACGAA	
BUX.s00460.319	TACTCTTAAATGCGGCGAGAAT	250
	GTCCTCATCATTTCCGAGTT	
BUX.s00532.10	GGCTCTACTGAGCCTTGCGA	230
	CTCCGGTAGCGCTTGGATCG	
BUX.s01147.177	CCTCTGCCGGATTGCTGGAT	284
	CCGAAGCGTTGCGAGGTGGT	
BUX.s01518.92	CAGGTGCTCAATTCTTCCCT	240
	GCAACTTTTCCTGCACGAGTTC	
BUX.s01662.77	CCACGGCCGATGACTTTCCA	230
	GCTGCTAACGTGACCCGACA	
BUX.s00600.45	ATGAGCACCGGGAAGAAAGTGTG	247
	TCGGCGTAGCGATCAAGGTA	
BUX.s00647.112	GCTGGTCTTCTGACAGCGTCGTA	340
	CCAATGATTGCTCGATACAGGGT	
BUX.s00116.699	CTAAGTAAGTATCCCTCTAGCGA	260
	AACGAAAATCTCCGCTGATT	
BUX.s00116.698	GAAATATCCTGGCCACGGCCAAG	463
	ACCACAACGTCTGCAGAGGTGGC	
BUX.s00422.680	GAGATTTCTAGCTCTACTATTC	303
	ACTTGGGGTCTAAAGTATTG	
BUX.s00298.34	GCCTGGCGCTCTTGACCTACAAT	304
	TCAGCTTCTTTTCGTGGACTCTCC	
BUX.s01109.178	GTTGGCTGTTTCAGCCCTAG	290
	CTTCAGCTTCTCCTCCGAAC	
BUX.s00713.953	TAACCGTCTTCCTTCTGCTG	290
	TAGGCGTCGTTGAATTGCCT	
BUX.s01144.122	TGGTTGTTCCAGGGCAGTCG	283
	TGAGCTGGAGGCAACAGCATGGC	
BUX.s01254.333	GTAGAGCCAAGCCATCTTGC	241
	CAGTCGCTTCTGTTGTTGGA	
BUX.s01281.37	GAACTCGGAGGTTTTGGTCA	177
	TTCGTCCGATCCAAAGAATC	
Bx-act	CGAGAAGTCCTATGAACTTC	300
	CACATCTGTTGGAAGGTGGAC	

CHAPTER III

TRANSCRIPTIONAL AND MORPHOLOGICAL CHANGES IN THE TRANSITION FROM MYCETOPHAGOUS TO PHYTOPHAGOUS PHASE IN THE PLANT-PARASITIC NEMATODE *B. XYLOPHILUS*

Data published in:
Isheng J. Tsai, Ryusei Tanaka, Natsumi Kanzaki, Mitsuteru Akiba, Toshiro Yokoi, Margarida Espada, John T. Jones, Taisei Kikuchi
Molecular Plant Pathology, 2016, 17(1),77-83

Summary

Drastic physiological and morphological changes in parasites are crucial for establishing a successful infection. The nematode *B. xylophilus* is the pathogenic agent of pine wilt disease and little is known about physiology and morphology in this nematode at the initial stage of infection. In this study, we devised an infection system using pine stem cuttings that allowed us to observe transcriptional and morphological changes in the host infecting phytophagous phase. We found 60 genes enriched in xenobiotic detoxification were both upregulated in two independent post-inoculation events, while down-regulation was observed in multiple members of collagen gene families. After 48 hours of inoculation, tails in some of adult females exposed to the host changed in morphology. These results suggest that *B. xylophilus* may change their physiology and morphology to protect themselves and to adapt to host pine wood environment.

Introduction

Pine wilt disease is one of the most serious forest pests in the world and has been responsible for timber losses of over 2 million cubic metres in some years (Jones *et al.*, 2013). The causal agent of pine wilt disease is the nematode *B. xylophilus*. This nematode is thought to be indigenous to North America but has subsequently been spread to East Asia and European countries (Jones *et al.*, 2008; Robertson *et al.*, 2011, Fonseca *et al.*, 2012). *B. xylophilus* has mycetophagous and phytophagous feeding stages in its life cycle. Nematodes are transmitted to healthy trees by a vector insect during maturation feeding of the insect.

The nematodes spread through the tree's vascular system and resin canals, feeding on epithelial cells and living parenchyma and start to reproduce. This phytophagous phase results in disruption of plant tissues leading to a lethal wilt within months of infection under certain environmental conditions. At this stage, fungi start colonizing the tree and the mycetophagous phase begins. When fungi become limiting, the nematodes locate pupae of their insect vectors and enter a survival (dauer) stage which invade the tracheal system of the insects and is subsequently transmitted to the next plant when the adult insect emerges (Jones et al., 2008).

The nutrient and environmental conditions encountered by the nematodes at the two phases of the life cycle vary. In the phytophagous phase the nematodes are resident in living plant material and exposed to a variety of pre-existing structural and chemical defences as well as induced responses. These are likely to include phenols, terpenes, and Reactive Oxygen Species (ROS) (Holscher *et al.*, 2014; Torres *et al*., 2006; Abdel-Rahman *et al*.; 2013; Sun *et al*., 2011). These threats are present throughout the phytophagous phase (i.e. until host death) because *B. xylophilus* moves continuously to fresher parts of the plant. To evade such host

defence responses, many parasites are known to change their morphology and physiology. This is reflected in changes in transcriptional patterns in a wide range of plant and animal parasitic nematodes such as *Globodera pallida* (Palomares-Rius *et al*., 2012), *Strongyloides stercoralis* (Stoltzfus *et al*., 2012), Haemonchus contortus (Jasmer *et al*., 2004), and *Ancylostoma caninum* (Williamson *et al*., 2006). Morphological changes that occur in response to the onset of parasitism can include changes to the surface coats (Akhkha et al., 2002, Proudfoot et al., 1993, Lopez de Mendoza et al., 2000).

In this study, we compared morphology and transcriptomes of mycetophagous and phytophagous *B. xylophilus* using a stem-cutting inoculation system. We show that the two phases have distinct morphological characters and have identified genes whose expression patterns suggest that they may be involved in these morphological changes. In addition, we show that the nematode undergoes rapid changes in gene expression in response to changes in its environment.

Results

Morphology changes

To imitate the environment that the nematodes are in the phytophagous phase (moving and feeding in a fresh part of the tree), we devised a stem cutting inoculation system as shown in Figure 1 (see Experimental Procedures). Nematodes were observed in the bottom of tubes about 30 min after inoculation and continued to be observed until 3 days post inoculation (Figure 2A). We sought to observe any morphological changes between the two stages. In the first 24 hours after inoculation, no morphological changes were observed in the nematodes. However, at 48 and 72 hours post inoculation, we observed that some female 4th stage larvae (L4) and female adults had different morphology in the tails (M-shape) (Figure 3). The proportion of female nematodes showing this change in tail shape increased until 72 hours post inoculation (Figure 2B). No such change was observed in other stages (males and larvae) of *B. xylophilus* (data not shown) and when the nematodes were incubated on fungi or in water (Figure 2B).

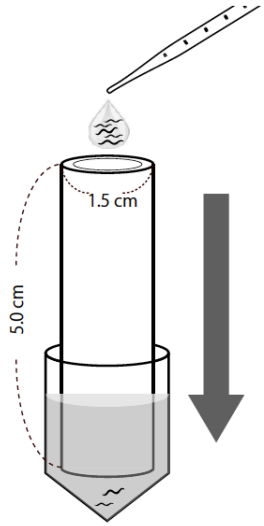


Figure 1- Stem-cutting inoculation system. *Pinus thunbergii* seedlings were cut (diameter, 1 cm; length, 5 cm) and placed in a plastic tube filled with distilled water. The nematodes were placed on top of the cutting and were collected from the water at the bottom at the appropriate times.

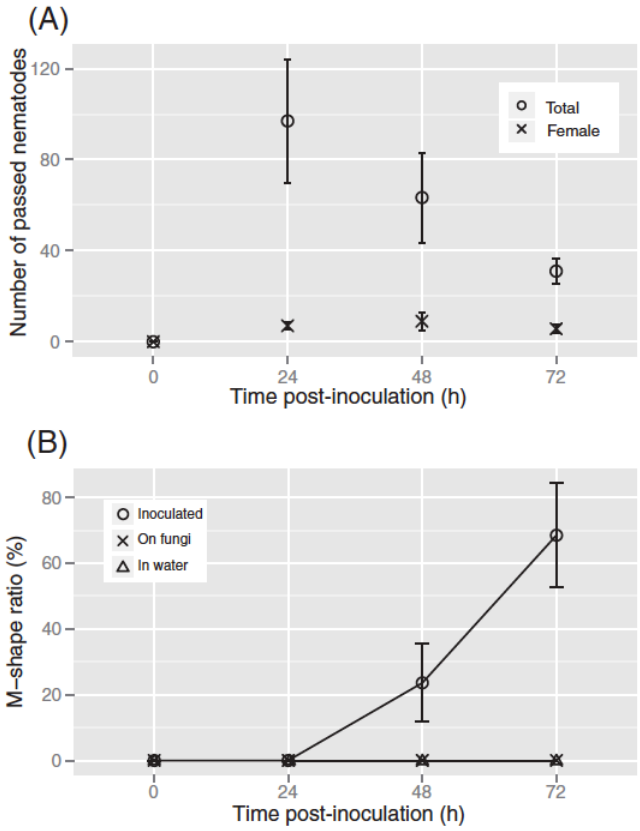


Figure 2- (A) Number of nematodes isolated from the stem cuttings at different time points after inoculation. Twenty microlitres of water containing 2000 *B. xylophilus* Ka4 were inoculated on top of the 5-cm-long stem cuttings (Fig. 1). The inoculated stem cuttings were incubated at 25 °C and 100% relative humidity. The numbers of nematodes which came through the stem cuttings were counted every 24 h. The number of females observed in each time interval is also shown in the graph. (B) Change in ratio of mucronated (M-shape) nematodes. Tips of females were observed under a microscope for the same samples as in (A). Nematodes incubated in water or growing on fungi were examined as controls. Error bars represent the standard error of the mean.

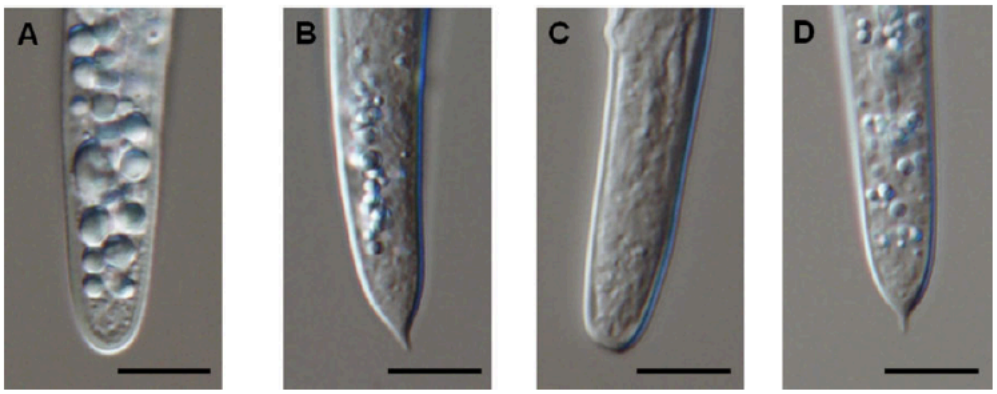


Figure 3 - Tail morphology of B. xylophilus Ka4 before and after inoculation. (A) L4 female with a round (R-shape) tail (in water). (B) L4 female with a mucronated (M-shape) tail (48 h after inoculation). (C) Adult female with a round (R-shape) tail (in water). (D) Adult female with a mucronated (M-shape) tail (48 h after inoculation). All scale bars are 10 µm.

RNAseq of *B. xylophilus*

To identify differentially expressed genes when *B. xylophilus* enters the phytophagous phase, and to account for additional factors associated with differential environmental conditions, RNA was extracted from mixed stage nematodes incubated for 2 hours in water and extracted 0.5-2.5 hours after inoculation at two different times of the year (August and September). The experimental design therefore used two post inoculation events and one control with two biological replicates per condition. The relative expression of *B. xylophilus* genes was then quantified by Illumina RNA sequencing. A total of 144.3 million 100 bp paired-end sequence reads were generated and on average 92% of these mapped to the *B. xylophilus* reference genome (Table S1). Differential expression of transcripts was calculated between pairwise comparisons of each condition. To avoid false positives on extremely lowly expressed genes, transcripts with RPKM of less than 1 in any of the conditions were excluded from further analysis. This still left 12,851 of 17,704 genes available for the analysis, and the full list of transcript RPKM counts is shown in Table S2.

The two biological replicates in each condition were clustered together indicating that the experiments were highly reproducible and gene expression in *B. xylophilus* is indeed influenced by life cycle stage and environmental factors (Figure S1). Our experimental setup allowed us to distinguish 1,143 genes that were upregulated in post inoculation samples (Figure 4). However, many of these genes were upregulated specifically in one of the two environmental samples (1061 and 22 genes in August and September, respectively). These genes may be differentially regulated due to environmental differences, for example, genes only upregulated in the August post-inoculation event was enriched in the embryo development, translation and positive growth rates (Table S3), suggesting that nematodes in this particular setup had a more optimal growth condition compared to the control and the September post-inoculation event.

Sixty and 384 genes were upregulated or downregulated in both post inoculation samples when compared to the control (Figure 4), respectively. It is possible that some of these differentially expressed genes underlie the mucronated morphological changes observed here (Figure 3). Indeed, GO term enrichment of down-regulated genes reveals a dominant term of "morphological change" when nematodes enter the phytophagous stage (Table S4). Within this term, collagen gene family members are predominant in the list (Table S5). Collagens are important in structural formations and modification in a range of species and are the dominant protein of the cuticle (Johnstone, 2000). Using qPCR, we further validated the downregulation of four *B. xylophilus* collagen genes after infection (Figure 5). In particular, col-5 shows a 16.7 fold decrease in expression. Ten heat shock proteins were also down-regulated after infection (Table S5), which may be responsible in dealing with osmotic stress when the nematodes are incubated solely in water. For the 28 of the 60 up-regulated genes that have functional annotations (Table S6), we found that 12 genes were largely involved in different phases of xenobiotic metabolism, including cytochrome P450, short chain dehydrogenase, UDP-glucuronosyl transferase and glutathione S transferase (GST). This increase in expression after inoculation may be involved in dealing with host derived stress (Lindblom and Dodd, 2006). In addition, we also identified 5 proteases that were up-regulated and may be involved in other aspects of the parasitic process, such as host tissue penetration. Interestingly, we further identified copies of P450 and proteases up-regulated only in the September post-inoculation event, again suggesting the more optimal growth conditions in the August post-inoculation event.

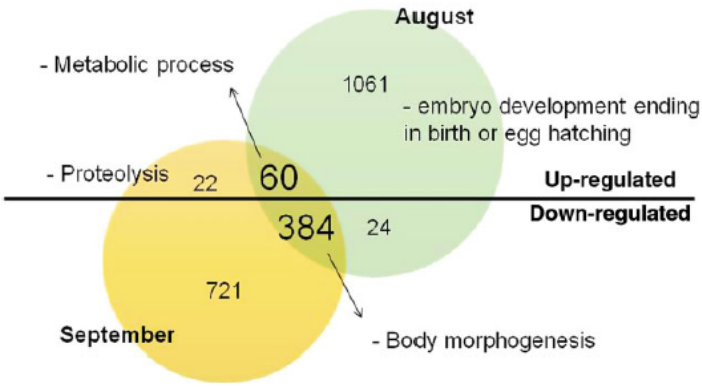


Figure 4- Venn diagram showing the up-regulated and down-regulated genes after infection with *B. xylophilus*.

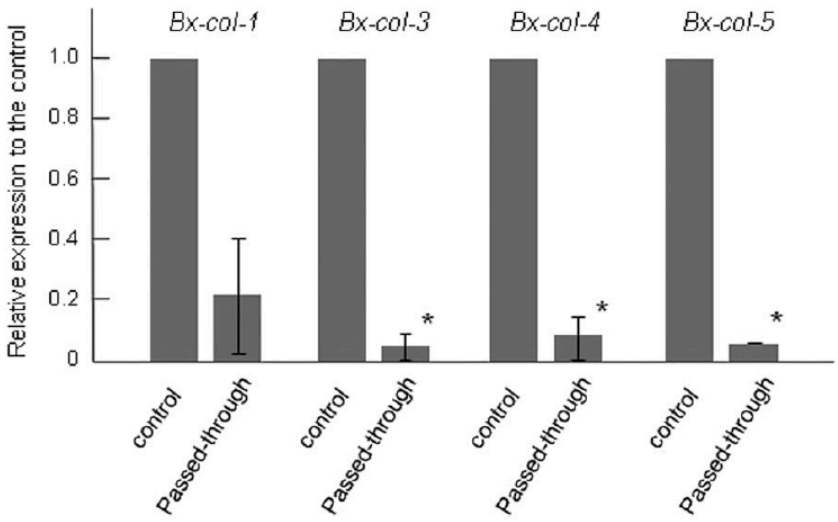


Figure 5 - Relative expression levels of collagen genes measured by quantitative polymerase chain reaction (qPCR). Expression of collagen genes (Bx-col-1, Bx-col-3, Bx-col-4 and Bx-col-5) was compared between the control and after inoculation (Passed-through). Error bars represent the standard deviations of the biological replicates. Asterisks indicate statistically significant differences between the two conditions (Student's t-test, $P < 0.05$).

Discussion

Plants produce a variety of chemicals to protect themselves from pathogens as part of the induced defence response (O'Brien *et al*., 2012). Reactive Oxygen Species (ROS) are thought to be a major part of these defense mechanisms and are known to be deployed against plant parasitic nematodes (Mehdy, 1994; Bolwell *et al*., 1995; Torres *et al*., 2006) as well as resins and reinforced cell walls. These mechanisms are thought to make nematode survival and movement difficult (Futai, 2013). However, *B. xylophilus* can survive, move and grow under these conditions in both living host trees (phytophagous) and in the dead host (mycetophagous) (Futai, 2013). In this study, we have employed an experimental setup to retrieve *B. xylophilus* as it is just entering the phytophagous phase. We have shown that various strategies are employed by *B. xylophilus* in order to establish a successful infection.

Many pathogens are known to change morphology when they infect the hosts. For example, heteroecious parasites, Toxoplasma, Plasmodium and Theileria change their morphology and physiology hosts by hosts (Plattner and Soldati-Favre, 2008). For the first time, we also observed morphological changes in *B. xylophilus* in the tails of L4 and adult females 48 hours after inoculation which were not seen in nematodes that were not exposed to the hosts. This suggests that the presence of the host is required to stimulate this specialised molting. The change in tail shape in females may be part of a response of the whole body surface as it adapts to host pine wood environment. Changes in tail structure were not observed across all life stages; in the case of the larvae, changes were not clear because of their small tail sizes. In addition, the adult male has a bursa at the end of tail and the presence of this structure would

obscure the morphological changes seen in females. Observation of these changes is important as the diagnosis of pine wilt disease is performed mainly by direct observation of nematodes isolated from pine trees, though some molecular techniques have been used as alternatives (Kikuchi *et al*., 2009). The discrimination of this nematode from other non-pathogenic nematodes is sometimes difficult. In particular, a non-pathogenic species *B. mucronatus*, which is closely related to *B. xylophilus*, shares most morphological characters with the pathogenic nematode except for female tail morphologies, i.e., rounded in *B. xylophilus* and mucronated (tipped) in *B. mucronatus*; such differences have been used as one of the main diagnostic characters (Kanzaki, 2008). *B. xylophilus* individuals with mucronated tail were sometimes observed i) in dead tree wood with high nematode population density, and ii) in newly killed or dying trees regardless of population density (data not shown). This observation can be explained by our observations in which adult *B. xylophilus* females in phytophagous phases also have mucronated tails.

Many members of the collagen gene family were downregulated in *B. xylophilus* 48 hours after inoculation. Low expression levels of collagen genes seem to be a common feature among infective life stages of various plant and animal parasitic nematodes (Mitreva *et al*., 2004, Elling *et al*., 2007). As collagens make up the majority of cuticle in many nematodes and expression of collagen genes lead to changes in nematode morphology (Johnstone, 2000), down regulation of this gene family may reflect changes in cuticle shape and structure. These changes may reflect the changes in morphology seen here, or alternatively, may reflect synthesis of new cuticle components designed to cope with the stresses of being inside a host. Clearly it remains to be determined to whether these collagens were expressed in specific parts of *B. xylophilus* and how long those downregulations continue. This may be a specific response of initial contact to the host plant, because in our preliminary RNAseq data from nematodes of 6 and 15 days post inoculation such downregulations were no longer observed (Margarida Espada and John Jones, unpublished result).

A likely explanation to the observation of major differences in gene expression between inoculations in August and September is underpinned by differences in the environmental conditions of the trees from which the cuttings were made. As tree cuttings were prepared from trees grown in the experimental nursery, the physiological conditions of the plants, which include immune activity and nutritional status, may be very different due to temperature, water and other environmental conditions. Furthermore, because the same procedures were used to prepare the two inoculums and there were two replicates in each stage of the RNAseq analysis, it is unlikely that condition differences between the two inoculums (e.g. stage ratio) caused the major differences in gene expression between inoculations in August and September. Our data, therefore, suggest that when the nematodes invade the tree they react flexibly according to the host conditions and may use different strategies to survive. In spite of these conditions, we identified 60 genes that were upregulated in both August and September.

These genes may be essential for the early stages of infection under all conditions and include several xenobiotic genes. Indeed, expression of genes encoding antioxidant proteins may be a response to host derived ROS. These genes were identified here and were also expressed *B. xylophilus* 7 days after infection in pine trees (Qiu *et al*., 2013). *B. xylophilus* is known to produce surface coat proteins that help them protect themselves from ROS (Shinya *et al*., 2013). Coping with host derived ROS is therefore likely to be a key factor underpinning *B. xylophilus* survival in the host tree.

A recent proteomic study in *B. xylophilus* by Shinya *et al*. (2010) compared surface coat proteins of nematodes isolated from plant after 15 days of infection with those grown on fungus, and the results were found to be consistent with our August post-inoculation condition. The authors have identified 12 proteins overrepresented in the phytophagous stage nematodes, which included glutathione S-transferase (GST), 14-3-3b protein and glyceraldehyde-3-phosphate dehydrogenase (Shinya *et al*., 2010), of which 10 were also up-regulated in our August phytophagous nematodes (Table S2). Additionally, all of the 4 surface coat proteins underrepresented in the phytophagous stage, including paramyosin and enolase (Shinya *et al*., 2010), were down-regulated in our August samples (Table S2). Conversely, only 1 out of 12 overrepresented and 3 out of 4 underrepresented proteins were up- or down-regulated, respectively, in our September RNAseq. These results suggest the conditions they used for the protein comparisons were similar to that of our August samples and the high degree of consistency between two studies (protein and RNAseq) emphasise the reliability of the findings in both studies. To conclude, we show that *B. xylophilus* undergo physiological changes to protect themselves from the host environment in the first stages of infection. In addition, morphological changes in the tail of female nematodes occur after infection. These results will assist in accurate diagnosis of pine wilt nematodes and controlling of pine wilt disease.

Experimental procedure

Nematodes used in this study

The *B. xylophilus* Ka4 and Ka4C1 strains (maintained in the Pathology laboratory of FFPRI Japan) were used in this study. Ka4C1 was generated from Ka4 by inbreeding for genome sequencing and RNAseq (Kikuchi *et al*., 2011). Nematodes were cultured for 7 days at 25°C on *Botrytis cinerea* grown on autoclaved barley grains with antibiotics (100 µg/ml streptomycin and 25 µg/ml chloramphenicol). The nematodes were then collected using the Baermann funnel method for 1 h at 25°C. The nematodes were washed three times in 0.5x phosphate buffered saline (PBS) before use in the experiments.

Stem cutting inoculation

Shoots (of diameter approx 1 cm) obtained from 5-year old Japanese black pine (*Pinus thunbergii*) were cut into 5 cm long sections and used for experiments immediately. Twenty µl

of water containing 2,000 mixed life-stage nematodes were inoculated on the top of the stem cuttings which were placed in 50 ml plastic tubes containing 4 ml distilled water (Figure 1). A small pit was made on the top of each stem cutting prior to nematode inoculation and the solution was carefully applied to this pit. The inoculated stem cuttings were incubated at 25°C and 100% relative humidity. Nematodes that came through the stem cuttings were collected from the bottom of the 50 ml tubes at various time points (0.5, 1, 2, 3, 6, 12, 24, 48, and 72 hours post inoculation) and were observed under a compound microscope (ECLIPSE 80i, Nikon). This system allowed nematodes to be obtained easily and quickly from the bottom of the tubes without requiring any disruption or maceration of the pine trees and we considered that this procedure would minimize any artificial effects during the recovery process. For control, 20 µl of water containing 2,000 nematodes were inoculated on *B. cinerea* grown on PDA or 4 ml of distilled water in glass tube, and incubated at 25°C. Nematodes were observed in the same way as the stem cutting experiments.

RNA-seq dataset generation

Approximately 40 000 nematodes were recovered from the base of 20 stem cuttings, as described previously, between 0.5 and 2.5 h after inoculation. The nematodes were pooled in a 1.5-mL tube and used for RNA extraction immediately or stored at −80 °C until use. After disruption of nematode bodies using zirconia beads (ø = 0.15–0.40 mm) in 500 µL of TRIzol (Invitrogen, Tokyo, Japan), total RNA was extracted according to a standard Trizol protocol (Invitrogen). Total RNA was also extracted from approximately 40 000 nematodes which were incubated in water at 25 °C for 2 h and used as a control. After RNA quality and quantity assessment using a Bioanalyser2100 (Agilent, Santa Clara, USA), 1 µg of total RNA was used to construct an Illumina sequencing library employing the TruSeq RNA-seq Sample Prep kit according to the manufacturer's recommended protocols (Illumina, San Diego, USA). The libraries were quantified by qPCR and sequenced for 200 cycles (100-bp paired ends) on an Illumina HiSeq2000 sequencer using the standard protocol (Illumina). RNAseq reads were mapped against the B. xylophilus genome reference (v1.3) using Tophat v.2.0.11 (Trapnell *et al*., 2009) and differential expression was called using EdgeR v3.2.4 (Robinson *et al*., 2010). A transcript was identified as differentially expressed in a pairwise comparison if the following criteria were met: false discovery rate (FDR) ≤ 0.01 and fold change ≥ 2.0. RNAseq experiments were conducted in duplicate for the test conditions and in triplicate for the control condition.

qPCR

To confirm the expression levels of collagen genes, 1000 mixed-stage nematodes (Ka4C1), which were either incubated in water (control) or collected from stem cuttings as described above, were used for qPCR. Total RNA was extracted from these nematodes using Trizol (Invitrogen), treated with DNase (TakaraBio, Otsu, Japan) and dissolved in 20 µL of water. Single-stranded cDNA was synthesized from 5 µL of the total RNA solution using the iScript cDNA Synthesis Kit according to the manufacturer's instructions (Bio-Rad,Tokyo, Japan). Full-

length sequences of four collagengenes (Bx-col-1, Bx-col-3, Bx-col-4 and Bx-col-5) were manually curated using the genome assembly (v1.3; available from http://parasite.wormbase.org/index.html) and the RNAseq data. Primers for target genes were designed using Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/) (Table S7, see Supporting Information). The qPCRs were performed using StepOnePlus (Applied Biosystems, Foster City, USA) with Power SYBR reagents (Applied Biosystems) in a reaction volume of 10 μL containing 5 μL of Power SYBR Green PCR Master Mix (2X), 0.5 μL of cDNA solution and 0.9 μM of each primer under the following conditions: 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. All qPCRs were conducted with two biological replicates, each having three technical measurements. The actin gene (Bxact-1) was used as an endogenous control. Relative transcript levels of the two samples were calculated using StepOne Software v2.3 (Applied Biosystems). Statistical analyses were performed using R packages (http://www.r-project.org/).

References

Abdel-Rahman, F.H., Alaniz, N.M. and Saleh, M.A. (2013) Nematicidal activity of terpenoids. *J. Environ. Sci. Health B*, 48, 16–22.

Akhkha, A., Kusel, J., Kennedy, M. and Curtis, R. (2002) Effects of phytohormones on the surfaces of plant-parasitic nematodes. *Parasitology* 125, 165–175.

Bolwell, G.P., Butt, V.S., Davies, D.R. and Zimmerlin, A. (1995) The origin of the oxidative burst in plants. *Free Radic. Res*. 23, 517–532.

Elling, A.A., Mitreva, M., Recknor, J., Gai, X., Martin, J., Maier, T.R., McDermott, J.P., Hewezi, T., Mc, K.B.D., Davis, E.L., Hussey, R.S., Nettleton, D., McCarter, J.P. and Baum, T.J. (2007) Divergent evolution of arrested development in the dauer stage of Caenorhabditis elegans and the infective stage of Heterodera glycines. *Genome Biol.* 8, R211.

Fonseca, L., Cardoso, J., Lopes, A., Pestana, M., Abreu, F., Nunes, N., Mota, M. and Abrantes, I. (2012) The pinewood nematode, *B. xylophilus*, in Madeira Island. *Helminthologia* 49, 96–103.

Futai, K. (2013) Pine wood nematode, *B. xylophilus*. *Annu. Rev. Phytopathol*. 51, 61–83.

Holscher, D., Dhakshinamoorthy, S., Alexandrov, T., Becker, M., Bretschneider, T., Buerkert, A., Bretschneider, T., Buerkert, A., Crecelius, A.C., De Waele, D., Elsen, A. and Heckel, D.G. (2014) Phenalenone-type phytoalexins mediate resistance of banana plants (*Musa* spp.) to the burrowing nematode *Radopholus similis*. *Proc. Natl. Acad. Sci. USA* 111, 105–110.

Jasmer, D.P., Mitreva, M.D. and McCarter, J.P. (2004) mRNA sequences for Haemonchus contortus intestinal cathepsin B-like cysteine proteases display an extreme in abundance and diversity compared with other adult mammalian parasitic nematodes. *Mol. Biochem. Parasitol.* 137, 297–305.

Johnstone, I.L. (2000) Cuticle collagen genes: expression in *Caenorhabditis elegans*. *Trends Genet.* 16, 21–27.

Jones, J.T., Moens, M., Mota, M., Li, H. and Kikuchi, T. (2008) *B. xylophilus*: opportunities in comparative genomics and molecular host–parasite interactions. *Mol. Plant Pathol.* 9, 357–368.

Jones, J.T., Haegeman, A., Danchin, E.G.J., Gaur, H.S., Helder, J., Jones, M.G.K., Kikuchi, T., Manzanilla-López, R., Palomares-Rius, J.E., Wesemael, W.M.L. and Perry, R.N. (2013) Top

10 plant-parasitic nematodes in molecular plant pathology. *Mol. Plant Pathol.* 14, 946–961.

Kanzaki, N. (2008) Taxonomy and systematics of the nematode genus Bursaphelenchus (Nematoda: Parasitaphelenchidae). In: Pine Wilt Disease (Zhao, B.G., Futai, K., Sutherland, J.R. and Takeuchi, Y., eds), pp. 44–66. Springer, Tokyo.

Kikuchi, T., Aikawa, T., Oeda, Y., Karim, N. and Kanzaki, N. (2009) A rapid and precise diagnostic method for detecting pinewood nematode *B. xylophilus* by loop-mediated isothermal amplification. *Phytopathology*, 99, 1365–1369.

Kikuchi, T., Cotton, J.A., Dalzell, J.J., Hasegawa, K., Kanzaki, N., McVeigh, P., Takanashi, T., Tsai, I.J., Assefa, S.A., Cock, P.J.A., Otto, T.D., Hunt, M., Reid, A.J., Sanchez-Flores, A., Tsuchihara, K., Yokoi, T., Larsson, M.C., Miwa, J., Maule, A.G., Sahashi, N., Jones, J.T. and Berriman, M. (2011) Genomic insights into the origin of parasitism in the emerging plant pathogen *B. xylophilus*. *Plos Pathog*. 7, e1002219.

Lindblom, T.H. and Dodd, A.K. (2006) Xenobiotic detoxification in the nematode Caenorhabditis elegans. *J. Exp. Zool. Part A Comp. Exp. Biol.* 305, 720–730.

Lopez de Mendoza, M., Modha, J., Roberts, M., Curtis, R. and Kusel, J. (2000) Changes in the lipophilicity of the surfaces of *Meloidogyne incognita* and *Haemonchus contortus* during exposure to host signals. *Parasitology* 120, 203–209.

Mehdy, M.C. (1994) Active oxygen species in plant defense against pathogens. *Plant Physiol.* 105, 467–472.

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

O'Brien, J.A., Daudi, A., Butt, V.S. and Bolwell, G.P. (2012) Reactive oxygen species and their role in plant defence and cell wall metabolism. *Planta*, 236, 765–779.

Palomares-Rius, J.E., Hedley, P.E., Cock, P.J., Morris, J.A., Jones, J.T., Vovlas, N. and Blok, V. (2012) Comparison of transcript profiles in different life stages of the nematode *Globodera pallida* under different host potato genotypes. *Mol. Plant Pathol*. 13, 1120–1134.

Plattner, F. and Soldati-Favre, D. (2008) Hijacking of host cellular functions by the Apicomplexa. *Annu. Rev. Microbiol*. 62, 471–487.

Proudfoot, L., Kusel, J., Smith, H., Harnett,W., Worms, M. and Kennedy, M. (1993) Rapid changes in the surface of parasitic nematodes during transition from pre- to post-parasitic forms. *Parasitology*, 107, 107–117.

Qiu, X., Wu, X., Huang, L., Tian, M. and Ye, J. (2013) Specifically expressed genes of the nematode *B. xylophilus* involved with early interactions with pine trees. *Plos ONE*, 8, e78063.

Robertson, L., Cobacho Arcos, S., Escuer, M., Santiago Merino, R., Esparrago, G., Abelleira, A. and Navas, A. (2011) Incidence of the pinewood nematode *B. xylophilus* Steiner and Buhrer, 1934 (Nickle, 1970) in Spain. *Nematology* 13, 755–757.

Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) EdgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140.

Shinya, R., Morisaka, H., Takeuchi, Y., Ueda, M. and Futai, K. (2010) Comparison of the surface coat proteins of the pine wood nematode appeared during host pine infection and in vitro culture by a proteomic approach. *Phytopathology* 100, 1289–1297.

Shinya, R., Morisaka, H., Kikuchi, T., Takeuchi, Y., Ueda, M. and Futai, K. (2013) Secretome analysis of the pine wood nematode *B. xylophilus* reveals the tangled roots of parasitism and its potential for molecular mimicry. *Plos ONE* 8, e67377.

Stoltzfus, J.D., Minot, S., Berriman, M., Nolan, T.J. and Lok, J.B. (2012) RNAseq analysis of the parasitic nematode *Strongyloides stercoralis* reveals divergent regulation of canonical dauer pathways. *Plos Negl. Trop. Dis*. 6, e1854.

Sun, Y., Yin, J., Cao, H., Li, C., Kang, L. and Ge, F. (2011) Elevated CO2 influences nematode-induced defense responses of tomato genotypes differing in the JA pathway. *Plos ONE* 6, e19751.

Torres, M.A., Jones, J.D. and Dangl, J.L. (2006) Reactive oxygen species signaling in response to pathogens. *Plant Physiol*. 141, 373–378.

Trapnell, C., Pachter, L. and Salzberg, S.L. (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111.

Williamson, A.L., Lustigman, S., Oksov, Y., Deumic, V., Plieskatt, J., Mendez, S., Zhan, B., Bottazzi, M.E., Hotez, P.J. and Loukas, A. (2006) *Ancylostoma caninum* MTP-1, an astacin-like metalloprotease secreted by infective hookworm larvae, is involved in tissue migration. *Infect. Immun*. 74, 961–967.

Supporting information

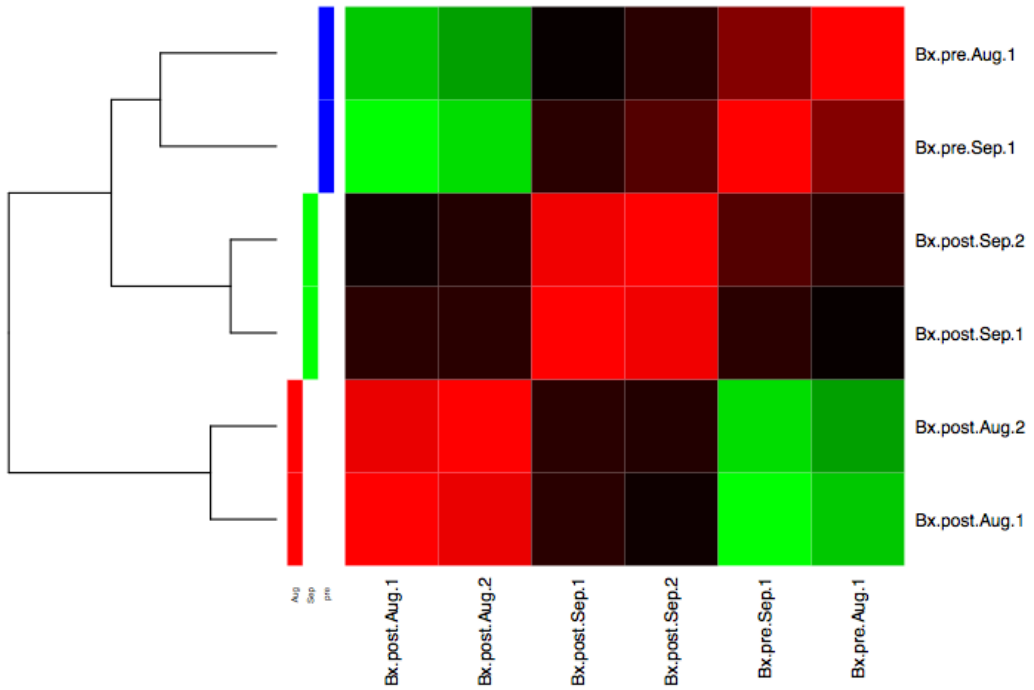


Figure S1 - Heatmap showing the hierarchically clustered correlation matrix by comparing the transcript expression values for each pair of samples.

Table S1 - *B. xylophilus* RNAseq mapping statistics

Samples	Sequence pairs	Total sequences	Aligned pairs	concordant rates (%)
August-1	9 733 258	19 466 516	8 751 583	86,3
August-2	10 751 478	21 502 956	9 831 084	93,9
Spetember-1	10 780 701	21 561 402	10 153 430	93,6
September-2	11 623 332	23 246 664	10 949 710	93,6
Control-1	13 736 064	27 472 128	12 730 368	91,1
Control-2	15 537 470	31 074 940	14 582 149	93,3

Table S2 - RPKMs of transcribed genes of *B. xylophilus* in this study.
(Table has more than 17000 lines; please consult:
http://onlinelibrary.wiley.com/doi/10.1111/mpp.12261/abstract;jsessionid=47BFA7588D1CC9063A965DC8B487D567.f03t03)

Table S3. Enriched GO terms of the solely upregulated genes in the August or the September phytophagous phase of *B. xylophilus* nematodes.

August only					
GO.ID	Term	Annotated	Significant	Expected	p-values
GO:0009792	embryo development ending in birth or egg hatching	1933	204	126,69	2,60E-16
GO:0006412	translation	235	45	15,4	7,90E-16
GO:0040010	positive regulation of growth rate	1076	123	70,52	3,10E-11
GO:0002119	nematode larval development	1530	154	100,27	2,80E-10
GO:0000003	reproduction	1743	166	114,23	1,10E-09
GO:0006825	copper ion transport	11	8	0,72	4,40E-08
GO:0006915	apoptotic process	351	49	23	1,60E-07
GO:0040007	growth	1766	183	115,74	3,60E-06
GO:0018996	molting cycle, collagen and cuticulin-based cuticle	221	32	14,48	1,50E-05
GO:0045039	protein import into mitochondrial inner membrane	4	4	0,26	1,80E-05
GO:0015986	ATP synthesis coupled proton transport	42	12	2,75	2,60E-05
GO:0006898	receptor-mediated endocytosis	543	60	35,59	3,70E-05
GO:0008340	determination of adult lifespan	571	61	37,42	6,60E-05
GO:0006465	signal peptide processing	5	4	0,33	8,60E-05
GO:0006506	GPI anchor biosynthetic process	14	6	0,92	0,00015
GO:0022600	digestive system process	5	3	0,33	0,00253
GO:0048569	post-embryonic organ development	205	17	13,44	0,00425
GO:0045747	positive regulation of Notch signaling pathway	6	3	0,39	0,00482
GO:0042787	protein ubiquitination involved in ubiquitin-dependent protein catabolic process	6	3	0,39	0,00482
GO:0043687	post-translational protein modification	26	6	1,7	0,00572
GO:0006633	fatty acid biosynthetic process	32	8	2,1	0,00577

September only					
GO.ID	Term	Annotated	Significant	Expected	p-values
GO:0006508	proteolysis	642	5	0,99	0,0015
GO:0005980	glycogen catabolic process	5	1	0,01	0,0077

Table S4. Enriched GO terms of the upregulated genes in both phases of *B. xylophilus* nematodes.

Upregulated in phytophagous phase					
GO.ID	Term	Annotated	Significant	Expected	p-values
GO:0008152	metabolic process	3585	19	13,79	0,0003
Downregulated in phytophagous phase					
GO.ID	Term	Annotated	Significant	Expected	p-values
GO:0010171	body morphogenesis	446	25	10,16	1,40E-05
GO:0040011	locomotion	1161	46	26,44	6,10E-05
GO:0006898	receptor-mediated endocytosis	543	25	12,36	0,00048
GO:0002119	nematode larval development	1530	49	34,84	0,00624

Table S5. RPKMs of downregulated genes in phytophagous phase of *B. xylophilus* nematodes. Collagens and heat shock proteins were selected from the complete Table S5 and are presented here. (available at http://onlinelibrary.wiley.com/doi/10.1111/mpp.12261/abstract;jsessionid=47BFA7588D1CC9063A965DC8B487D567.f03t03)

	RPKMs							
	Inoculation							
GeneID	August -1	August -2	September -1	September -2	Control -1	Control -2	Fold	Product Description
BXY_0097000	51,6	41,5	128,6	169,2	970,7	642,1	0,2	cel:COLlagen;b2g:protein col-170
BXY_0097300	277,9	193,3	198,3	285,9	1 017,4	730,4	0,5	cel:COLlagen;b2g:nematode cuticle collagen n-terminal domain containing protein
BXY_0165400	63,9	128,9	560,6	576,4	2 209,6	2 241,4	0,3	cel:Stress Induced Protein;b2g:heat shock protein beta-1
BXY_0329300	9,0	5,6	26,2	35,8	302,3	76,4	0,2	cel:COLlagen;b2g:briggsae cbr-col-125 protein
BXY_0412000	1,6	1,7	2,7	2,5	92,3	18,5	0,1	cel:COLlagen;b2g:nematode cuticle collagen domain-containing protein
BXY_0437200	2,2	2,7	2,3	2,6	59,4	9,0	0,1	cel:COLlagen;b2g:nematode cuticle collagen n-terminal domain containing protein
BXY_0507400	5,4	6,5	6,5	7,5	253,4	32,7	0,1	cel:SQuaT;b2g:nematode cuticle collagen n-terminal domain containing protein
BXY_0586000	977,9	1 078,6	1 892,0	1 929,4	5 938,4	4 457,2	0,6	cel:Heat Shock Protein;b2g:NA
BXY_0640100	8,6	6,1	117,0	133,7	1 543,2	1 518,0	0,1	cel:NA;b2g:heat shock protein 70 b2
BXY_0651700	1,8	5,7	2,1	3,5	114,7	14,9	0,1	cel:COLlagen;b2g:collagen structural probable cuticle from embryo to adult (kd) (col-3)
BXY_0724800	3,6	8,7	5,6	6,8	205,3	23,4	0,1	cel:COLlagen;b2g:collagen structural probable cuticle from embryo to adult (kd) (col-3)
BXY_0768000	5,4	6,9	157,2	208,1	1 417,5	1 254,2	0,1	cel:NA;b2g:heat shock protein 70 b2
BXY_0957900	655,3	596,1	1 115,7	1 245,4	5 206,2	4 062,6	0,4	cel:Heat Shock Protein;b2g:NA
BXY_0974900	2,5	1,5	6,2	5,8	124,5	14,4	0,1	cel:COLlagen;b2g:nematode cuticle collagen domain-containing protein
BXY_1087500	18,5	9,2	12,8	20,6	93,0	53,8	0,4	cel:COLlagen;b2g:cuticle collagen 13
BXY_1152300	4,3	4,1	7,0	3,8	61,5	9,8	0,3	cel:COLlagen;b2g:nematode cuticle collagen n-terminal domain containing protein
BXY_1158600	51,9	24,6	19,0	40,1	546,3	194,9	0,2	cel:COLlagen;b2g:collagen structural probable cuticle from embryo to adult (kd) (col-3)
BXY_1161100	49,7	29,3	55,3	80,8	186,0	152,3	0,6	cel:NH;b2g:heat shock protein 105 kda
BXY_1274600	218,8	401,3	2 707,5	2 383,4	7 487,6	6 525,2	0,4	cel:Stress Induced Protein;b2g:small heat shock protein
BXY_1385000	6,3	4,9	96,3	146,5	1 374,3	194,0	0,2	cel:COLlagen;b2g:nematode cuticle collagen n-terminal domain containing protein
BXY_1563600	2,1	1,5	71,5	88,2	1 066,9	916,0	0,1	cel:NA;b2g:heat shock protein 70 b2
BXY_1672900	36,9	19,6	32,0	50,4	91,7	163,7	0,5	cel:DNaJ domain (prokaryotic heat shock protein);b2g:dnaj homolog subfamily c member 2-like
BXY_1690800	644,7	670,2	775,7	1 119,6	4 715,8	2 575,3	0,4	cel:COLlagen;b2g:collagen structural probable cuticle from embryo to adult (kd) (col-3)
BXY_1702400	10,5	6,5	8,4	13,0	25,4	26,0	0,7	cel:DNaJ domain (prokaryotic heat shock protein);b2g:protein tumorous imaginal mitochondrial-like
BXY_1722300	15,2	14,0	53,1	65,5	460,1	134,1	0,2	cel:COLlagen;b2g:cre-col-77 protein
BXY_1722400	6,7	6,2	14,7	23,9	235,8	24,9	0,2	cel:COLlagen;b2g:protein col-77
BXY_1766200	71,3	37,9	104,6	175,3	1 473,8	368,5	0,2	cel:COLlagen;b2g:collagen structural probable cuticle from embryo to adult (kd) (col-3)

Table S6. RPKMs of upregulated genes in the phytophagous phase of *B. xylophilus* nematodes.

	RPKMs Inoculation							
GeneID	**August -1**	**August -2**	**September -1**	**September -2**	**Control -1**	**Control -2**	**Fold**	**Product Description**
BXY_0039700	125,4	235,3	62,3	50,7	7,6	12,0	24,2	cel:NA;b2g:class 3 lipase protein
BXY_0043900	658,4	591,2	448,2	407,9	182,6	203,3	5,5	cel:NH;b2g:NA
BXY_0096800	82,0	123,8	237,6	61,5	9,2	7,8	29,7	cel:DeHydrogenases, Short chain;b2g:retinol dehydrogenase 16-like
BXY_0111700	9,4	15,3	14,4	9,6	2,7	4,3	6,9	cel:CYtochrome P450 family;b2g:protein cyp-33c9
BXY_0175500	12,1	16,2	14,2	6,2	1,1	1,8	16,4	cel:NA;b2g:NA
BXY_0182200	84,9	89,6	63,2	71,1	15,3	5,1	15,2	cel:ASpartyl Protease;b2g:protein asp- isoform a
BXY_0206000	54,1	133,1	117,0	126,7	33,6	32,4	6,5	cel:ASpartyl Protease;b2g:protein asp- isoform a
BXY_0286000	187,9	170,9	236,3	193,1	68,1	62,7	6,0	cel:NA;b2g:NA
BXY_0304700	92,8	64,6	15,5	20,2	2,5	2,1	42,0	cel:NH;b2g:NA
BXY_0307100	49,9	53,2	19,3	25,2	3,9	1,8	26,1	cel:NH;b2g:NA
BXY_0352600	22,1	55,8	33,2	32,2	15,9	9,3	5,7	cel:NA;b2g:NA
BXY_0483700	53,8	39,9	19,2	21,5	5,6	3,0	15,6	cel:NH;b2g:NA
BXY_0492900	82,2	86,2	263,2	64,9	33,4	34,4	7,3	cel:SOrbitol DeHydrogenase family;b2g:alcohol dehydrogenase 2
BXY_0521800	166,7	167,4	44,4	40,3	11,3	6,7	23,2	cel:NH;b2g:NA
BXY_0588800	61,9	95,8	54,3	21,7	1,2	6,7	29,7	cel:DAF-16/FOXO Controlled, germline Tumor affecting;b2g:NA
BXY_0594100	142,8	161,0	109,5	84,4	27,7	25,0	9,5	cel:Glutathione S-Transferase;b2g:NA
BXY_0597200	27,4	30,0	23,8	18,4	8,5	7,0	6,4	cel:PaNtothenate Kinase;b2g:pantothenate kinase
BXY_0610200	1 641,9	3 551,6	2 810,7	1 410,1	68,0	80,7	63,3	cel:NH;b2g:NA
BXY_0610400	47,3	110,5	71,3	53,5	21,0	14,2	8,0	cel:NA;b2g:NA
BXY_0610500	138,1	213,1	206,8	180,0	75,9	37,1	6,5	cel:NA;b2g:NA
BXY_0689100	7,0	13,1	26,4	7,2	1,6	3,4	10,6	cel:PaTched Related family;b2g:patched family protein
BXY_0694500	1 153,3	1 028,7	849,5	808,8	339,1	184,9	7,3	cel:Cysteine PRotease related;b2g:cathepsin b

BXY_0712100	411,6	310,1	44,7	68,5	8,9	2,9	71,1	cel:NA;b2g:nudix hydrolase 6
BXY_0727300	100,9	151,0	102,2	115,2	31,0	52,2	5,6	cel:Glutathione S-Transferase Kappa protein;b2g:2-hydroxychromene-2-carboxylate isomerase
BXY_0777600	35,5	72,4	9,0	7,4	1,1	2,0	40,8	cel:NA;b2g:NA
BXY_0779700	268,8	207,0	96,7	89,3	37,5	46,6	7,9	cel:NA;b2g:phytanoyl- peroxisomal-like
BXY_0784800	16,2	14,2	12,6	5,9	2,0	1,8	12,7	cel:fatty Acid CoA Synthetase family;b2g:protein acs-14
BXY_0798400	79,3	99,2	26,2	31,4	9,8	7,4	13,7	cel:NH;b2g:NA
BXY_0800100	5,1	9,2	21,9	9,0	1,1	2,4	12,8	cel:CYtochrome P450 family;b2g:protein cyp-33d3
BXY_0827900	77,6	107,0	102,7	103,9	25,0	23,6	8,0	cel:NH;b2g:NA
BXY_1014700	281,1	318,4	279,2	256,6	46,5	28,4	15,2	cel:NH;b2g:protein nep- isoform a
BXY_1014800	113,6	152,1	48,3	19,1	2,0	2,3	77,1	cel:NH;b2g:protein nep- isoform a
BXY_1032600	826,9	1 222,0	106,5	109,3	34,1	21,8	40,5	cel:NA;b2g:cholesterol 25-hydroxylase-like protein member 2-like
BXY_1074200	427,8	774,3	977,6	382,2	187,9	113,5	8,5	cel:NA;b2g:epoxide hydrolase 1-like
BXY_1088300	653,5	771,5	1 130,7	541,4	74,4	34,8	28,4	cel:UDP-GlucuronosylTransferase;b2g:protein ugt-15
BXY_1088400	6,0	9,9	11,0	3,7	2,0	1,1	9,9	cel:UDP-GlucuronosylTransferase;b2g:cre-ugt-21 protein
BXY_1088500	38,6	65,9	100,0	42,3	7,6	2,7	23,8	cel:UDP-GlucuronosylTransferase;b2g:protein ugt- isoform a
BXY_1088600	45,8	61,9	35,4	14,6	1,8	1,7	45,2	cel:UDP-GlucuronosylTransferase;b2g:protein ugt- isoform a
BXY_1108500	10,4	2,2	5,4	9,6	1,4	2,0	8,1	cel:NA;b2g:NA
BXY_1166100	276,9	257,9	55,9	66,2	14,2	6,5	31,7	cel:NH;b2g:NA
BXY_1209200	33,5	26,5	20,5	21,8	4,8	2,7	13,8	cel:NH;b2g:NA
BXY_1217800	83,3	73,0	13,1	19,1	2,4	1,7	46,5	cel:NA;b2g:hypothetical protein CAEBREN_00583
BXY_1225100	383,4	333,7	58,3	72,5	6,9	4,9	72,2	Hypothetical protein
BXY_1244700	23,7	34,4	64,6	19,3	8,2	13,0	6,7	Hypothetical protein
BXY_1285700	97,4	142,2	82,9	75,2	31,9	33,4	6,1	Hypothetical protein
BXY_1336500	561,3	537,9	33,4	52,2	6,8	18,6	46,5	cel:NH;b2g:NA
BXY_1346500	280,5	358,0	67,8	82,1	14,7	15,7	25,9	cel:NH;b2g:NA

BXY_1398800	35,3	34,9	21,7	26,1	6,3	1,6	15,0	cel:NH;b2g:NA
BXY_1549700	65,0	131,2	140,7	61,6	6,5	2,6	43,7	cel:DeHydrogenases, Short chain;b2g:3-oxoacyl-(acyl-carrier-protein) reductase
BXY_1559600	16,9	9,9	18,3	24,9	2,7	3,9	10,7	cel:NH;b2g:NA
BXY_1559700	107,4	104,3	77,0	108,0	5,0	16,2	18,7	cel:NH;b2g:NA
BXY_1559800	31,8	26,7	20,5	31,6	2,0	9,9	9,3	cel:NH;b2g:NA
BXY_1560000	39,8	32,5	35,3	57,0	1,0	13,7	11,1	cel:NH;b2g:NA
BXY_1562600	196,8	152,5	36,6	21,4	2,0	9,5	35,2	cel:Glutathione S-Transferase;b2g:NA
BXY_1573800	18,4	30,4	11,5	23,3	3,6	3,6	11,6	cel:NH;b2g:---NA---
BXY_1591300	63,5	134,5	40,2	32,9	11,3	4,6	17,0	cel:Temporarily Assigned Gene name;b2g:NA
BXY_1655900	83,8	127,7	108,4	96,3	27,7	40,1	6,1	cel:ARRestin Domain protein;b2g:NA
BXY_1683600	128,9	412,2	143,8	150,1	55,1	39,2	8,9	cel:NH;b2g:NA
BXY_1685000	71,0	132,0	31,6	24,5	3,1	3,4	39,6	cel:NH;b2g:NA
BXY_1689800	102,6	163,0	54,4	69,4	24,2	25,7	7,8	cel:NH;b2g:NA

Table S7. Primers used in the qPCR experiments.

Target gene symbol	NCBI/EMBL/DDBJ Accsession numbers	Forward primer sequence	Reverse primer sequence	Amplicon length	On Exon/Intron junction?	In silico specificity (BLASTn evalue < 1e-5)	Specificty check post amplification	Cq of NTC
Bx-col-1	LC033885	TCCATGAAACCCGCGAAAGA	TGGACAGTTCTGTGGCTGAC	113	No	Specific	Melt curve	ND
Bx-col-3	LC033886	GCCCATGTTGCACAACTACG	GGCGAATGGGATCTGCTTCA	284	Yes	Specific	Melt curve	ND
Bx-col-4	LC034169	CACTCTGCCCATGGTCTACA	CGGATTTGGTTGACCTCGGA	214	Yes	Specific	Melt curve	ND
Bx-col-5	LC034170	GATGCCAATGCCAAACCGAG	TGATGTCTTCGGCGTCCTTG	132	No	Specific	Melt curve	38,77
Bx-act-1	AB500147	CCACCAGAGCGCAAATACTC	CATCTGTTGGAAGGTGGACA	72	Yes	Specific	Melt curve	36,94

CHAPTER IV

CHARACTERIZATION OF GLUTATHIONE S-TRANSFERASES FROM THE PINEWOOD NEMATODE, *BURSAPHELENCHUS XYLOPHILUS*

Data published in:
Margarida Espada, John T. Jones, Manuel Mota
Nematology, 2016, 18, 697-709

Summary

We have previously identified two secreted glutathione S-transferases (GST) expressed in the pharyngeal gland cell of *B. xylophilus*, which are upregulated post infection of the host. This study examines the functional role of GSTs in *B. xylophilus* biology. We analysed the expression profiles of all predicted GSTs in the genome and the results showed that they belong to kappa and cytosolic subfamilies and the majority are upregulated post infection of the host. A small percentage is potentially secreted and none is downregulated post infection of the host. One secreted protein was confirmed as a functional GST and is within a cluster that showed the highest expression fold change in infection. This enzyme has a protective activity that may involve host defences, namely in the presence of terpenoid compounds and peroxide products. These results suggest that GSTs secreted into the host participate in the detoxification of host-derived defence compounds and enable successful parasitism.

Introduction

Glutathione S-transferases (GST, EC 2.5.1.18) are enzymes involved in detoxification metabolism and are present in a range of different organisms including bacteria, plants and animals. The main function of this large family of enzymes is the detoxification of potentially damaging endogenous stress products and exogenous xenobiotic compounds and also an important role in drug metabolism. This is achieved by the ability to catalyse the conjugation of the reduced form of glutathione (GSH) to potential toxins in order to increase their solubility and thus enable them to be metabolised or excreted from the host (Brophy and Pritchard, 1994; Campbell *et al*., 2001; Torres-Rivera and Landa, 2008; Matouskova *et al*., 2016). GST does not act directly on reactive oxygen species (ROS), but on the oxidized products of their activity, including lipid hydroperoxides and reactive carbonyls (Torres-Rivera and Landa, 2008). In parasitic species GST is an important detoxification enzyme, especially in helminths where GSTs provide initial defence against oxidative damage and protect the worm from the host immune response, as well as acting as drug-binding proteins (Precious and Barrett, 1989; Brophy and Barrett, 1990; Brophy and Pritchard, 1994; Matouskova *et al*., 2016). Therefore, the roles of these enzymes in the host-parasite interaction have been studied extensively. Recent studies on GSTs from animal parasitic helminths showed that sigma-GSTs have prostaglandin synthase activity, and bind to toxins to a suppression of the host immune response to the benefit of the parasite (van Rossum *et al*., 2004; Dowling *et al*., 2010; LaCourse *et al*., 2012). In addition, analysis of the secretome of the animal parasitic trematode *Fasciola hepatica*, revealed sigma class-GST in extracellular vesicules that are deployed during parasitism (Cwiklinski *et al*., 2015). In the plant parasitic nematode *Meloidogyne incognita*, one GST has been identified as being secreted from the pharyngeal gland cells (*Mi-gst*-1) and plays an important role in the interaction with the host as evidenced by the fact that silencing of this

gene by RNAi leads to a reduction in parasitism. This GST may protect the nematode against host derived ROS or may modulate plant responses that are triggered by nematode attack (Dubreuil *et al*., 2007).

Parasitic helminths contain several forms of GSTs which can be grouped in subfamilies on the basis of their subcellular location: kappa (mitochondrial), microsomal and cytosolic (soluble GSTs from the mu, alpha, pi, theta, sigma, zeta and omega classes) (Frova, 2006; Torres-Rivera and Landa, 2008). Several GSTs have been identified in migratory plant-parasitic nematodes (PPN), including *Bursaphelenchus xylophilus*, *Ditylenchus africanus*, *Pratylenchus coffeae*, *Radopholus similis* and from the sedentary species *Meloidogyne* spp. and *Globodera pallida* (Bellafiore *et al*., 2008; Dubreuil *et al*., 2007; Haegeman *et al*., 2009; Haegeman *et al*., 2011; Kikuchi *et al*., 2011; Cotton *et al*, 2014; Espada *et al*., 2016; Jacob *et al*., 2008). A total of 65 potential GSTs were predicted from the genome of *B. xylophilus*, a similar number to that in *C. elegans,* but higher than seen in other PPN (Kikuchi *et al*., 2011).

When the pinewood nematode (PWN), *B. xylophilus*, infects a tree it triggers several physical and chemical alterations leading to the symptoms of pine wilt disease (PWD). Kuroda et al., (1991) hypothesised a mechanism of cavitation, in which terpenoids synthesised in xylem ray cells induce cavitation and embolisms in tracheids leading to failure of water transport. Previous studies have shown that levels of plant terpenes in *P. thunbergii*, particularly α-pinene and β-pinene, increase when the tree is infected by *B. xylophilus* (Fukuda *et al*., 1997; Kuroda *et al*., 1991; Kuroda, 1991). However, a recent study examining infection of plant material maintained in tissue culture, suggested that terpenoid compounds do not significantly increase after infection with PWN although levels were maintained after infection, with α-pinene making up between 26%-32% of total terpenoid content and β-pinene between 34%-47% (Faria *et al*., 2015). Several of these compounds have nematicidal activity, although no study has been made in *B. xylophilus*. Chemical compounds including terpenoids have been tested against filarial nematode GST and one study showed that α-pinene has an inhibitory effect on the nematode GST (Azeez *et al*., 2012).

In a previous study, we identified two secreted glutathione S-transferases that were upregulated in an early stage of infection and which are expressed in the dorsal pharyngeal gland cell (Espada *et a.l*, 2016). It was suggested that these enzymes could be involved in detoxification of plant endogenous compounds, helping *B. xylophilus* to overcome host defences. Here we demonstrate that at least one of these is a functional GST and that the presence of this enzyme provides protection against stresses likely to be encountered during infection of the host tree. We show that biochemically active GST is secreted by nematodes. In addition, we examine the global changes in expression of *B. xylophilus* GSTs upon infection of the host.

Results and Discussion

Global analysis of B. xylophilus *GST expression profiles*

An analysis of GSTs performed as part of the *B. xylophilus* genome project (Kikuchi *et al.,* 2011) identified 65 potential GSTs. Our BLASTP based analysis of the *B. xylophilus* genome revealed that five more sequences, which could encode proteins similar to GSTs, were present (Figure 1). Analysis of the protein domains present in each sequence confirmed all the protein sequences as GSTs, as described in Table S1. The majority of these sequences have a thioredoxin-like fold domain (IPR012336) followed by glutathione S-transferase N-terminal and C-terminal domains, both of which are features of cytosolic subfamily (reviewed in Frova, 2006). Five sequences contained domains similar to maleylacetoacetate isomerase (IPR005955), which is a feature of the zeta class of GSTs. The other 4 sequences were identified as kappa subfamily GSTs, due to the presence of the DSBA-like thioredoxin domain (IPR01853) (a feature of the HCCA isomerase/GST kappa family – IPR01440) (Frova, 2006).

Six of the GST sequences have a predicted signal peptide, suggesting a role in detoxification of extracellular compounds, including host derived toxins (Figure 1). These potentially secreted proteins included the two sequences (BUX.s00647.112 and BUX.s01254.333) that were previously identified as being expressed in the pharyngeal gland cells (Espada *et al*., 2016). Next we used our previously described RNAseq dataset to examine global changes in expression profiles of the GST sequences, by using log2 of the fold changes. This showed that 42 of the GST sequences are upregulated in nematodes after infection of trees as compared to nematodes grown on fungi (Figure 1), including four of those sequences with a signal peptide. None of the secreted GSTs were downregulated after infection. The maximum-likelihood phylogenetic analysis (Figure 1) of the *B. xylophilus* GSTs and sequences from other nematodes showed that the pharyngeal gland cell sequences clustered with other sequences upregulated after infection. One (BUX s01254.333) formed a cluster with other secreted and upregulated protein while the other (BUX s00467.112) clustered with another secreted protein and two other upregulated proteins. This cluster includes the sequences that show the highest increases in expression during the infection of the host. Neither the pharyngeal gland cell GSTs, nor the secreted GSTs formed a single cluster (although the secreted GSTs were present as pairs in three clusters). These clusters were consistent in a neighbour-joining phylogenetic analysis (Figure S1). These data suggest that a range of the GSTs present in *B. xylophilus* have been recruited independently to play a role in protection against host derived toxins and that the range of secreted GSTs has not evolved as a result of duplication of a single secreted ancestor.

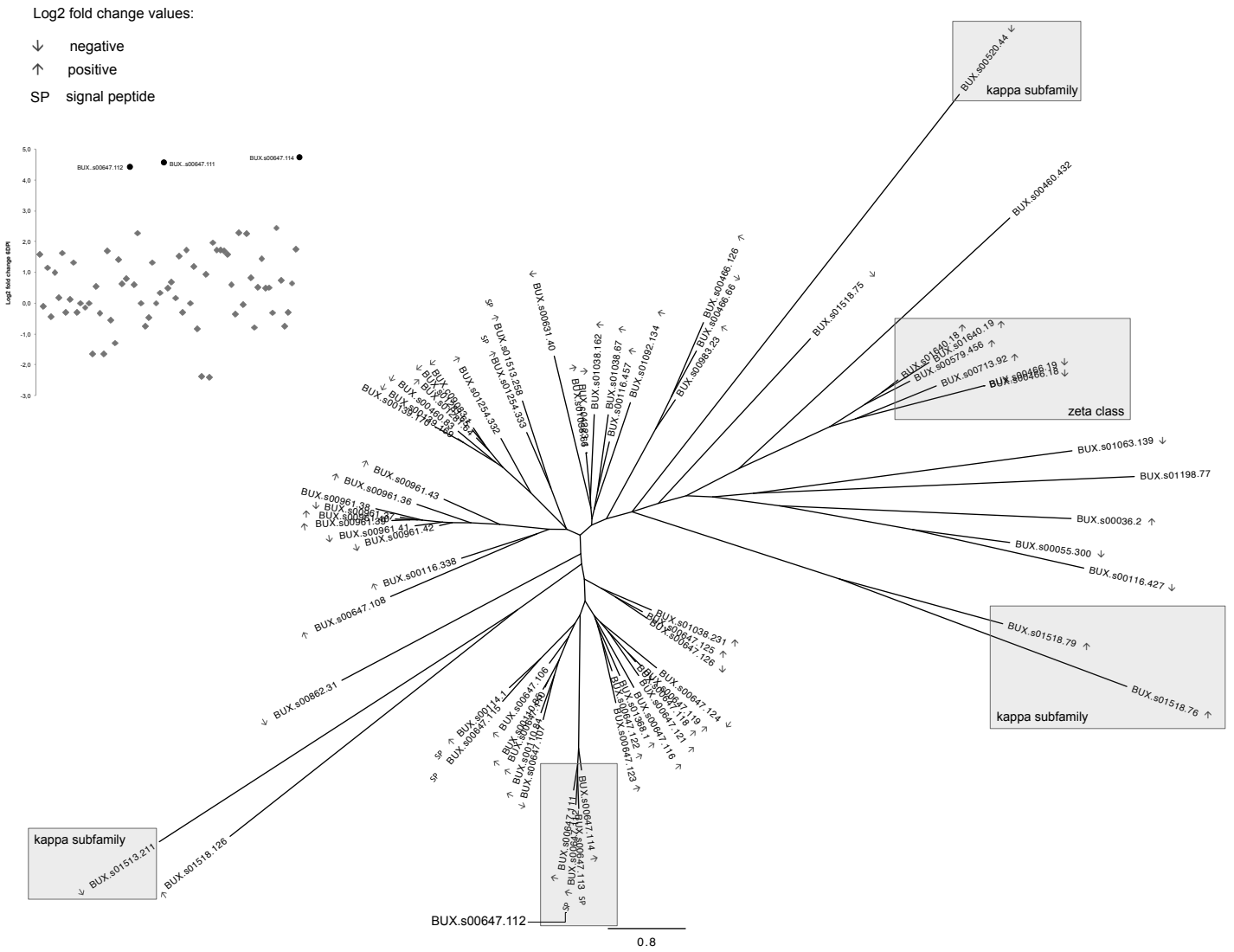


Figure 1 – Maximum-likelihood phylogenetic tree that represents the protein sequence similarity between all 70 PWN predicted GSTs. The GSTs belonging to the kappa subfamily and to the cytosolic zeta class are represented within grey boxes. For each gene the log2 of the fold changes (6 days post infection) values of the expression levels are represented by arrows. The highest log2 fold change values belong to the genes BUX.s00647.112, BUX.s00647.111, BUX.s00647.114 that cluster together, represented within a grey box. The dot plot on the top left of the figure is a representative chart of the expression values of all genes. SP represents the presence of a signal peptide.

Enzymatic and protective activity of GSTs involved in the host-parasite interaction

We next examined the biochemical activity of one of the pharyngeal gland cell GST sequences. The recombinant BUX.s00647.112 protein was cloned into an expression vector with an N-terminal His tag and purified from bacterial cell lysate, yielding a protein of approximately 25KDa, in agreement with the size predicted from the amino acid sequence (Figure 2). The recombinant protein had glutathione transferase activity (using CDNB as a substrate) very similar to that observed for the positive control (Table 1). These data confirm that the BUX.s00647.112 protein is a functional GST.

Our previous data showed that several GSTs (including the BUX s00647.112 sequence) are expressed in the pharyngeal gland cells, from where they could be secreted into the host. In addition, a larger scale proteomic analysis of *B. xylophilus* secreted proteins identified several peptides that could be derived from GST sequences (Shinya *et al*., 2013) further suggesting that GSTs form an important component of the *B. xylophilus* repertoire of secreted proteins. In keeping with this, we were able to detect GST activity (albeit at low levels) in secretions collected from *B. xylophilus* (Table 1). The RNAseq data suggest that it would have been possible to detect higher GST activity in secretions harvested from nematodes extracted from trees but technical limitations prevented us from attempting this analysis.

Table 1 – Glutathione transferase activity results using CDNB as substrate (Blank) in recombinant BUX.s00647.112 protein, *B. xylophilus* protein extract and secretions. Each value is represented by mean ± SD.

Sample (CDNB as substrate)	**GST activity (µmol ml $^{-1}$min^{-1})**
GST activity in crude extracts of PWN proteins and secretions	
GST (control)	133.7 ± 62.3
PWN secretions	31.2 ± 1.9
PWN proteins	37.1 ± 0.2
GST activity in the recombinante BUX.s00647.112 protein	
GST (control)	1509.8 ± 73.4
Recombinant BUX.s00647.112	2096.3 ± 312.5

We next sought to analyse whether the *B. xylophilus* pharyngeal gland cell GST can provide protection against the toxins likely to be encountered by a nematode infecting a pine tree. Testing the function of the GST in pine trees is not possible due to technical limitations. We therefore compared the ability of bacterial cells in which the GST was either induced or not induced to grow in the presence of hydrogen peroxide and several terpenoid compounds. The peroxide was intended to represent the products of ROS while the terpenoids were chosen to mimic toxic compounds likely to be present in an infected pine tree. In the presence of the

GST, bacteria showed significantly higher growth in an environment with a (-) and (+)-α-pinene (-)-β-pinene, 0.5% limonene and up to 3% hydrogen peroxide (Figure 3). There were no significant differences in the 1% limonene treatment or in the control (induced vs. non-induced). The difference in growth rate was most apparent in the presence of 0.5% (-)-β-pinene. A Western blot (Figure 2) showed that the recombinant GST was present in all IPTG-induced samples while the non-induced bacterial cells showed no signal in the blot (Figure 2).

These data confirm that *B. xylophilus* secretes functional GST proteins into the host, which may be important for allowing the nematode to overcome host defences. This may be a strategy that is widely used by plant-parasitic nematodes: a secreted GST has been identified from *M. incognita* (*Mi-gst-1*) which has been shown to promote infection of this nematode (Dubreuil *et al*., 2007) and which is also thought to function by protecting the nematode from host defences. Like the *B.xylophilus* sequence, the *M. incognita* GST is upregulated upon infection and expressed in the pharyngeal gland cells. GSTs also form a significant component of the strategy used by animal-parasitic nematodes to neutralise host defence responses. It is likely that GSTs used for internal metabolic processes have become adapted for a role in the host-parasite interaction in both plant- and animal-parasitic nematodes. Similar adaptation of housekeeping proteins for roles in parasitism in animal and plant parasites has been described previously with peroxiredoxins, glutathione peroxidases and lipid binding proteins all known to be deployed by plant-parasites and animal parasites in order to provide protection from host defences (reviewed by Jasmer *et al*., 2003). Convergent evolution between animal- and plant-parasitic nematodes is therefore a recurring theme in terms of how they cope with host derived stresses.

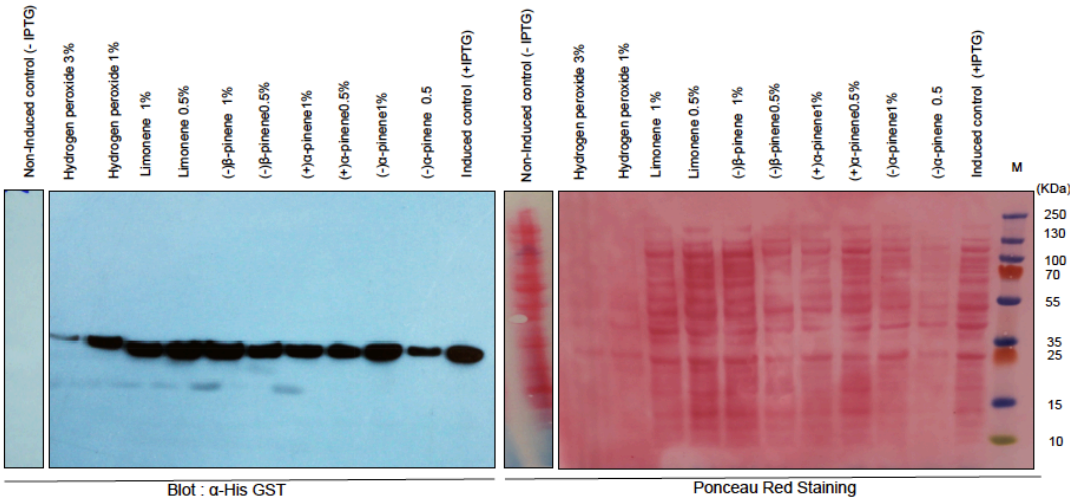


Figure 2 – The results of the immuno-detection of anti-Histag on the recombinant BUX.s00647.112 protein resistance assays. On the right, the Ponceau Red staining and on the left the results of the blot detected by chemilumencence. M: protein ladder (GeneRuler, Thermofisher).

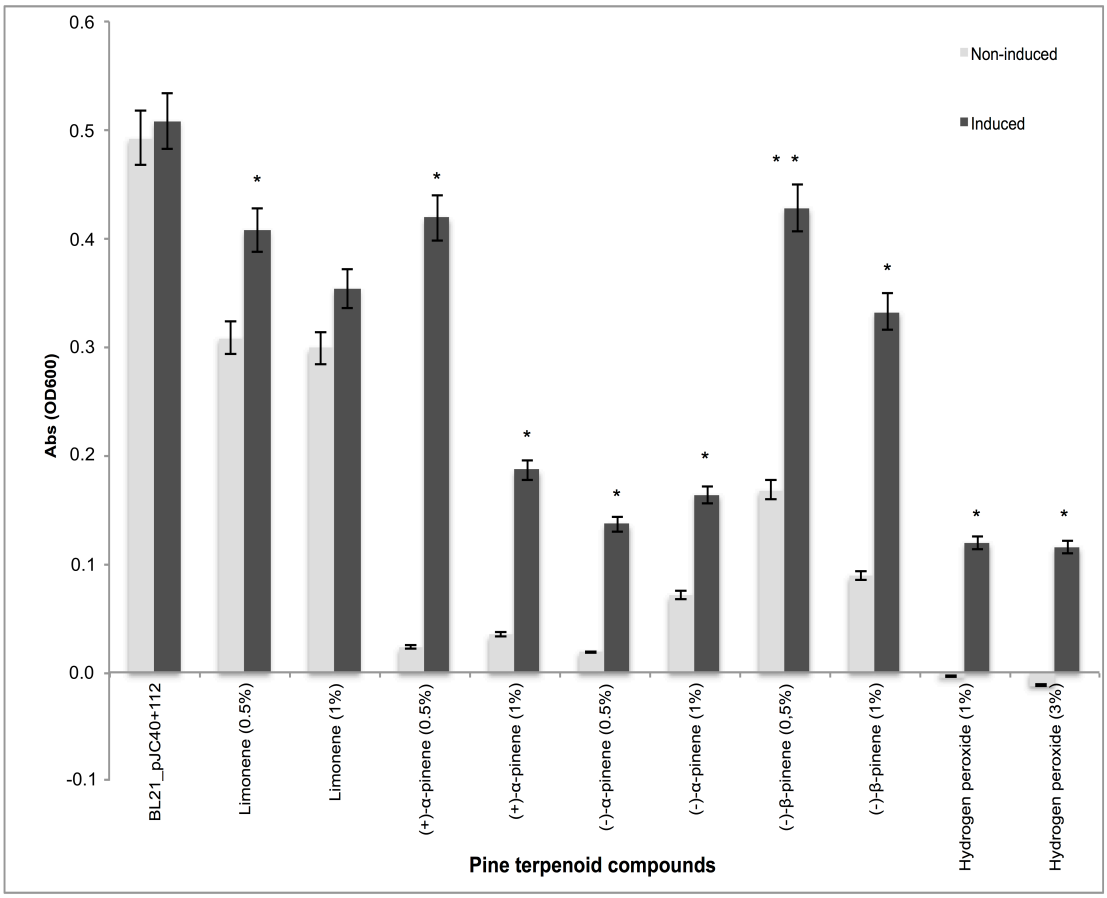


Figure 3 – Resistance test in BL21(DE3) cells. Induced vs. non-induced BUX.s00647.112 protein using different pine terpenoid compounds and different concentrations of each (X axis). The values in the Y axis correspond to values of absorbance (OD_{600}). The LB media was used to grow the bacteria. Protein expression was induced with 0.5mM IPTG (see experimental procedures). Significant differences between induced and non-induced treatments were analysed by ANOVA (* p-value<0.05; ** p-value<0.01).

Experimental procedures

Phylogenetic analysis of GST sequences

Potential GST-encoding sequences were identified using the previous data from Kikuchi *et al.* (2011) and by BLASTP searching the gene calls from the *B. xylophilus* genome against the NR database (cutoff $1e^{-5}$). Any sequences for which at least one of the top three hits included the expression "glutathione S-transferase", were selected for further analysis. This analysis was performed using BLAST+ wrappers for Galaxy (v0.1.01) (Cock *et al.*, 2015). The expression levels of the transcripts at various life stages were predicted from RNAseq data generated in a previous study (Espada *et al.*, 2016) and log2 of the fold change for each gene was calculated. For all the predicted GSTs the subfamilies and protein domains were identified using InterProScan 5 (Jones *et al.*, 2014) (http://www.ebi.ac.uk/interpro/search/sequence-search). Secreted GSTs were predicted based on the presence of signal peptide as predicted by SignalP (v3.0) (Petersen *et al.*, 2011) and the absence of a transmembrane domain. All the alignments

of the full-length protein sequences were performed with the software SeaView (Gouy *et al*., 2010). The Maximum-likelihood phylogenetic tree was generated by PhyML (in SeaView) from the alignment of the sequences (protein distance measure: Jukes-cantor; aLRT SH-like for branch supporting). The phylogenetic tree was edited in FigTree (v1.4.0) (http://tree.bio.ed.ac.uk/software/figtree/). A neighbour-joining phylogenetic tree was generated in the software CLC Sequence Viewer (v7.6.1) (protein distance measure: Jukes-cantor; one thousand replicates for bootstrap for branch supporting).

Biological material

Nematodes were cultured on *Botrytis cinerea* and harvested using a Baermann funnel as previously described (Espada *et al*., 2016). Secreted proteins were collected as described in Kikuchi *et al*., (2004). Briefly, mixed life stage nematodes were collected in a 15ml-tube, by centrifugation at 2844g for 15 minutes, suspended in 1ml M9 buffer and incubated for 2 days at 18C. After this time, the sample was centrifuged at 2844g to pellet the nematodes, the supernatant containing secreted proteins was collected and stored in aliquots at -80^{o}C until used in enzyme assays.

Cloning in expression vector and protein purification

The primers to amplify the full-length of one of the *B. xylophilus* GSTs shown to be expressed in the dorsal pharyngeal gland cells (BUX.00647.112) were designed from the cDNA sequence lacking the signal peptide (as predicted by SignalP 3.0). The gene specific primers included the Kozak sequence (ACCATG) in the forward primer (5'ACCATGTTAGAGCTGTATTATTTCAACGAGAAG) and a Stop codon (TGA) in the reverse primer (5'TCATTGAGTGGCATTGAAATAATTGTAAATCG). The full length gene was amplified using KOD Hot Start proof-reading DNA polymerase and purified using the QIAquick gel extraction Kit (Qiagen). The gene was cloned into the pCR8 TOPO vector and transformed in one shot TOP10 competent cells following the manufacturer's instructions (Invitrogen). The recombinant clones were screened by colony PCR and one clone was confirmed by sequencing. Purified entry plasmid (approximately 140ng) was transferred to the destination vector pJC40 (a 10xHis-tag N-terminus fusion vector) using the LR cloning kit following the manufacturer's instructions (Invitrogen). The cloning reaction was transformed into BL21(DE3) chemical competent cells. Positive transformants (construct pJC40+00647.112) were analysed by colony PCR and confirmed by sequencing. The His-tagged protein was induced by adding 1mM IPTG (isopropyl-beta-D-thiogalactopyranoside) to a bacterial culture grown from a single colony in 10ml LB with 100µg/ml ampicillin, at 37C until the concentration reached an OD_{600} of 0.6. The protein was then purified using Ni-NTA resin columns (Ni-NTA Spin kit, Qiagen) following the manufacturer's protocol.

Resistance test in BL21(DE3) cells

To induce expression of the recombinant protein, a single colony was grown in 45ml LB and 100µg/ml of ampicillin, at 37C with agitation, until the concentration reached an OD_{600} of 0.6. At this point 100µl aliquots of the bacterial suspension were placed in new sterile 15ml-tubes containing 4ml LB and to which the terpenes (limonene, (+) and (-)-α-pinene, (-)-β-pinene) or hydrogen peroxide were added. For each treatment, two different concentrations were tested and two replicates were used: 0.5% and 1% for limonene, (+) and (-)-α-pinene, (-)-β-pinene (Sigma-Aldrich); 1% and 3% for hydrogen peroxide. Protein expression was induced in the remaining bacterial suspension by adding 0.5mM IPTG and incubating at 37C, with agitation, for 2 hours. After this time the terpenoid and hydrogen peroxide treatments were repeated using 100 µl aliquots of the bacterial suspension as described above. The respective control tubes were also grown in the same conditions. All the treatments were subsequently grown overnight at 37C with agitation. The OD_{600} was measured for all treatments in a spectrophotometer (Spekol 1500, Analytik Jena). The results were analysed with an ANOVA test using the statistical software GenStat (version 17th; VSN International, 2012).

Western Blotting

Aliquots of the bacterial cells from test described above were used in a Western-blot using an antibody against a poly-histidine tag (Sigma-Aldrich) to demonstrate the presence of the recombinant protein in the assay. The bacterial extracts were heated at 90C for 10 minutes in NuPage LDS sample buffer (Invitrogen). The proteins were separated on a 4-12% NuPage Bis-Tris gel and transferred to nitrocellulose membrane (GE Healthcare). Immuno-detection of the protein was performed using anti-His antibody (Sigma) at 1:5,000 dilution as primary antibody and detected using a secondary antibody conjugated to peroxidase (α-mouse IgGxHRP at 1:50,000) (Sigma) by chemiluminescence using the Pierce Supersignal West Pico kit (Thermo-Scientific).

Enzyme assay

The Glutathione-S-transferase assay kit (Sigma-Aldrich) was used with 1-chloro-2,4-dinitrobenzene (CDNB) as the standard substrate to test activity of recombinant protein and activity present in collected secretions. All assays were replicated three times. A solution containing 2mM reduced L-glutathione and 1mM CNDB in Dulbecco's phosphate buffered saline was prepared and used within an hour of preparation. 50 µl aliquots of this solution were mixed with 1µl of control GST enzyme or with test enzyme preparations and transferred to a quartz cuvette. Absorbance was measured at 340nm, each 30 seconds over a period of 5 minutes, after a lag time of 1 minute, following the manufacturer's procedure. GST activity was calculated for each sample as described by the kit manufacturer.

References

Azeez, S., Babu, R.O., Aykkal, R. and Narayanan, R. (2012) Virtual screening and in vitro assay of potential drug like inhibitors from spices against glutathione-S-transferase of filarial nematodes. *J. Mol Mod* 18, 151-163.

Bellafiore, S., Shen, Z.X., Rosso, M.N., Abad, P., Shih, P. and Briggs, S.P. (2008). Direct identification of the *Meloidogyne incognita* secretome reveals proteins with host cell reprogramming potential. *PLoS Pathog* 4, e1000192.

Brophy, P.M. and Barrett, J. (1990) Glutathione transferase in helminths. *Parasitology* 100 (2), 345-349.

Brophy, P.M. and Pritchard, D.I. (1994) Parasitic-helminth Glutathione S-transferases: an update on their potential targets for immuno- and chemotherapy. *Exp. Parasitol.* 79, 89-96.

Campbell, A.M., Teesdale-Spittle, P.H., Barrett, J., Liebau, E., Jefferies, J.A. and Brophy, P.M. (2001) A common class of nematode glutathione S-transferase GST/ revealed by the theoretical proteome of the model organism *Caenorhabditis elegans*. *Comp Biochem Physiol B* 128, 701-708.

Cock, P.J.A., Chilton, J.M., Grüning, B., Johnson, J.E. and Soranzo, N. (2015) NCBI BLAST+ integrated into Galaxy. *GigaScience* 4, 39.

Cotton, J.A., Lilley, C.J., Jones, L.M., Kikuchi, T., Reid, A.J., Thorpe, P., Tsai, I.J., Beasley, H., Blok, V., Cock,P.J.A., Eves-van den Akker, S., Holroyd, N., Hunt, M., Mantelin, S., Naghra, H., Pain, A., Palomares-Rius, J.E, Zarowiecki, M., Berriman, M., Jones, J.T. and 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 Biol.*, 15:R43.

Cwiklinski, K., de la Torre-Escudero, E., Trelis, M., Bernal, D., Dufresne, P.J., Brennan, G.P., O'Neill, S., Tort, J., Paterson, S., Marcilla, A., Dalton, J.P. and Robinson, M.W. (2015) The Extracellular Vesicles of the Helminth Pathogen, *Fasciola hepatica*: Biogenesis Pathways and Cargo Molecules Involved in Parasite Pathogenesis. *Mol. Cel Proteomics* 12,3258-73.

Dowling, D.J., Hamilton, C.M., Donnelly, S., La Course, J., Brophy, P.M., Dalton, J. and O'Neill, S.M. (2010) Major secretory antigens of the helminth *Fasciola hepatica* activate a suppressive dendritic cell phenotype that attenuates Th17 cells but fails to activate Th2 immune responses. *Infect. Immun*, 78, 793-801.

Dubreuil, M., Magliano, E., Deleury, P., Abad, P. and Rosso, M.N. (2007) Transcriptome analysis of root-knot nematode functions induced in the early stages of parasitism. *New Phytol.* 176, 426–436.

Espada, M., Silva, A.C., Eves-van den Akker, S., Cock,P.J.A., Mota, M. and Jones, J.T. (2016) Identification and characterization of parasitism genes from the pinewood nematode *Bursaphelenchus xylophilus* reveals a multi-layered detoxification strategy. *Mol. Plant Pathol.* 17, 286-295.

Faria, J.M.S., Sena, I., Vieira da Silva, I., Ribeiro, B., Barbosa, P., Ascensão, L., Bennett, R.N., Mota, M. and Figueiredo, A.C.F. (2015) In vitro co-cultures of *Pinus pinaster* with *Bursaphelenchus xylophilus*: a biotechnological approach to study pine wilt disease. *Planta* 241, 1325-1336.

Frova, C. (2006) Glutathione transferases in the genomics era: new insights and perspectives. *Biomol. Eng.* 23, 149-169.

Fukuda, K. (1997) Physiological Process of the Symptom Development and Resistance Mechanism in Pine Wilt Disease. *J For Res* 2, 171-181.

Gouy M., Guindon S. and Gascuel O. (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* 27(2), 221-224.

Haegeman, A., Jacob, J., Vanholme, B., Kyndt, T., Mitreva, M. and Gheysen, G. (2009) Expressed sequence tags of the peanut pod nematode *Ditylenchus africanus*: The first transcriptome analysis of an Anguinid nematode. *Mol. Biochem. Parasitol.* 167, 32-40.

Haegeman, A., Joseph, S. and Gheysen, G. (2011) Analysis of the transcriptome of the root lesion nematode *Pratylenchus coffeae* generated by 454 sequencing technology. *Mol. Biochem. Parasitol.* 178, 7-14.

Jacob, J., Mitreva, M., Vanholme, B. and Gheysen, G. (2008) Exploring the transcriptome of the burrowing nematode *Radopholus similis*. *Mol. Genet. Genomics* 280, 1-17.

Jasmer, D.P., Goverse, A. and Smant, G. (2003) Parasitic nematode interactions with mammals and plants. *Annu. Rev. Phytopathol.* 41, 245-270.

Jones, P., Binns, D., Chang,H., Fraser,M., Li ,W, McAnulla, C., McWilliam,H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M.,

Yong, S., Lopez, R. and Hunter, S. (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30(9), 1236–1240.

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

Kikuchi, T., Cotton, J.A., Dalzell, J.J., Hasegawa, K., Kanzaki, N., McVeigh, P., Takanashi, T., Tsai, I.J., Assefa, S.A., Cock, P.J.A., Otto, T.D., Hunt, M., Reid, A.J., Sanchez-Flores, A., Tsuchihara, K., Yokoi, T., Larsson, M.C., Miwa, J., Maule, A.G., Sahashi, N., Jones, J.T. and Berriman, M. (2011) Genomic insights into the origin of parasitism in the emerging plant pathogen *Bursaphelenchus xylophilus*. *PLoS Pathog* 7, e1002219.

Kuroda, K. (1991) Mechanism of cavitation development in the pine wilt disease. *Eur. J. For. Pathol.* 21, 82–89.

Kuroda, K., Yamada, T. and Ito, S. (1991) *Bursaphelenchus xylophilus* induced pine wilt: factors associated with resistance. *Eur. J. For. Pathol.* 21, 430-438.

LaCourse, E.J., Perally, S., Morphew, R.M., Moxon, J.V., Prescott, M., Dowling, D.J., O'Neill, S.M., Kipar, A., Hetzel, U., Hoey, E., Zafra, R., Buffoni, L., Arévalo, J.P. and Brophy, P.M. (2012) The sigma class glutathione transferase from the liver fluke *Fasciola hepatica*. *PLoS NeglTrop. Dis*. 6(5): e1666.

Matoušková, P., Vokřál, I., Lamka, J. and Skálová, L. (2016) The Role of Xenobiotic-Metabolizing Enzymes in Anthelmintic Deactivation and Resistance in Helminths. *Trends Parasitol.*, 32, 481-491.

Petersen, T. N., Brunak, S., von Heijne, G. and Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785-786.

Precious, W.Y. and Barrett, J. (1989) Xenobiotic metabolism in helminths. *Parasitol.Today* 5(5), 156-160.

Shinya, R., Morisaka, H., Kikuchi, T., Takeuchi, Y., Ueda, M. and 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, e67377.

Torres-Rivera, A. and Landa, A. (2008) Gluthathione transferases from parasites: a biochemical view. *Acta tropica* 105, 99-11.

van Rossum, A.J., Jefferies, J.R., Rijsewijk, F.A.M., LaCourse, E.J., Teesdale-Spittle,P., Barrett,J., Tait, A. and Brophy, P.M. (2004) Binding of Hematin by a New Class of Glutathione Transferase from the Blood-Feeding Parasitic Nematode *Haemonchus contortus*. *Infect. Immune.* 72, 2780-2790.

Supporting information

Table S1 - Protein domains predicted for all 70 putative GSTs from the nematode. Each domain is represented by the InterProScan identification code. For some of the proteins the family was identified.

Gene code*	Protein family	# Domain 1	#Domain 2	# Domain 3
BUX.c04223.1	none	IPR010987 Glutathione S-transferase, C-terminal-like	IPR004046 Glutathione S-transferase, C-terminal	
BUX.c09083.1	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00110.84	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00110.85	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00114.1	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00116.427	failed axon connections IPR026928	IPR012336 Thioredoxin-like fold	IPR010987 Glutathione S-transferase, C-terminal-like	IPR004046 Glutathione S-transferase, C-terminal
BUX.s00116.457	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00116.338	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00139.169	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00139.170	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00460.83	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00466.126	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00466.66	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00631.40	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00647.107	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00647.123	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	
BUX.s00647.116	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	
BUX.s00647.121	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00647.112	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00647.118	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00647.119	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00647.113	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	
BUX.s00647.126	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00647.124	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00647.108	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	
BUX.s00647.115	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	
BUX.s00647.111	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00647.110	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00647.125	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00647.122	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00862.31	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	
BUX.s00961.40	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00961.38	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	
BUX.s00961.36	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00961.42	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00961.37	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00961.43	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00961.41	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s00983.23	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s01038.67	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	
BUX.s01038.162	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s01038.66	none	IPR010987 Glutathione S-transferase, C-terminal-like	IPR004046 Glutathione S-transferase, C-terminal	
BUX.s01038.231	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s01092.134	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s01254.333	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s01254.332	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s01281.65	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s01281.64	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s01368.1	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s01513.258	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s01518.75	none	IPR012336 Thioredoxin-like fold	IPR010987 Glutathione S-transferase, C-terminal-like	IPR004046 Glutathione S-transferase, C-terminal
BUX.s00520.44	none	IPR012336 Thioredoxin-like fold	IPR001853 DSBA-like thioredoxin domain	
BUX.s00647.114	none	IPR010987 Glutathione S-transferase, C-terminal-like		
BUX.s01513.211	none	IPR012336 Thioredoxin-like fold	IPR001853 DSBA-like thioredoxin domain	
BUX.s01518.79	none	IPR001853 DSBA-like thioredoxin domain	IPR012336 Thioredoxin-like fold	

BUX.s01518.76	HCCA isomerase/glutathione S-transferase kappa (IPR014440)	IPR012336 Thioredoxin-like fold	IPR001853 DSBA-like thioredoxin domain	
BUX.s00036.2	none	IPR010987 Glutathione S-transferase, C-terminal-like	IPR004046 Glutathione S-transferase, C-terminal	
BUX.s00055.300	none	IPR012336 Thioredoxin-like fold	IPR010987 Glutathione S-transferase, C-terminal-like	IPR004046 Glutathione S-transferase, C-terminal
BUX.s00460.432	none	IPR2109 Glutaredoxin	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal
BUX.s00466.18	Maleylacetoacetate isomerase (IPR005955)	IPR005955 Maleyacetoacetate isomerase	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal
BUX.s00466.19	Maleylacetoacetate isomerase (IPR005955)	IPR005955 Maleyacetoacetate isomerase	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal
BUX.s00579.456	Maleylacetoacetate isomerase (IPR005955)	IPR005955 Maleyacetoacetate isomerase	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal
Bux.s00647.106	none	IPR012336 Thioredoxin-like fold	IPR010987 Glutathione S-transferase, C-terminal-like	
BUX.s00713.92	Maleylacetoacetate isomerase (IPR005955)	IPR005955 Maleyacetoacetate isomerase	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal
BUX.s00961.39	none	IPR010987 Glutathione S-transferase, C-terminal-like		
BUX.s01063.139	none	IPR010987 Glutathione S-transferase, C-terminal-like	IPR004046 Glutathione S-transferase, C-terminal	
BUX.s01198.77	none	IPR012336 Thioredoxin-like fold	IPR010987 Glutathione S-transferase, C-terminal-like	
BUX.s01518.126	none	IPR012336 Thioredoxin-like fold	IPR010987 Glutathione S-transferase, C-terminal-like	IPR004046 Glutathione S-transferase, C-terminal
BUX.s01640.19	none	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like
BUX.s01640.18	Maleylacetoacetate isomerase (IPR005955)	IPR012336 Thioredoxin-like fold	IPR004045 Glutathione S-transferase, N-terminal	IPR010987 Glutathione S-transferase, C-terminal-like

*according to *B. xylophilus* gene calls of the genome (version 1.2); available in GeneDB (www.genedb.org/Homepage/Bxylophilus)

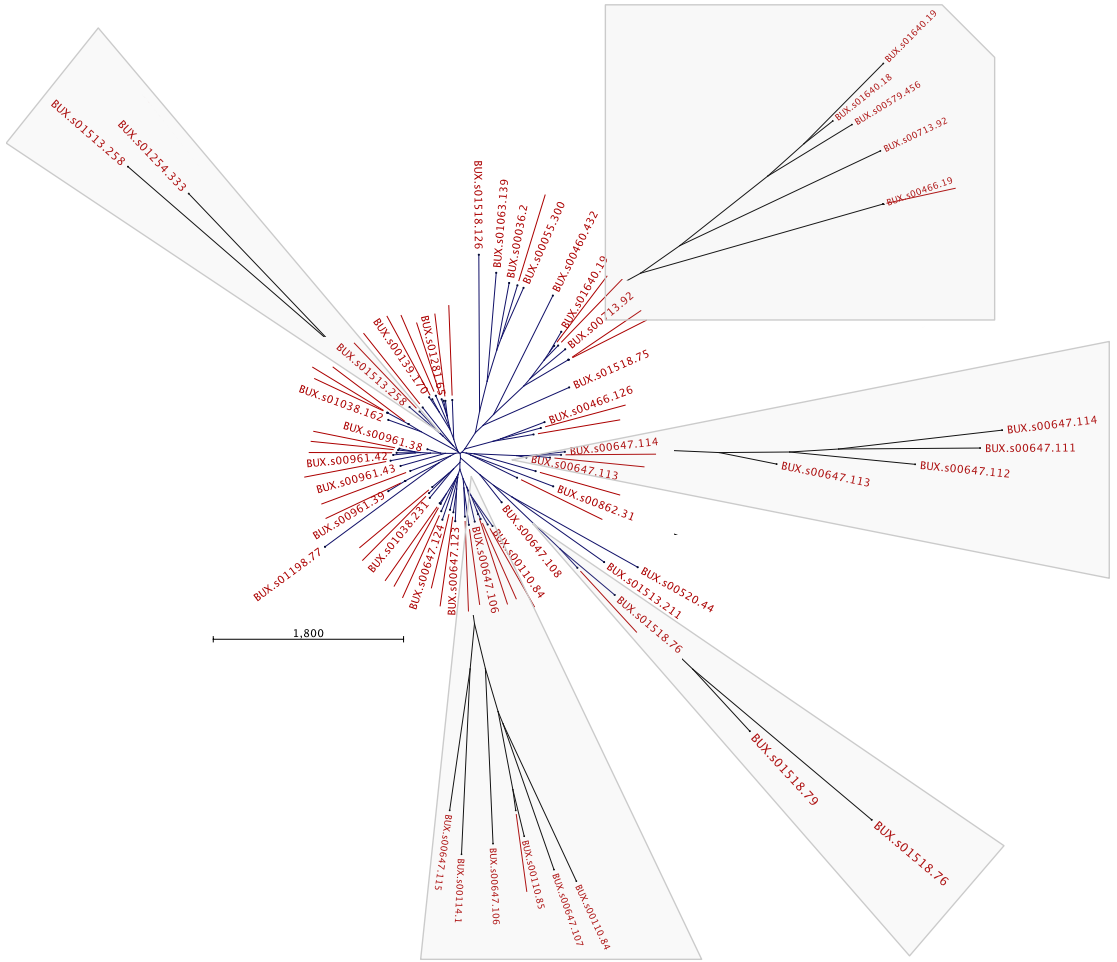


Figure S1- Neighbour-joining phylogenetic tree of all 70 protein sequences from PWN. The highlighted clusters in grey boxes represent the kappa subfamily, the zeta classe and the clusters with the protein of interest (BUX.s00647.112) and the proteins with predicted signal peptide. This tree confirms that the clusters are not and artefact of the maximum-likelihood phylogenetic tree.

CHAPTER V

DIRECT SEQUENCING OF THE TRANSCRIPTOME OF PHARYNGEAL GLAND CELLS AND ANALYSIS OF PROMOTER MOTIFS ASSOCIATED WITH GLAND CELL GENES AS TOOLS FOR IDENTIFICATION OF EFFECTORS FROM THE PINE WOOD NEMATODE *B. XYLOPHILUS*

Manuscript in preparation

Summary

Plant-parasitic nematodes cause severe damage to a wide range of crops and forestry plant species worldwide. The migratory endoparasitic nematode, *Bursaphelenchus xylophilu*s is a quarantine pathogen that affects pine species and consequently has a major impact on the forestry industry. The interactions of this nematode with the plant host are mediated by secreted effector proteins produced in the pharyngeal gland cells. Identification of effectors is important to understand the parasitism mechanisms and develop new control measures for the pathogens. Using an approach pioneered in cyst nematodes, we have analysed the promoter regions of a small panel of previously validated effector genes from *B. xylophilus* and identified a DNA motif associated with these sequences. Analysis of the whole genome subsequently showed that this DNA sequence motif is present in the promoter region of approximately, 600 genes of the nematode We subsequently analysed an RNAseq dataset derived from purified gland cells. Abundance in this data set was correlated with an increased probability of the presence of a signal peptide. In addition, the novel DNA motif followed the same distribution and was more likely to be found associated with the most abundant gland cell genes. This study described a new promoter regulation element from a migratory plant-parasitic nematode, *B. xylophilus* and allowed identification of 118 candidate effectors that are represented in the gland cell transcriptome, that have at least one iteration of the motif in their promoter region and that have a signal peptide. These sequences represent a powerful resource for the study of *B. xylophilus* infection biology.

Introduction

Plant-parasitic nematodes (PPN) infect a broad range of plants of agricultural and economic importance. They display a wide range of interactions with their hosts and many are biotrophic pathogens. The pine wood nematode, *B. xylophilus*, is a migratory endoparasitic nematode that causes extensive damage to forestry across many parts of the world. The life cycle of this nematode is complex and includes fungal- and plant- feeding stages, as well as a stage that is vectored to new hosts by an insect, most often the longhorn beetle *Monochamus* spp. (reviewed by Jones *et al.,* 2008). The fungal feeding stage of the nematode feeds on fungi present in dead or dying pine trees. As food availability declines, the nematode enters a survival stage which locates pupae of *Monochamus* and settles within the tracheae or beneath the elytra of the adult beetle as it emerges from the pupal chamber. The beetle may migrate to another tree colonized by fungi or may feed on living trees. In the latter case the nematode leaves the beetle and infects the host tree, feeding on parenchymal and epithelial cells. Nematodes migrate, feed and reproduce within the host causing extensive damage both directly, due to their feeding activities, and indirectly as a result of disruption of water transport due to cavitation in infected tissues. Under appropriate environmental conditions, most notably in hot climates, death of infected trees can occur within weeks of infection (Jones *et al.*, 2008; Mamiya, 2012).

Like other plant pathogens, the interactions of PPN with their host plants are mediated by effectors; secreted proteins originating from pharyngeal gland cells that are secreted into the host through the stylet (Haegeman *et al.*, 2012; Bird *et al.*, 2015). These proteins enable the nematode to successfully feed, reproduce and migrate inside the host. Advances in genomics and transcriptomics have allowed insights into the types of effectors required for parasitism by *B. xylophilus*. A range of plant cell-wall degrading enzymes and modifying proteins, which presumably facilitate invasion and migration, have been identified including cellulases (Kikuchi *et al.,* 2004), pectate lyases (Kikuchi *et al.,* 2006) and expansins (Kikuchi *et al.,* 2009). More recently, RNAseq analysis of nematodes after infection of trees revealed that a range of antioxidant and detoxification proteins are deployed as effectors during infection (Espada *et al.,* 2016). This analysis also identified a number of pioneer effector sequences that have no similarity to other previously identified sequences but that encode secreted proteins that are specifically expressed in the gland cells of the nematode. The importance of effectors in the life cycle of PPN has led to a range of approaches for their identification. Perhaps the most efficacious of the methods used to date has been direct analysis of the genes expressed in the pharyngeal gland cells. Initially this was achieved through Expressed Sequence Tag analysis of cloned cDNA made from RNA extracted from these tissues (*e.g.* Gao *et al.,* 2003; Huang *et al.,* 2003). A method was subsequently developed for microaspiration of gland cells followed by RNAseq analysis and has been used to identify effectors from a range of PPN (Maier *et al.,* 2013).

Genes encoding effectors of PPN are primarily expressed in the pharyngeal gland cells and this tissue specificity offers the possibility of using promoter sequences associated with expression in gland cells to identify candidate effectors. The feasibility of identifying genes associated with specific nematode tissues was demonstrated by the identification of DNA motifs present in the promoter regions of genes expressed in muscles of *C. elegans* and *C. briggsae* (GuhaThakurta *et al.,* 2004). Recently a DNA sequence motif (the DOGbox) was identified that is associated with genes expressed in the dorsal gland cell of the potato cyst nematode *Globodera rostochiensis* and was subsequently used to predict novel effectors which were validated by *in situ* hybridization (Eves-van den Akker *et al.*, 2016). Analysis of promoter elements therefore offers a powerful tool for identification of novel effectors.

In spite of the progress described above, our understanding of the effectors produced by *B. xylophilus* and the mechanisms by which it infects its hosts remain sketchy. The greatest progress in terms of identification and functional characterisation of effectors has been made with the sedentary endoparasitic cyst forming and root knot nematodes. *Bursaphelenchus xylophilus* is not directly related to either of these groups and has a very different mode of parasitism, suggesting that there is unlikely to be extensive overlap in effector repertoires used by these nematodes, something that has been borne out by the studies of *B. xylophilus* to date (Kikuchi *et al.,* 2011; Espada *et al.,* 2016). Here we identify a promoter element associated with genes expressed in the pharyngeal gland cells of *B. xylophilus* and use this to identify novel candidate

effector sequences from the *B. xylophilus* genome. We have also undertaken direct transcriptome analysis of the gland cells of this nematode and have combined the promoter and transcriptome datasets for verification and use this to identify a comprehensive effector list from this species.

Results

Identification of a DNA motif associated with genes expressed in the pharyngeal gland cells

Recent analysis of the genome sequence of *G. rostochiensis* allowed identification of a DNA sequence motif (the DOG box) associated with genes expressed in the dorsal gland cell (Eves-van den Akker *et al*., 2016) which has subsequently been used as a resource for predicting effectors in this species. We sought to determine whether a similar approach could be used to predict a motif associated with genes expressed in the gland cells of *B. xylophilus* which, although it is also a plant-parasite, is not directly related to *G. rostochiensis.* To identify potential regulatory elements associated with genes expressed in the gland cells, we used a training set based on a selected input list of 40 genes for which expression was previously validated in the gland cells (Kikuchi *et al*., 2004, Kikuchi *et al*., 2005, Kikuchi *et al*., 2006; Kang *et al*., 2012; Espada *et al*., 2016). These sequences included a range of plant cell wall-degrading enzymes as well as novel effectors identified in our previous work. This analysis identified one DNA sequence motif that was represented in 70% of the input list genes and peaked between 40 to 70bp upstream of the coding region (Figure 1A). This sequence motif (STATWWAWRS**)** has 6 variable *loci* indicated by the correspondent nucleotide according to the DNA ambiguity code ([C|G]TATA[T|A]AA[G|A][C|G]). This STATWWAWRS motif is present upstream of 597 of the genes predicted in the *B. xylophilus* genome. The most represented in this list are peptidases (cysteine or aspartic families), and genes without sequence similarity to others in databases-pioneer genes. The vast majority (556) of these genes have a single occurrences of the motif in their promoter region (Figure 1B). We used the TATA box motif (TATAAA), which is present in many eukaryotic organisms, as a control for the DNA motif analysis. The TATAAA motif was present in 6417 *B. xylophilus* genes and its position in relation to the start codon of associated genes was different to that of the STATWWAWRS motif described here (Figure 1A). Effectors are secreted proteins and therefore almost always have a signal peptide for secretion at their N-terminus. We therefore analysed the genes associated with the STATWWAWRS motif in order to determine whether they were enriched for predicted secreted proteins. 34% of these sequences have a predicted signal peptide, compared to 15.6% of those associated with the TATAA box and 12.7% of all of the genes predicted in the *B. xylophilus* genome (Figure 1C). Secreted proteins are therefore over-represented in the sequences that carry the STATWWAWRS motif, as would be expected if this motif is associated with genes expressed in the secretory gland cells.

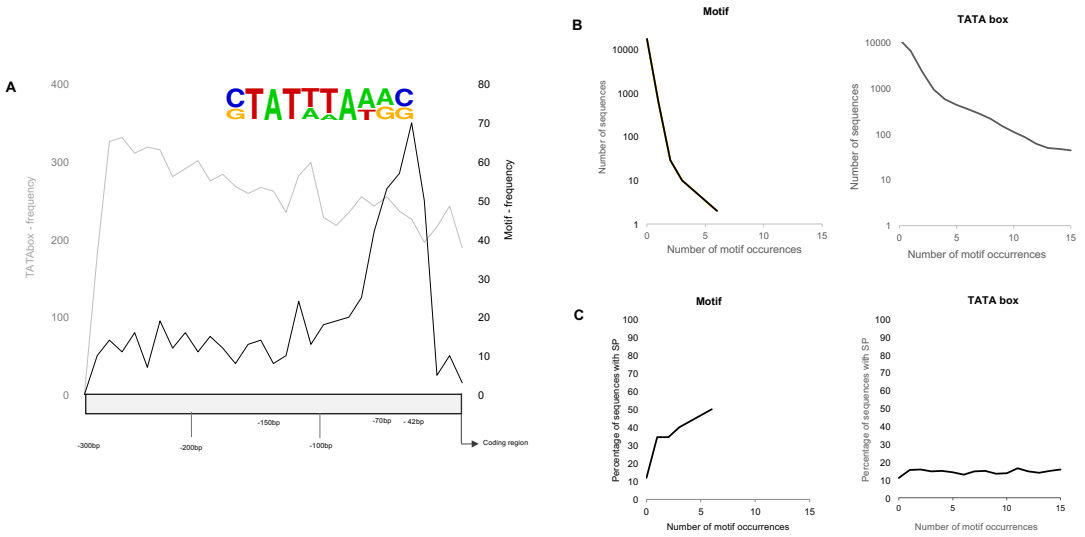


Figure 1 – Features of the DNA sequence motif [STATWWAWRS] and associated genes in the *B. xylophilus* genome. (A) Sequence logo of the motif (in colour) showing the consensus and variable sites. The motif peaks around 70bp upstream of the coding region and TATA box peaks around the first 20 bp upstream the coding region (grey line); (B) Number of sequences with the motif and the number of occurrences as compared to the TATA box sequence; (C) Proportion of genes with the motif or TATA box that have a signal peptide.

In order to demonstrate that the STATWWAWRS motif can act as a predictor of sequences expressed in the gland cells of *B. xylophilus* we used *in situ* hybridization to examine expression of novel genes (*i.e.* those that had not previously been studied) in mixed stage nematodes. This analysis showed that genes for which a signal was detected were expressed specifically in the gland cells (Figure 2).

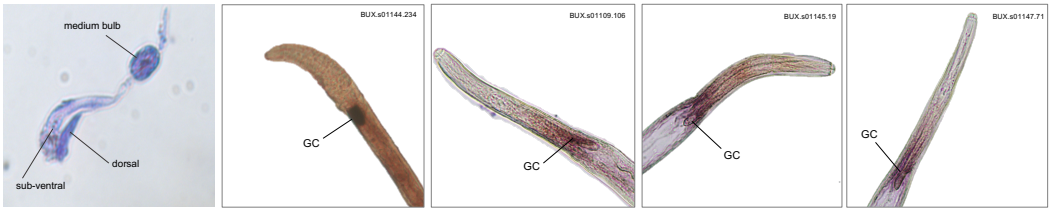


Figure 2 – *B. xylophilus* pharyngeal gland cells obtained by microaspiration (left). Localisation of candidate effector genes in the pharyngeal gland cells (GC) by *in situ* hybridization. Each of these genes is associated with the STATWWAWRS motif. Sequence similarity analysis with BlastP showed that BUX.s01144.234 have sequence similarity to a thaumatin-like protein, BUX.s01109.106 and BUX.s01147.71 have similarity to a transthyretin-like protein and BUX.s01145.19 is similar to a lipase found in *C. briggsae* (CBR-LIPL-1).

The STATWWAWRS motif is unrelated in sequence to the DOG box predicted for *G. rostochiensis*. We subsequently analysed the presence and number of occurrences of the STATWWAWRS motif in *G. rostochiensis* and a RKN, *Meloidogyne hapla*. Although the STATWWAWRS motif is present in the promoter region (up to 500 base pairs) of genes from both

RKN and CN the presence of the motif is not correlated with the presence of a signal peptide (Figure 3).

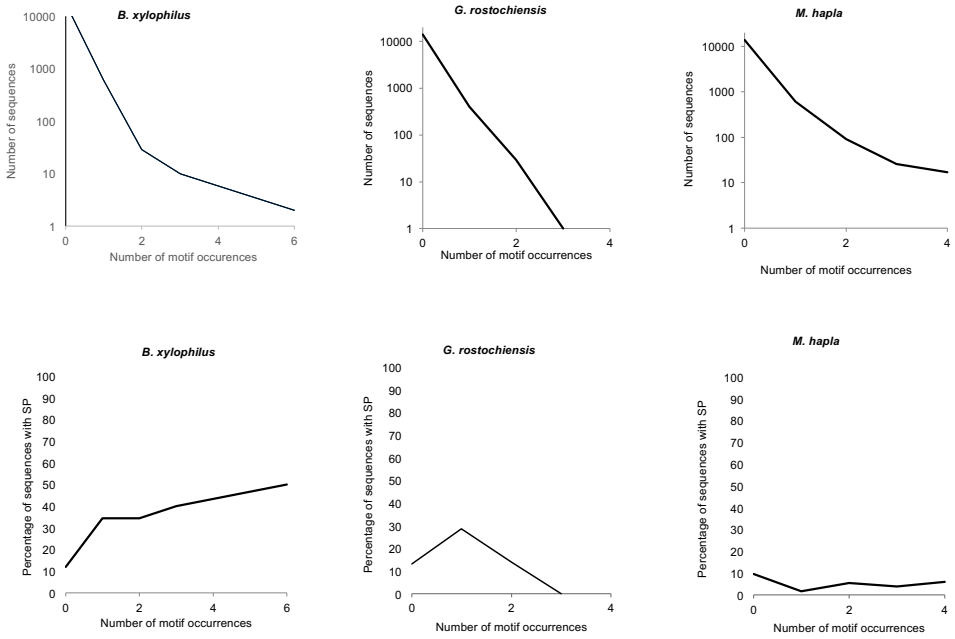


Figure 3 - The STATWWAWRS motif in other sedentary plant-parasitic nematodes. The motif is present in *Globodera rostochiensis* and *Meloidogyne hapla* but it is not associated with the presence of a signal peptide (SP) (bottom graphs).

Analysis of the transcriptome of purified B. xylophilus pharyngeal gland cells

A method for extraction and sequencing of RNA from dissected gland cells of plant-parasitic nematodes was recently reported (Maier *et al.,* 2013). We therefore sought to further validate the STATWWAWRS motif and expand the repertoire of effectors from *B. xylophilus* by using this approach. Gland cells were dissected from the nematode (Figure 2) and used for RNA extraction. The two RNAseq replicates (BX-1 and BX-2) yielded 124 and 143 million reads respectively and in each case approximately 30% of the reads mapped to the *B. xylophilus* genome (Table 1). This relatively low mapping rate is probably related to the amplification required to generate sufficient material for sequencing with the majority of the unmapped reads derived from RNA used for removal of rRNA from the sample (not shown). However, the depth of read coverage obtained allowed subsequent analysis of genes present in the gland cells.

Table 1 – RNAseq mapped data from the two samples sequenced (BX-1, BX-2).

	Trimmed reads input	**Mapped reads**
Bx-1	124218810	37820234 (*30.4% of input*)
Bx-2	143259452	34916231 (*24.4% of input*)

In order to provide an indication of expression levels, the genes that had RNAseq reads mapped to them were sorted into bins according to their depth of RNAseq coverage (0, 2, 10, 100, 1000, 10000 and 20000 FPKM). The proportion of the sequences that have a predicted signal peptide arose with increasing expression level (Figure 4) and was significantly higher than that found across all *B. xylophilus* sequences in the three highest expression bins. Abundant representation in the gland cell transcriptome is therefore positively correlated with the presence of a signal peptide. A comparison of the sequences carrying the STATWWAWRS motif and those in the gland cell transcriptome showed that more than half of the sequences with the motif were represented in the RNAseq dataset (not shown). Encouragingly, the presence of the motif is correlated with expression level in the gland cell transcriptome, with sequences in the three highest expression bins more likely to have at least one interaction of the STATWWAWRS motif upstream of their coding region (Figure 4). Abundant representation in the gland cell transcriptome is therefore correlated with increased probability of both a signal peptide and the STATWWAWRS motif.

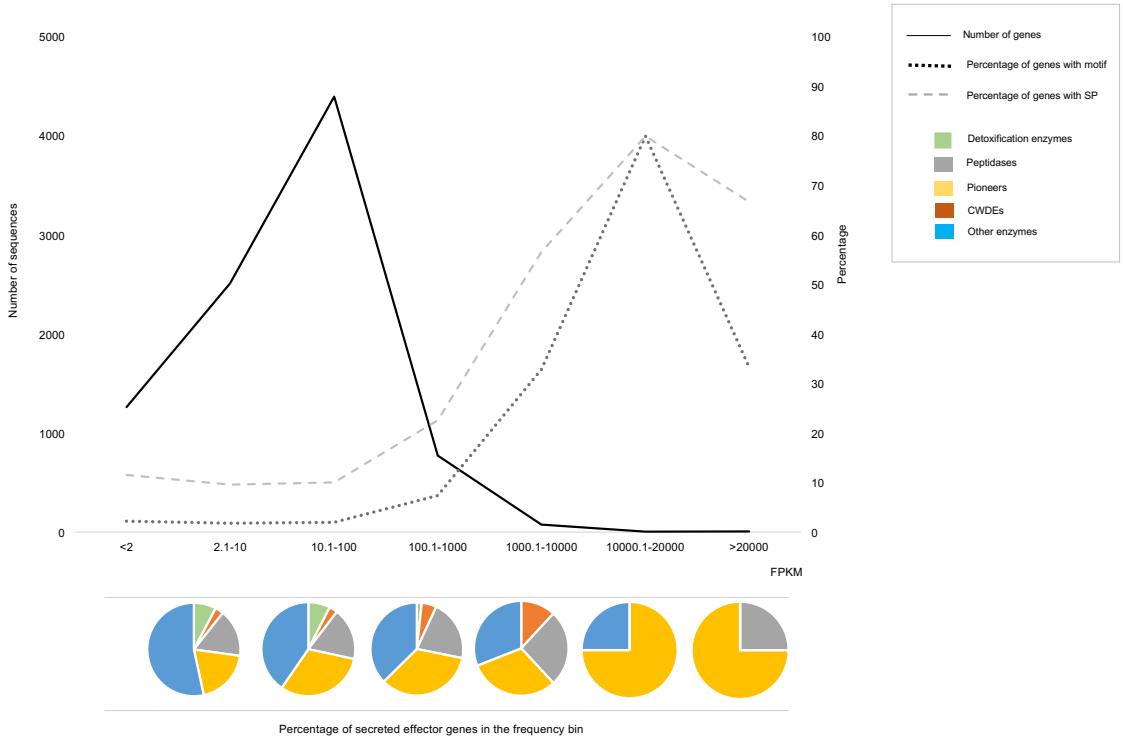


Figure 4 – Distribution of the genes expressed in the GC and percentage of those genes that have SP. Abundance of the expressed genes in the GC data set (black line) and the percentage of secreted proteins (grey line). The normalized expression profile are distributed in bins according to the value of FPKM (Fragments per kilobase million). The presence of signal peptide is increased in the most abundant expressed gland cell genes. The pie charts represent the distribution (percentage of genes) of the different gland cell- expressed proteins within the most abundantly secreted genes. SP: signal peptide. “Other enzymes” include several different proteins like allergen v5/Tpx, Metridin toxins, thaumatin, transthyretin, fatty acid metabolism, lysozyme, calreticulin and different GH families, FMRF amide-related, saposin B and galectin.; “CWDEs“ include PEL, endoglucanases and chitinase; “Detoxification enzymes” included carboxylesterase B, thioredoxin, oxireductase, GST and others; “Peptidases” included aspartic, cysteine, serine and metallo families and proteinase inhibitor.

Further validation and analysis of the motif

The availability of the gland cell transcriptome allowed us to investigate the presence of other promoter sequences associated with gland cell genes. For this, the 30 most abundant genes (Table S2) in the gland cell transcriptome were analysed as described above for the presence of DNA motifs in their promoter regions. The STATWWAWRS motif had the best representation in this input dataset, being present 40% of the genes. Since there is only limited overlap between this gene set and the original training set used to identify the STATWWAWRS motif, this provides a strong validation for the association of the motif with genes expressed in the gland cells. No other novel motifs were revealed as a result of this analysis. The STATWWAWRS motif has four variable *loci* meaning that a number of potential variants of this sequence are present in *B. xylophilus*. We analysed each of the variants present individually but found no specific patterns of association with specific gene classes (not shown). In addition, no individual variant showed significantly greater association with secreted proteins or with genes present in the gland cell transcriptome than the consensus sequence itself.

Candidate effectors of B. xylophilus

The transcriptome data and list of genes associated with the STATWWAWRS motif were subsequently used to generate a list of *B. xylophilus* genes that are likely to encode effectors. Genes that are included on this list were represented in the gland cell transcriptome dataset, encode a protein with a predicted signal peptide at the N-terminus and have at least one occurrence of the STATWWAWRS motif in the genomic region 1000bp upstream of the predicted start codon. A total of 118 sequences are present on the resulting *B. xylophilus* effector list (Table S3). A comparison with our previous transcriptome analysis of this nematode shows that almost half of these sequences are upregulated *in planta* (Figure 5). In addition, 42% of these effector sequences were identified in collected secreted proteins of *B. xylophilus* in a proteomic analysis of secreted proteins (Shinya *et al.,* 2013), a considerable enrichment compared to the 8.4% of all *B. xylophilus* proteins that were identified in this analysis.

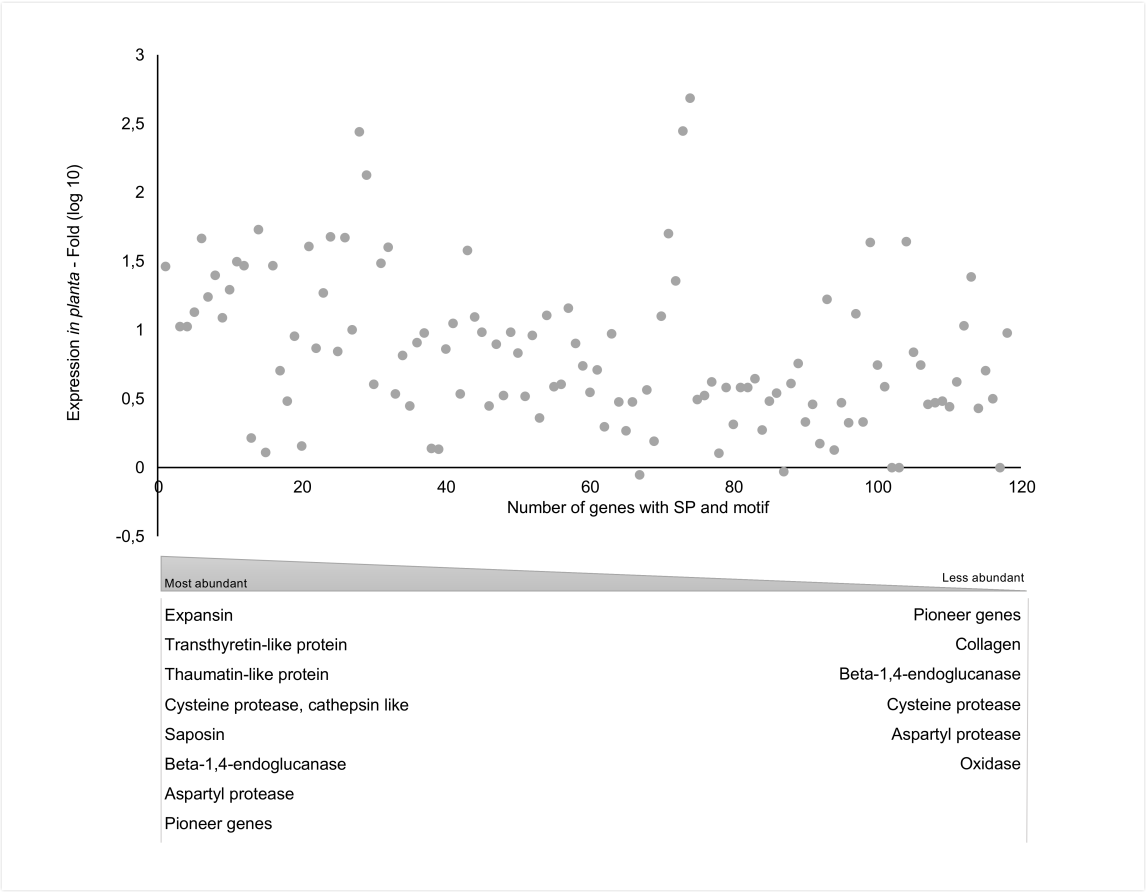


Figure 5 – Comparison between the transcriptomic profile (fold ten of the expression upon infection of the plant host) and the most abundant genes with signal peptide and presence of the STATWWAWRS motif.

The *B. xylophilus* effector list includes many previously verified effectors, including several cellulases, beta-1,3-endoglucanases, pectate lyases, expansins (one of which is the most abundant sequence in the gland cell transcriptome dataset), venom allergen proteins and several pioneer sequences for which gland cell expression was subsequently experimentally verified (Figure 2, Figure 4). This provides a level of reassurance that the effector list reflects the biology of *B. xylophilus* and that as yet uninvestigated sequences merit further study. In addition, several proteinases and three transthyretin-like proteins (including the second most abundant sequence in the gland cell transcriptome) are present on the list. More than half of the sequences on the effector list are pioneers, which have no similarity to other sequences in various publically accessible databases. This reflects similar studies on other plant-parasitic nematodes which have shown that a large proportion of effectors are novel sequences (reviewed by Kikuchi *et al.,* 2017).

This list is unlikely to include all effectors, as some genuine effectors may not have been expressed at the time we sampled gland cell material and/or may be under the control of other as yet undetected gland cell promoters. However, this provides a comprehensive and robust list of candidate effectors from *B. xylophilus* for future studies.

Discussion

Effector identification and prediction are key issues to understand the mechanisms underlying parasitism. In the last decades, there have been considerable efforts made in finding new approaches to identify new parasitism proteins and further understand their functional role in disease. For PPN these include direct sequencing of gland cell RNA and analysis of promoter elements. In this study, we have combined these two approaches to identify a high-confidence effector list from *B. xylophilus*. Using a validated pharyngeal gland cell effector subset, we identified a putative regulatory element (STATWWAWRS) that is associated with expression in the gland cells. This sequence was distinct from the DOG box sequence identified from *G. rostochiensis* (Eves-van den Akker *et al*., 2016). The new STATWWAWRS motif is present in all three PPNs that we examined; *B. xylophilus*, *Meloidogyne hapla* and *Globodera rostochiensis*. However, the motif is preferentially associated with secreted proteins in *B. xylophilus*, but not in RKN or CN, suggesting that it is not related to expression in the gland cells in these other species. Similar comparative studies in other nematodes may be informative. However, there are no published genome sequences for related species such as *B. mucronatus*. Although other migratory endoparasitic nematodes have been sequenced, the assemblies are not publically available. In addition, although these nematodes share some features of parasitism with *B. xylophilus* they are not directly related and the presence of a similar motif is therefore unlikely.

Following the identification of the STATWWAWRS motif we were able to confirm that previously uncharacterized genes associated with the motif were expressed in the gland cells. This confirms that, as for the DOG box of *G. rostochiensis,* the STATWWAWRS promoter motif can be used to predict new candidate effectors from *B. xylophilus*. A similar approach, based on identification of promoters associated with genes expressed in specific tissues and/or at specific life cycle stages, may be of benefit if applied to other pathogens for which identifying effectors is difficult. For example, although it is relatively straightforward to identify effectors from oomycetes (based on the presence of an RxLR motif associated with a signal peptide) (Whisson *et al*., 2007) and bacterial plant pathogens (based on the presence of a type 3 signal sequence) (Alfano and Collmer, 2004), identifying effectors from fungal plant pathogens is considerably more difficult as no protein motif is known that is associated with effectors from these organisms. Given a sufficiently robust training set of known effectors it may be feasible to identify novel fungal effectors using this approach.

We also applied an alternative approach to identification of *B. xylophilus* effectors by sequencing RNA extracted directly from dissected gland cells. This approach has been used successfully with other PPN (Maier *et al.,* 2013). The main benefit of this approach is the ability to directly analyse gland cell materials, giving a high probability of identifying genuine effectors. This is reflected by the presence of known effectors amongst the most abundant sequences represented in the sequences obtained from *B. xylophilus* gland cells. This provides a degree of confidence that other abundantly represented secreted proteins of unknown function merit further

investigation. Several sequences in this category were subsequently validated as effectors by *in situ* hybridisation. However, it is clear that this approach will not identify all effectors and suffers from other potential drawbacks. Extraction of gland cells and purification of RNA is extremely technically challenging. In addition, we were unable to sample parasitic stage nematodes extracted from trees. Both of these factors mean that it is very unlikely that all gland cell genes will be sampled. In addition, no gland cell purification method will work perfectly and contamination with other nematode body parts is likely. False positives are therefore likely to be included in the dataset. However, these data provide a valuable resource from which effector candidates can be identified.

Given the potential drawbacks of each approach used in isolation, we aimed to identify a high confidence list of effectors by bringing together the gland cell transcriptome and promoter data. This analysis demonstrated that the two datasets cross validated one another as sequences with the motif were over-represented in the gland cell dataset and the presence of the motif was linked to abundance in the gland cell dataset. The final list of effectors consisted of 118 sequences that are represented in the gland cell transcriptome with the motif in the region upstream of the coding sequence and which had a predicted signal peptide. Many of the sequences (approx. 30%) on this list are pioneer sequences that have no sequence similarity to others characterized in the databases. This is in keeping with studies on other PPN which have shown that a large proportion of effectors are novel sequences. For example, 38 of 53 confirmed effectors of *H. glycines* and 28 of 37 effectors from *M. incognita* identified in the first studies of these nematodes were pioneers (Gao *et al.,* 2003; Huang *at al.,* 2003). Similarly, analysis of *G. pallida* (Thorpe *et al.,* 2014) and *G. rostochiensis* (Eves-van den Akker *et al.,* 2016) genome sequences suggests that there is limited overlap between effectors from CN and RKN sequences. Some of the other sequences on the effector list are consistent with a role in parasitism, and include cell wall degrading enzymes, proteinases and venom allergen proteins. In addition, several different transthyretin-like sequences are present. Similar sequences are present in many nematodes, often as large gene families of secreted proteins; *C. elegans* contains more than 60 such sequences. Although a small number of transthyretin-like proteins have been identified as being expressed in the gland cells of several different PPN (reviewed in Haegeman, *et al*., 2012) their function remains unknown.

The developments described in this study represent a unique opportunity to develop a better understanding of the mechanisms underpinning infection of plants by migratory plant-parasitic nematodes. Moreover, these genes represent potential new targets for control of PPN.

Experimental procedures

In silico *identification of DNA motifs in the promoter regions*

To identify putative promoters, sequences up to 1000bp upstream of the predicted start codon of *B. xylophilus* genes were extracted from the genome data for this species (available at www.genedb.org) using the script get_up_stream_region.py (https://github.com/peterthorpe5). To identify potential motifs associated with effectors a list of verified effectors were used as the input list for motif identification in the motif discovery algorithm HOMER (Heinz *et al*., 2010). We searched for motif occurrences in the promoter region of all PWN genes using the FIMO webserver (http://meme-suite.org/tools/fimo). The number of occurrences of the identified motif was analysed for each gene. The presence of a signal peptide in the associated genes was analysed using SignalP version 4.1 (Petersen *et al.,* 2011).

In situ *hybridisation*

The spatial expression patterns of selected genes associated with the predicted motif and/or that were present in the gland cell transcriptome dataset (below) was determined by *in situ* hybridisation as previously described (de Boer *et al*., 1998; Espada *et al.*, 2016). The primers used for this analysis are shown in the Table S1.

Microaspiration of pharyngeal gland cells from B. xylophilus

A Portuguese isolate of *B. xylophilus* was cultured on *Botrytis cinerea* in flasks for 7 days at 25C (Evans *et al*., 1970). Mixed life stages of the nematodes were collected using the Baerman funnel technique (Evans, 1970) and fixed in 100% ethanol at -80C overnight. Fixed nematodes were cut into pieces using a razor blade. Fixed, cut nematodes were stained in Histogene acid staining solution (for nucleic acids) and resuspended in Halocarbon oil 700 (Sigma). The stained tissues were spread on RNAse free glass cover slips and stored at -80C before further processing. Microaspiration of the pharyngeal gland cells was performed under vacuum on an inverted microscope as previously described (Maier *et al.,* 2013) and extracted gland cells were stored under oil at -80C before RNA extraction (Maier *et al.,* 2013).

RNA sequencing

Total RNA was isolated from approximately 200 mixed dorsal and subventral gland cells using the Arcturus PicoPure RNA isolation kit (Thermofisher Scientific) following the manufacturer's instructions. Approximately 6ng of total RNA was isolated from these gland cells following this process. The total RNA was amplified before sequencing and subsequently used for library preparation with the SMARTer Stranded total RNA-seq Pico input kit (Clontech, USA). The quality of the RNA was assessed using a Bioanalyzer. Two paired-end libraries (BX-1 and BX-2) were sequenced using the NextSeq Illumina service from Admera Health (USA). These two technical replicates represent one biological replicate (gland cells from mixed life-stage nematodes). The run was spiked with 15-20% PhiX.

Analysis of gland cell transcriptome

The RNAseq data from the two libraries generated approximately 268 million paired end reads per library. The reads were trimmed of adapter sequences and low quality bases (Phred < 25) and aligned to the *B. xylophilus* genome using Tophat2 (Kim *et al.,* 2013). Version 1.2 of the genome was used to build the reference genome and is available at Gene DB (http://www.genedb.org/Homepage/Bxylophilus) (Kikuchi *et al*., 2011). The number of reads aligned to each gene were counted using Bedtools, and TMM normalized as described in Espada *et al*. (2016). *B. xylophilus* genes were sorted into bins of ascending numbers of reads mapped (e.g. 0, 2, 10, 100, 1000, 10000 and >20000 FPKM, Fragments per kilobase million) as we hypothesized that since effectors are produced in large quantity by the nematode, genes that are abundantly represented in the gland cell transcriptome were more likely to be effector candidates. The proportion of sequences in each bin with a signal peptide (identified using SignalP v4.1) was compared to the proportion of secreted proteins in the whole genome (Bendtsen *et al*., 2004). Secreted proteins above this threshold were selected for refinement of the regulatory element predictions. The analysis of gene/protein function was based on sequence similarity and performed against non-redundant database by BlastP and Blastn (*p-value*<1e-04), using a local installation of the Galaxy platform (Cock *et al.,* 2014).

References

Alfano, J.R. and Collmer, A. (2004) Type III secretion system effector proteins: double agents in bacterial disease and plant defense. *Annu. Rev. Phytopathol.* 42, 385–414.

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

Bird, McK.D., Jones, J.T., Opperman, C.H., Kikuchi, T. and Danchin, E.G.J. (2015) Signatures of adaptation to plant parasitism in nematode genomes. *Parasit.* 142, S71–S84.

Cock, P.J.A. and Pritchard, L. (2014) Galaxy as a platform for identifying candidate pathogen effectors. In Plant-pathogen Interactions: methods and protocols, 2nd edition. Methods in Molecular Biology volume 1127 (Birch, P., Jones, J. and Bos, J., eds), pp. 3-15.

de Boer, J.M., Yan, Y., Smant, G., Davis, E.L. and Baum, T.J. (1998) *In-situ* hybridization to messenger RNA in *Heterodera glycines. J Nemat.* 30, 309-12.

Espada, M., Silva, A.C., Eves-van den Akker, S., Cock, P.J.A., Mota, M. and Jones, J.T. (2016) Identification and characterization of parasitism genes from the pinewood nematode *Bursaphelenchus xylophilus* reveals a multi-layered detoxification strategy. *Mol. Plant Pathol.* 17, 286-295.

Evans, A.A.F. (1970) Mass culture of mycophagous nematodes. *J Nematol.* 2, 99–100.

Eves-van den Akker, S., Laetsch, D. R., Thorpe, P., Lilley, C.J., Danchin, E.G.J., da Rocha, M., Rancurel, C., Holroyd, N.E., Cotton, J.A., Szitenberg, A., Grenier, E., Montarry, J., Mimee, B., Duceppe, M., Bo+, I., Marvin, J.M.C., Jones, L.M., Yusup, H.B., Lafond-Lapalme, J., Esquibet, M., Sabeh, M., Rott, M., Overmars, H., Finkers-Tomczak, A., Smant, G., Koutsovoulos, G., Blok, V., Mantelin, S., Cock, P.J.A., Phillips, W., Henrissat, B., Urwin, P.E., Blaxter, M. and 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.

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

GuhaThakurta, D., Lawrence, A.S., Waterston, R.H., Stormo, G.D. (2004) Novel transcription regulatory elements in *Caenorhabditis elegans*. *Genome Res.* 14(12), 2457–2468.

Haegeman, A., Mantelin, S., Jones, J.T. and Gheysen, G. (2012) Functional roles of effectors of plant-parasitic nematodes. *Gene* 492, 19-31.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H. and Glass, C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38(4), 576-589.

Huang, G., Gao, B., Maier, T., Allen, R., Davis, E.L., Baum, T.J. and Hussey, R.S. (2003) A profile of putative parasitism genes expressed in the esophageal gland cells of the root-knot nematode *Meloidogyne incognita*. *Mol Plant Microbe Interact*. 16(5), 376-381.

Jones, J.T., Li. H., Moens, M., Mota, M. and Kikuchi, T. (2008) *Bursaphelenchus xylophilus*: opportunities in comparative genomics and molecular host-parasite interactions. *Mol Plant Pathol* 9, 357-368.

Kang, J.S., Koh, Y.H., Moon, Y.S. and Lee, S.H. (2012) Molecular properties of a venom allergen-like protein suggest a parasitic function in the pinewood nematode *Bursaphelenchus xylophilus*. *Int J Parasitol.* 42, 63-70.

Kikuchi, T., Jones, J.T., Aikawa, T., Kosaka, H. and Ogura, N. (2004) A family of GHF45 cellulases from the pine wood nematode *Bursaphelenchus xylophilus*. *FEBS Lett.* 572, 201-205.

Kikuchi, T., Shibuya, H. and Jones, J.T. (2005). Molecular and biochemical characterization of an Endo-ß-1,3-glucanase from the pine wood nematode *Bursaphelenchus xylophilus. Biochem. J.* 389, 117-125.

Kikuchi, T., Shibuya, H., Aikawa, T. and Jones, J.T. (2006) Cloning and characterization of pectate lyases expressed in the esophageal gland of the pine wood nematode *Bursaphelenchus xylophilus*. *Mol Plant Microbe Interact.* 19(3), 280–287.

Kikuchi, T., Li, H., karim, N., Kennedy, M.W., Moens, M. and Jones, J.T. (2009) identification of putative expansin-like genes from the pinewood nematode, *Bursaphelenchus xylophilus*, and evolution of the expansin gene family within Nematoda. *Nematology* 11, 355-364.

Kikuchi, T., Cotton, J.A., Dalzell, J.J., Hasegawa, K., Kanzaki, N., McVeigh, P., Takanashi, T., Tsai, I.J., Assefa, S.A., Cock, P.J.A., Otto, T.D., Hunt, M., Reid, A.J., Sanchez-Flores, A., Tsuchihara, K., Yokoi, T., Larsson, M.C., Miwa, J., Maule, A.G., Sahashi, N., Jones, J.T. and Berriman, M. (2011) Genomic insights into the origin of parasitism in the emerging plant pathogen *Bursaphelenchus xylophilus*. *PLoS Pathogens* 7, e1002219.

Kikuchi, T., Eves-van den Akker, S. and Jones, J.T. (2017) Genome evolution of plant-parasitic nematodes. *Annu. Rev. Phytopathol*. 55, 333-354.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S.L. (2013) Tophat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 14:R36.

Maier, T.R., Hewezi, T., Peng, J.Q. and Baum, T.J. (2013) Isolation of whole esophageal gland cells from plant-parasitic nematodes for transcriptome analyses and effector identification. *Mol Plant Microbe Interact.* 26, 31-35.

Mamiya, Y. (2012). Scanning electron microscopy of pine seedling wood tissues sections inoculated with the pinewood nematode, *Burspahelenchus xylophilus*, previously prepared for light microscopy. *J. Nematol* 44, 255-259.

Petersen, T.N., Brunak, S., von Heijne, G. and Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8: 785–786.

Shinya, R., Morisaka, H., Kikuchi, T., Takeuchi, Y., Ueda, M., and 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(6), e67377.

Thorpe, P, mantelin, S., Cock, PJ., blok, V.C., Coke, M.C., Eves-van den Akker, S., Guzeeva, E., Lilley, C.J., Smant, G., Reid, A.J., Wright, K.M., Urwin, P.E and Jones, J.T. (2014) Genomic characterization of the effector complement of the potato cyst nematode *Globodera pallida*. *BMC Genomics* 15:923.

Whisson, S.C, Boevink, P.C., Moleleki, L., Avrova, A.O., Morales, J.G., Gilroy, E.M., Armstrong, M.R., Grouffaud, S., van West, P., Chapman, S. Hein,I., Toth, I.K., Pritchard, L. and Birch, P.R.J. (2007) A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature* 450, 115-118.

Supporting information

Table S1 - Primers used for ISH validation.

Gene	Forward primer	Reverse primer
BUX.s01144.234	TGTCAAGATCACGGTCGTCA	TTCCACAAGCACCAGTTTCG
BUX.s01109.106	TATGACGTGGACACCCTCAG	GCGGCCTTGTCAGATTCTTT
BUX.s01145.19	ACAGCTGCCCCAATGATTAC	CCGCATTGATTACGTTGATG
BUX.s01147.71	CAAGGAGTAGCGGTGAGAGG	TTCTCAGTTCGGGTTCGATT

Table S2 - Top 30 expressed genes in the gland cells tissues. SP: presence or absence of signal peptide; motif: presence of at least one repetition of the STATWWAWRS motif; ISH: validated the spatial expression; NA: no signal; GC; signal in gland cells.

Gene ID[1]	Gene expression (FPKM)	SP	Motif	ISH[2]	Secretome[3]	Expression post infection[4]	Gene Description (based on sequence similarity)
BUX.s01226.3	43610,67	+	+			-	Reverse transcriptase
BUX.s01226.2	34330,2	+	-			DPI	Reverse transcriptase
BUX.s00083.32	28320,67	+	-		+	FF	Cysteine peptidase, family C1A (Proteinase inhibitor I29)
BUX.s01513.259	28133,97	-	-			DPI	Not known
BUX.s00364.143	20947,55	-	-			DPI	Hypothetical protein; Metridin-like ShK toxin
BUX.s01281.223	20907,19	+	+	GC		DPI	Not known
BUX.s00782.2	16572,97	-	+			FF	Small HSP21-like protein; HSP20-like
BUX.s01063.193	14530,48	+	+			-	Transthyretin-like family protein
BUX.s01109.570	13699,99	+	-	NA		DPI	Not known
BUX.s01332.1	12613,75	+	-			DPI	Not known
BUX.s01639.10	12613,75	+	-			-	Not known
BUX.s01167.27	9072,31	-	-			DPI	Not known
BUX.s01144.234	8343,09	+	+	GC	+	DPI	Thaumatin-like protein 1b
BUX.s01144.128	8146,98	+	-		+	DPI	Not known
BUX.s00532.10	7085,05	+	+	NA		DPI	Aspartic Peptidase, family A1
BUX.s00036.112	6830,64	+	+	GC	+	DPI	Beta-1,4-endoglucanase; GH45
BUX.s01259.45	6760,58	+	+	NA	+	DPI	Cysteine protease family cathepsin 1; Proteinase inhibitor I29
BUX.s01143.167	6745,18	-	-			DPI	Not known
BUX.s00647.61	5542,85	+	+		+	DPI	Not known
BUX.s01167.26	5473,3	-	-			DPI	
BUX.s01109.169	5433,02	+	-			-	Saposin B domain
BUX.s01226.4	5245	-	-			DPI	Hypothetical protein - common roundworm retrotransposon R4
BUX.s01226.1	5037,37	-	-			DPI	Hypothetical protein - common roundworm retrotransposon R4
BUX.s00036.113	4504,83	+	+	GC	+	DPI	Beta-1,4-endoglucanase (B. xylophilus)
BUX.s00298.157	4309,6	-	-			DPI	Eukaryotic translation elongation factor 1A protein (B. xylophilus)
BUX.s01144.122	4148,47	+	+	GC		DPI	Not known
BUX.s00713.953	3997,31	+	+	NA	+	DPI	Peptidase aspartic, family A1
BUX.s00139.22	3881,41	+	+		+	-	Not known
BUX.s01147.119	3703,45	-	-			-	Calreticulin (B. xylophilus)
BUX.s01063.196	3408,99	+	-			DPI	Not known

[1] According to *Bursaphelenchus xylophilus* genome version 1.2 (available in Gene DB)

[2] *in situ* hybridization

[3] based on Shinya et al., 2013

[4] based on results from Espada *et al.*, 2016, *Mol. Plant Pathol.*

Table S3 - Candidate effectors from *B. xylophilus:* 118 genes are represented in the gland cell transcriptome and have both a signal peptide and the STATWWAWRS motf (most abundant to less abundant in gland cells from top left to bottom right).

GeneID	Upregulated in planta?	Expression in planta	GeneID	Upregulated in planta?	Expression in planta
BUX.s01281.223	+	103,9	BUX.s00713.837	-	112,17
BUX.s01063.193	-	1934,67	BUX.s01261.1	-	36,57
BUX.s01332.1	+	55,73	BUX.s00813.76	-	80,49
BUX.s01639.10	+	55,73	BUX.s00060.1	+	233,06
BUX.s01144.234	+	171,06	BUX.s01281.7	-	1,90
BUX.s00532.10	+	128,23	BUX.s00579.150	-	170,53
BUX.s00036.112	+	393,48	BUX.s01518.89	-	49,37
BUX.s01259.45	+	84,44	BUX.s00351.204	-	51,01
BUX.s00647.61	+	297,20	BUX.s00579.74	-	0,70
BUX.s00036.113	+	322,72	BUX.s00466.37	-	222,01
BUX.s01144.122	+	127,83	BUX.s01145.19	+	38,79
BUX.s00713.953	+	176,45	BUX.s01066.142	+	78,13
BUX.s00139.22	-	2224,90	BUX.s01109.178	+	352,81
BUX.s01147.176	+	85,58	BUX.s01283.1	+	10,72
BUX.s01662.95	-	251,02	BUX.s01281.78	+	25,91
BUX.s01147.177	+	119,82	BUX.s00358.16	-	339,34
BUX.s01066.8	+	987,43	BUX.s01281.179	-	17,43
BUX.s01063.106	-	45,41	BUX.s00139.149	-	557,60
BUX.s01066.63	+	138,92	BUX.s01092.187	-	148,33
BUX.s01144.305	-	25,97	BUX.s01109.106	-	364,02
BUX.c07686.1	+	5,18	BUX.s00240.36	-	126,94
BUX.s00713.1076	-	30,03	BUX.s01063.30	-	158,73
BUX.s01281.215	+	52,28	BUX.s01147.71	-	175,78
BUX.s01259.20	+	209,06	BUX.c08843.2	-	7,44
BUX.s01259.83	-	21,80	BUX.s00240.34	-	163,98
BUX.s01259.22	-	6,95	BUX.s00713.1016	-	230,36
BUX.s00116.606	+	33,60	BUX.s00116.696	-	16,93
BUX.s01147.175	+	17,44	BUX.s01149.55	-	19,42
BUX.s01259.23	+	34,02	BUX.s01259.24	-	93,17
BUX.s01254.165	-	122,36	BUX.s01147.22	+	154,07
BUX.s01259.69	+	7,37	BUX.s00240.23	-	104,82
BUX.s00647.68	+	50,73	BUX.s00240.41	-	5,21
BUX.s00579.208	-	59,63	BUX.s00713.666	-	15,07
BUX.c08842.2	-	3,06	BUX.s01038.116	+	8,14
BUX.s01254.96	-	414,47	BUX.s01149.57	-	3,63
BUX.s00116.607	-	26,29	BUX.s01259.43	-	110,05
BUX.s00116.604	+	4,32	BUX.s00139.24	-	319,36
BUX.s01518.90	-	0,95	BUX.s00770.68	+	22,03
BUX.s01147.188	-	100,22	BUX.s00713.958	-	0,18
BUX.c08843.1		2,99	BUX.s01656.1	+	2,69
BUX.s00116.969	+	21,14	BUX.s01144.22	+	93,97
BUX.s00422.677	-	2685,78	BUX.s01281.180	-	17,24
BUX.s01281.230	+	98,77	BUX.c06430.1	-	22,98
BUX.s00116.597	+	1182,49	BUX.s00351.33	+	15,86
BUX.s01066.145	+	68,54	BUX.s01147.2	+	1,91
BUX.s01066.65	-	0,09	BUX.s00116.86	+	30,82
BUX.s00713.1002	+	28,14	BUX.s01147.178	-	2,03
BUX.s00358.21	-	536,25	BUX.s01063.187	-	0,88
BUX.s00117.41	+	35,10	BUX.s01518.99	-	0,57
BUX.c08842.1	+	13,08	BUX.s00803.9	-	68,01
BUX.s00358.19	-	551,47	BUX.s00116.893	-	209,28
BUX.s00364.45	+	13,86	BUX.s01066.1	-	9,76
BUX.s00036.107	-	0,52	BUX.s00813.52	+	34,10
BUX.s00116.596	+	199,43	BUX.s01038.234	+	71,58
BUX.s01661.67	-	31,49	BUX.s00460.290	-	11,91
BUX.s00579.44	-	30,91	BUX.s01254.196	+	31,11
BUX.s00460.341	+	242,32	BUX.s00713.223	-	6,31
BUX.s00579.188	+	1,52	BUX.s01115.1	-	0,00
BUX.s01659.4	+	4,02	BUX.s00813.51	+	13,77

CHAPTER VI

GENERAL DISCUSSION

The pinewood nematode, *B. xylophilus*, is a threat to forestry ecosystems worldwide. Many forestry pathogens are emerging and serious pests. In the native range of a pathogen, co-evolution ensures that the local tree population is able to resist or tolerate infection by local pathogens. However, increased global trade allows spread of pathogens to new regions, in which the trees have no prior exposure to the pathogen. The devastating effects of an introduced tree pathogen can be illustrated by the effects of Dutch Elm disease on the UK elm population (Harwood *et al.*, 2011) and, more recently, by the ash dieback pathogen (forestry.gov.uk). The potentially devastating effects of *B. xylophilus* can be seen from the impact of this pathogen on forestry in the Far East since its introduction in their early part of the 20th Century. In forests, pathogens can affect economic trade and, where serious losses occur, the ability of forestry ecosystems to store carbon, reduce flood risk or purify water (Boyd *et al.*, 2013).

In the context of PWD, the parasitic nematode is one key element that needs to be studied in terms of its interaction with the plant host. However, its phoretic interactions also need to be studied in order to gain a complete picture of the biology of disease. *B. xylophilus* is a migratory endoparasite with a unique and complex life cycle. It is both a plant cell and fungal feeder and has several modes, including dispersal and propagative stages, in the life cycle. It is one of the few plant-parasites that infects the stem of gymnosperms, most notably coniferous species of the genus *Pinus*. *B. xylophilus* is only distantly related to other migratory and sedentary plant-parasitic nematodes and has evolved the ability to infect plants stems independently. Although these features make this nematode a scientifically fascinating organism, they also make it one of the most challenging to study.

Neutral approaches towards identifying *B. xylophilus* effectors

This work represents the most comprehensive study to date on the molecular mechanisms underlying the interaction between *B. xylophilus* and its host and is the first comprehensive and pragmatic study of effector identification in this nematode. Previous studies on *B. xylophilus* effectors relied on *a priori* assumptions about the nature of effectors required by sedentary nematodes, which may not be appropriate (reviewed in Chapter I). Although this approach allowed identification of some *B. xylophilus* cell wall degrading enzymes it has also given rise to a number of studies on *B. xylophilus* orthologues of genes identified as effectors in root-knot and cyst nematodes. However, there is no guarantee that a protein adapted as an effector in root-knot or cyst nematodes has a similar function in *B. xylophilus.* In addition, the biology of *B. xylophilus* and the sedentary endoparasites is entirely different and these nematodes are not directly related. With these factors in mind, we aimed to identify novel effectors from *B. xylophilus* using non-biased *in silico* approaches based on changes in gene expression occurring during infection of the tree host and in the specialised parasitism tissues, as well as analysis of potential gene regulation elements. These approaches exploited the opportunities offered by using next

generation sequencing and allowed the identification/prediction of effectors and provided insights into the adaptations underlying parasitism of *B. xylophilus*.

We first compared replicated transcriptome datasets generated from nematodes grown *in vitro* on fungi (mycetophagous phase) and from nematodes after infection of trees (phytophagous phase) in order to identify secreted proteins upregulated after infection. Our analysis showed that expression levels of the genes can be significantly influenced by environmental factors, such as the time of the year (e.g., inoculation in different months), and that this can contribute to differences in the genes that are differentially expressed between conditions (data from Chapters II and III). Our analysis showed that this was a significant factor in two completely unrelated RNAseq datasets generated in different laboratories on different continents (Portugal and Japan). This shows that in future studies absolute consistency between samples is required in order to generate datasets that can be used to identify changes in gene expression occurring after infection. In spite of these issues, we were able to use these datasets to identify and predict new effector genes using a bioinformatic pipeline that identifies genes with a signal peptide at the N-terminus, that lack a transmembrane domain and that were upregulated *in planta*. Our confirmation of the gland cell localization of several genes identified using this approach suggest that it is an efficient method to identify nematode effectors. Similar approaches, identifying secreted proteins upregulated after infection as candidate effectors have been used for other pathogens such as fungi or oomycetes (Petre *et al.*, 2014; Sonah *et al.,* 2016; Sun *et al.*, 2017) and has also been shown to be effective in cyst nematodes (Thorpe *et al.,* 2014).

A more direct approach based on RNAseq analysis of material extracted directly from excised pharyngeal gland cells was also used to identify effectors in combination with promoter analysis (below). Sequencing of gland cell mRNA has the advantage of being a direct approach that is likely to generate useful data. A similar approach has been used for a variety of other plant-parasitic nematodes (Gao *et al*., 2003; Huang *et al*., 2003; Rutter *et al*., 2014). However, this is an extremely technically challenging procedure. The presence of known effectors amongst the most abundantly represented sequences in the gland cell RNAseq dataset is reassuring but it is clear that this approach does not identify all effectors. Genes that are only expressed after infection were not present in our dataset as we were not able to extract nematodes from trees for gland cell purification. This may explain the rarity of detoxification proteins in the gland cell RNAseq dataset, which our previous approach using RNAseq analysis of nematodes before and after infection showed to be important effectors.

The final approach was based on the presence of gene regulation elements upstream of effectors. We used a list of sequences that our previous studies had shown to be expressed in the pharyngeal gland cells as a training set to identify a promoter motif DNA sequence associated with effector genes. This was subsequently used to predict novel effectors, some of which were validated by *in situ* hybridization and by presence in the gland cell RNAseq dataset. The

identification of a conserved feature in the promoter/regulation region of effector genes provides a huge opportunity for studies in the genomics, evolution and understanding of the parasitism mechanism of this migratory nematode. Future studies could include the identification of transcription factors that bind to the promoter motif associated with effectors using, for example, mass spectrometry. Additionally, this finding provides the possibility of predicting new genes that are not identified using the bioinformatics pipelines described above, and represents a breakthrough in the study of the effectorome of parasitic nematodes. A similar approach allowed identification of a promoter motif from the cyst nematode, *Globodera rostochiensis* that is associated with genes expressed in the dorsal gland cell (Eves van den Akker *et al.,* 2016). This approach can be used to predict novel parasitism-related proteins and may allow discovery of regulation elements common to several effectors, it can be used to modify the expression of these genes *in planta*. The idea of manipulating the regulation of the expression of effector genes in order to disrupt infection and provide control is a new line of research that needs to be addressed.

Genomic organization of effectors in *B. xylophilus*

Next generation sequencing – genomic and transcriptomic data- allowed large-scale comparative studies to predict effectors and revealed some genetic adaptations to parasitism. In *B. xylophilus*, from the 18074 predicted proteins (according to version 1.2 of the genome, Kikuchi *et al*., 2011) approximately 12% have a signal peptide. This figure is consistent with the proportion of secreted proteins seen in other nematodes. However, within the genes upregulated in the phytophagous phase, a greater proportion of genes with a signal peptide is seen. A similar pattern occurs within the most abundantly expressed gland cell genes. With this study, we can conclude that the most highly abundant and expressed genes in parasitic stage can be identified by transcriptomic approach, through next generation sequencing platforms and using different bioinformatics scripts/tools to prioritise, select and analyse the data allows the identification of candidate effectors.

As seen in other PPNs the number of genes in *B. xylophilus* is smaller than *C. elegans* (18074 versus 20317predicted genes) which suggests that this feature is a consistent adaptation to parasitism (Kikuchi *et al*., 2017). One intriguing feature of our analysis of *B. xylophilus* effectors is the absence of expanded families of effectors. This is in contrast to other PPNs, most notably cyst nematodes, in which it is common to have large effector gene families. For example in *G. pallida*, a large expansion of both the SPRY domain effector and of glutathione synthetases is observed (Mei *et al*., 2015; Kikuchi *et al.,* 2017). These expanded families are a consequence of the genomic selection or may reflect the variability of target proteins in the various host plants of these nematodes. In the case of *B. xylophilus*, the nematode hosts are restricted to coniferous trees, specifically from the genus *Pinus* and this may explain the absence of gene families. An alternative explanation is that the nematode may not have a requirement to produce effectors that interact directly with host proteins to modify their function; cyst and root-knot nematodes are

biotrophic pathogens whereas *B. xylophilus* migrates and feeds destructively. Alternatively, *B. xylophilus* may have a broader panel of effectors rather than specific proteins for each function. It is notable however, that comparing the two transcriptomic studies (described in chapter II and III) showed that certain clusters of genes are specifically up regulated in all Japanese (JP) samples compared to Portuguese (PT), and vice versa. This may reflect host or population specific effectors. The differentially expressed genes that have a high expression in PT samples and low expression in JP samples include protein degradation enzymes (such as peptidases and cathepsin); TTR, pectate lyase, acid phosphatase and several pioneer proteins. Similarly, genes that have low expression in PT samples and high expression in JP samples include casein kinases, pectate lyases, several detoxification enzymes (e.g. GST, CYP-33C9, UGT, SDR), cysteine proteinases and several pioneer proteins among others. Some of these proteins have a potential role in parasitism, but the majority are unknown.

Our analysis shows that *B. xylophilus* has a range of genomic adaptations to parasitism. Characterisation of effectors reveals that:
- Large groups of peptidases and detoxification enzymes are highly upregulated after infection. Spatial expression profiles suggest that these are secreted in a way that provides multiple levels of protection to the nematode;
- Novel secreted proteins that are highly expressed upon infection are possibly species-specific;
- A specific regulation (promoter) element is associated with parasitism-genes expressed in the pharyngeal gland cell;
- 0.2% of *B. xylophilus* genes were acquired by horizontal gene transfer, including the plant and fungal cell wall degrading enzymes. These are important to its success inside the host;
- Morphological changes in the nematode body may be related to the downregulation of collagen proteins in the phytophagous phase compared to the mycetophagous phase.
- No expanded gene families are found in the *B. xylophilus* effector genes repertoire.

Phytophagous life-stage: effectors for successful parasitism

Migration in the host tissues

In the initial parasitic stage, the most abundant proteins expressed *in planta* are peptidases (mainly cysteine and aspartic families, but also metallo- and serine families) and a range of cell wall degrading enzymes. *B. xylophilus* was described as the PPN with the highest number of peptidases known to date (Kikuchi *et al.,* 2011); with aspartic and cysteine proteinase families the most abundant. This may reflect its migratory lifestyle and the need to degrade a range of host peptides, as well as its feeding behaviour. Some of these proteins have been identified as being secreted from the gland cells as described in Chapter V, specifically catalytic families A1 (aspartic) and C1A (cysteine). There is also another function of the peptidases that are highly expressed in the intestine of the nematode during infection. A secreted aspartic peptidase of *M.*

incognita is found in the plant roots during the migration phase but also in the sedentary phase, near the anterior part of the nematode and the apoplasm of the giant cells, suggesting a role in digestion of host peptides (Vieira *et al.*, 2011).

The cell wall degrading enzymes are highly up-regulated *in planta* and are highly abundant in the gland cells (reviewed in Chapter I). As expected, these enzymes are highly expressed at the early stages of infection, and include cellulases, expansins and pectate lyases needed for disruption and modification of the plant cell wall. In the parasitic stage, the expression of fungal cell wall degrading enzymes, including endo-1,3- glucanase, are also upregulated which may reflect a switch to the fungal feeder mode.

Protection of the nematode

Many detoxification and anti-oxidant proteins are secreted *in planta*, and are expressed in the gland cells and intestine. Many of these proteins are highly expressed in parasitic stages and some are expressed in the gland cells, and belong to phase I and II of the metabolic pathways of xenobiotic compounds (more detailed in Chapter IV). In this study, we demonstrate that the migratory nematode uses a range of several detoxification enzymes to modify the endogenous compounds produced by the host during infection. Therefore, the role of the metabolism pathways in the nematode defence against the host are likely to be of major importance for survival. We have shown that a gland cell secreted-GST can degrade of some of the terpenoid compounds and reactive oxygen species released by the host when infected by the nematode. A previous study showed that a *B. xylophilus* catalase is an important detoxification enzyme in tolerance to oxidative stress, however these proteins were secreted in the intestine (Vicente *et al.*, 2015). We identified other detoxification enzymes such as UDP-glucuronosyl transferase (UGT), cytochrome P33 and multicopper oxidase (Chapter II). This extensive repertoire may reflect the adaptation of the nematode to cope with the range of different environments that the nematode experiences during its life cycle, such as the insect body during transmission, pupal chambers and wilting tree.

There are several proteins highly expressed in early stage of infection whose functional role is unknown. During this work, several proteins were identified with a toxin protein domain (metridin-like ShK toxin) but the function of such proteins in PPN is unknown. Similar proteins were previously reported in the filarial nematode *Litomosoides sigmodontis* (Armstrong *et al.*, 2014) and in the animal parasitic nematode, *Trichostrongylus colubriformis* (Strongylida) (Cantacessi *et al.*, 2010) and have been put forward as a vaccine targets for human filarial diseases. In *B. xylophilus* these proteins were mostly expressed in the gland cells, and several have a signal peptide. Genes encoding thaumatins are abundantly present and secreted in the gland cells of the nematode, but are not differentially expressed during infection. These have been also reported in the secretome of *B. xylophilus* (Shinya *et al.*, 2013), but their functional role is unknown.

A substantial proportion of the candidate effectors identified were pioneers. The presence of novel genes as effectors is common for PPNs and is considered as a evolutionary adaptation of the pathogens to their different environments and conditions (Kikuchi et al., 2017). In *B. xylophilus*, many of these pioneer effector sequences are present in the gland cells and highly expressed during infection of the tree host, suggesting a role in parasitism. Some of these may be host or species-specific. Next generation sequencing of new PPNs has allowed the identification of these novel effector genes, however the determination of the functional role of these genes remains challenging, particularly in a tree pathogen.

All together, these effectors contribute to successful migration, development and reproduction inside the tree. Parasitism mechanisms should not be explained in terms of one or a small number of effectors, but by complex sets of different proteins that work together to allow the nematode to infect. However, if a single effector or regulator can be identified that is essential for parasitism this will represent an attractive control target. New functional characterization of more effectors of the nematode will give new insights into the development of PWD. Additionally, the interaction with the host must be assessed regarding the plant response to nematode parasitism. Few previous studies have sequenced pine transcripts after infection of the PWN (Santos *et al*., 2012; Gaspar *et al*., 2017), providing new data for complementary studies.

Plant-parasitic nematodes are challenging and amazing organisms to study. The availability of new sequencing technologies has revolutionized the field and has provided significant opportunities to improve our understanding of the biology of these pathogens. It may also be beneficial to integrate such studies with those that use other developing technologies such as proteomics. This could give a more detailed panel of the functional effectorome during parasitism.

.

References

Armstrong, S.D., Babayan, S.A., Lhermitte-Vallarino, N., Gray, N., Xia, D., Martin, C., Kumar, S., Taylor, D.W., Blaxter, M.L., Wastling, J.M. and Makepeace, B.L. (2014) Comparative analysis of the secretome from a model filarial nematode (*Litomosoides sigmodontis*) reveals maximal diversity in gravid female parasites. *Mol. Cell Proteomics* 13(10), 2527-44.

Boyd, I.L., Freer-Smith, P.H., Gilligan, C.A. and Godfray, H.C.J. (2013) The consequence of tree pests and diseases for ecosystem services. *Science* 342, 1235773.

Cantacessi, C., Mitreva, M., Campbell, B.E., Hall, R.S., Young, N.D. Jex, A.R. , Ranganathan, S. and Gasser, R.B.*(2010) First transcriptomic analysis of the economically important parasitic nematode, *Trichostrongylus colubriformis*, using a next-generation sequencing approach. *Infect. Genet. Evol.* 10(8), 1199–1207.

Eves-van den Akker, S., Laetsch, D. R., Thorpe, P., Lilley, C.J., Danchin, E.G.J., da Rocha, M., Rancurel, C., Holroyd, N.E., Cotton, J.A., Szitenberg, A., Grenier, E., Montarry, J., Mimee, B., Duceppe, M., Bo+, I., Marvin, J.M.C., Jones, L.M., Yusup, H.B., Lafond-Lapalme, J., Esquibet, M., Sabeh,M., Rott, M., Overmars, H., Finkers-Tomczak, A., Smant, G., Koutsovoulos, G., Blok, V., Mantelin, S., Cock, P.J.A., Phillips, W., Henrissat, B., Urwin, P.E., Blaxter, M. and 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.

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

Gaspar, D., Trindade, C., Usié, A., Meireles, B., Barbosa, P., Fortes, A.M., Pesquita, C., Costa, R.L. and Ramos, A.M. (2017) Expression profiling in pinus pinaster in response to infection with the pinewood nematode, *Bursaphelenchus xylophilus*. *Forests* 8, 279.

Harwood, T.D., Tomlinson, I., Potter, C.A. and Knight, J.D. (2011) Dutch elm disease revisited: past, present and future management in Great Britain. *Plant Pathol*. 60, 545-555.

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

Kikuchi, T., Cotton, J.A., Dalzell, J.J., Hasegawa, K., Kanzaki, N., McVeigh, P., Takanashi, T., Tsai, I.J., Assefa, S.A., Cock, P.J.A., Otto, T.D., Hunt, M., Reid, A.J., Sanchez-Flores, A., Tsuchihara, K., Yokoi, T., Larsson, M.C., Miwa, J., Maule, A.G., Sahashi, N., Jones, J.T. and Berriman, M. (2011) Genomic insights into the origin of parasitism in the emerging plant pathogen *Bursaphelenchus xylophilus*. *PLoS Pathog.* 7, e1002219.

Kikuchi, T., Eves-van den Akker, S., Jones, J.T. (2017) Genome evolution of plant-parasitic nematodes. *Annu. Rev. Phytopathol*. 55, 333-354.

Mei, Y., Thorpe, P., Guzha, A., Haegeman, A., Blok, V.C., MacKenzie, K., Gheysen, G., Jones, J.T and Mantelin, S. (2015) Only a small subset of the SPRY domain gene family in *Globodera pallida* is likely to encode effectors, two of which suppress host defences induced by the potato resistance gene *Gpa2*. *Nematology* 17, 409-424.

Petre, B., Joly, D.L. and Duplessis, S. (2014) Effector proteins of rust fungi. *Front. Plant Sci.* 5, 416.

Rutter, W.B., Hewezi, T., Abubucker, S., Maier, T.R., Huang, G., Mitreva, M., Hussey, R.S. and Baum, T.J. (2014) Mining novel effector proteins from the esophageal gland cells of *Meloidogyne incognita*. *Mol. Plant Microbe Interact*. 27(9), 965-974.

Santos, C.S, Pinheiro, M., Silva, A.I., Egas, C. and Vasconcelos, M.W. (2012) Searching for resistance genes in Bursaphelenchus xylophilus using high throughput screening. *BMC Genomics* 13, 599.

Shinya, R., Morisaka, H., Kikuchi, T., Takeuchi, Y., Ueda, M. and 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, e67377.

Sonah, H., Deshmukh, R.K. and Bélanger, R.R. (2016) Computational prediction of effectors proteins in fungi: opportunities and challenges. *Front Plant Sci.* 7, 126.

Sun, J., Gao, Z., Zhang, X., Zou, X., Cao, L. and Wang, J. (2017) Transcriptome analysis of *Phytophthora litchii* reveals pathogenicity arsenals and confirms taxonomic status. *PLoS ONE* 12(6): e0178245.

Thorpe, P., Mantelin, S., Cock, P.J.A., Blok, V.C., Coke, M.C., Cotton, J.A., Eves van den Akker, S., Guzeeva, E., Lilley, C.J., Reid, A.J., Wright, K.M., Urwin, P.E and Jones, J.T. (2014) Characterisation of the full effector complement of the potato cyst nematode *Globodera pallida*. *BMC Genomics* 15, 923.

Vicente, C.S.L., Ikuyo, Y., Shinya, R., Mota, M. and Hasegawa, K. (2015) Catalases Induction in High Virulence Pinewood Nematode *Bursaphelenchus xylophilus* under Hydrogen Peroxide-Induced Stress. *PLoS ONE* 10(4), e0123839.

Vieira, P., Danchin, E.G.J., Neveu, C., Crozat, C., Jaubert, S., Hussey, R.S., Engler, G., Abad, P., Almeida-Engler, J., Castagnone-Sereno, P. and Rosso, M. (2011) The plant apoplasm is an important recipient compartment for nematode secreted proteins. *J. Exp. Bot*. 62, 1241–1253.

UNIVERSIDADE DE ÉVORA
INSTITUTO DE INVESTIGAÇÃO
E FORMAÇÃO AVANÇADA

Contactos:

Universidade de Évora

Instituto de Investigação e Formação Avançada - IIFA

Palácio do Vimioso | Largo Marquês de Marialva, Apart. 94

7002-554 Évora | Portugal

Tel: (+351) 266 706 581