Identification and characterization of parasitism genes from the pinewood nematode Bursaphelenchus xylophilus reveals a multi-layered detoxification strategy
Espada, Margarida; Silva, Ana Cláudia; Eves-van den Akker, Sebastian; Cock, Peter J. A.; Mota, Manuel; Jones, John T.

Published in:
Molecular Plant Pathology

DOI:
10.1111/mpp.12280

Publication date:
2016

Document Version
Peer reviewed version

Link to publication in Discovery Research Portal

Citation for published version (APA):
Identification and characterization of parasitism genes from the pine wood nematode Bursaphelenchus xylophilus reveals a multilayered detoxification strategy.

Running title: Effectors of B. xylophilus.

Margarida Espada¹,², Ana Cláudia Silva¹,²,³, Sebastian Eves-van den Akker¹, Peter J.A. Cock², Manuel Mota¹ & John T. Jones²,³,⁵

²Cell and Molecular Sciences Group/ Information and Computer Sciences group (PJAC), The James Hutton Institute, Invergowrie, Dundee, DD2 5DA, UK
³Biology Department, University of Gent, KL Ledeganckstraat 35, B9000 Gent, Belgium
⁴College of Life Sciences, University of Dundee, Dundee, DD1 5EH
⁵School of Biology, University of St Andrews, North Haugh, St Andrews, KY16 9TZ

Corresponding authors: John.Jones@hutton.ac.uk; mmota@uevora.pt

Keywords: Bursaphelenchus xylophilus, transcriptome, effectors, xenobiotic metabolism, gland cells.

Word count: 4713
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 pine wood 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 compounds (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 these interactions are mediated by 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).

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 (Supplementary Figure 1). 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), GHF45 cellulases, peptidases and GH16 endoglucanases (Supplementary Table 1).

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.

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 (Supplementary Figure 3). The RT-PCR showed confirmed that all the ten putative effector genes were more highly 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, compared to nematodes feeding on fungi (Supplementary Figure 3). Most of the genes showed highest expression levels at 6 dpi. However, one gene (BUX.s00422.201) was equally highly expressed at both 6 and 15-dpi while three of the genes (BUX.01144.122; BUX.s01147.177 and BUX.s01066.2) showed highest expression at 15 dpi.

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.
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*, BxPr75OH, 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 25°C (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 60°C) 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 40°C. Nematodes were then transferred to a solution of 75% glycerine: 25% lactophenol for approximately 24 hours at 40°C, 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) (Supplementary Table 2). 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 tool 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 Supplementary Table 3.

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 (Supplementary Table 3). 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.

Acknowledgements

This work was supported by the REPHRAME project (KBBE.2010.1.4-09). ME is funded by FCT (Fundação para a Ciência e a Tecnologia, IP) under a PhD grant (SFRH/BD/84541/2012). ME and MM are also funded by FEDER Funds through the Operational Programme for Competitiveness Factors - COMPETE and National
Funds through FCT - Foundation for Science and Technology under the Strategic Projects PEst-C/AGR/UI0115/2011 and PEst-OE/AGR/UI0115/2014. AS was funded by an ERASMUS MUNDUS Category B scholarship awarded through project 2008-2102 (EUMAINE). The James Hutton Institute receives funding from the Scottish Government Rural and Environmental Science and Analytical Services division (RESAS).

References


Mann, H. B. and Whitney, D. R. (1947) On a test of whether one of two random


RNAseq data described in this manuscript are available through ENA under accession numbers X-Y.
Supporting information legends

Supplementary Figure 1 - Differential expression analysis of the transcripts. The heatmap resulting from the RNAseq analysis, using 8 samples in three different conditions – pre-invasive/mycetophagous (Fungal Feeding) and post invasive/Phytophagous (6 and 15 DPI).

Supplementary Figure 2 - Analysis of the most represented molecular function (level 3) in the Top 200 set of up regulated genes obtained by a bioinformatics pipeline.

Supplementary Figure 3 - Validation of the expression of the secreted effectors by semi-quantitative-PCR using the actin housekeeping 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 (FPKM) predicted by RNAseqTMM for each candidate gene.

Supplementary Table 1 – List of the twenty-nine differential expressed transcripts between mycetophagous and phytophagous stages. Detailed description of the twenty-nine transcripts includes the presence or absence of putative signal peptide, their putative protein domain (according to Gene DB annotation of the 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 nr (non redundant) database (cutoff value of 1e-03) and their orthologous protein sequences in other organisms,
and also the normalised expression profile in the three different conditions (fungal feeding nematodes [FF], nematodes 6 and 15 days post infection [dpi]). The normalized expression values are in FPKM (Fragments Per Kilobase of exon per Million mapped fragments).

**Supplementary Table 2** - Summary of RNAseq data.

**Supplementary Table 3** - List of pair of primers used for amplification of probes for *in situ* hybridisation. Gene model are according to Kikuchi *et al.*, 2011 and sequences available at http://www.genedb.org/Homepage/Bxylophilus. bp: base pair.
### 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.
<table>
<thead>
<tr>
<th>Predicted function</th>
<th>Putative protein domain (GeneDB annotation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROTEASES (10)</td>
<td>Aspartic protease A1 (5)</td>
</tr>
<tr>
<td></td>
<td>Cysteine proteases C1A (1); C46 (1)</td>
</tr>
<tr>
<td></td>
<td>Serine-type protease (2)</td>
</tr>
<tr>
<td></td>
<td>Metallo-type protéase M13 (1)</td>
</tr>
<tr>
<td>FATTY ACID METABOLISM (2)</td>
<td>Fatty acid retinoid binding proteins</td>
</tr>
<tr>
<td>DETOXIFICATION OF XENOBIOTIC COMPOUNDS (12)</td>
<td>FMO (flavin monoxygenase) (2)</td>
</tr>
<tr>
<td></td>
<td>UDP-glucuronosyl transferase (2)</td>
</tr>
<tr>
<td></td>
<td>Multicopper putative acid oxidase (1)</td>
</tr>
<tr>
<td></td>
<td>Glutathione S-transferase (2)</td>
</tr>
<tr>
<td></td>
<td>Cytochrome P450 (3)</td>
</tr>
<tr>
<td></td>
<td>Acid phosphatase (1)</td>
</tr>
<tr>
<td></td>
<td>Epoxide hydrolase (1)</td>
</tr>
<tr>
<td>UNKNOWN PROTEINS DOMAIN (PIONEERS) (16)</td>
<td>None</td>
</tr>
<tr>
<td>PROTEIN WITH TOXIN DOMAIN (2)</td>
<td>Metridin-like Shit toxin domain</td>
</tr>
<tr>
<td>ALLERGENS (1)</td>
<td>Putative allergen V5/TPx1</td>
</tr>
<tr>
<td>GLYCOSYL HYDROLASE CLASSES (2)</td>
<td>GH29 (alpha-L-fuco domain)</td>
</tr>
<tr>
<td></td>
<td>GH30- GH2</td>
</tr>
<tr>
<td>LYSOZYME ACTIVITY (1)</td>
<td>Lysozyme 7,8</td>
</tr>
</tbody>
</table>

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

**Figure 1** - Positions of pharyngeal gland cells in *juvenile-adult* (A) and *adult-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)

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

**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.