Characterisation of arabinogalactan endo β 1,4 galactanases from Globodera rostochiensis, Globodera pallida and Rotylenchulus reniformis

Kerry Leslie, Sophie Mantelin, Sebastian Eves-van den Akker, John T Jones

HAL Id: hal-04230005

https://hal.inrae.fr/hal-04230005

Submitted on 5 Oct 2023

Characterisation of arabinogalactan endo β 1,4 galactanases from *Globodera rostochiensis, Globodera pallida* and *Rotylenchulus reniformis*

Kerry Leslie [a,b], Sophie Mantelin [c], Sebastian Eves-van den Akker [d], John T. Jones [a,b,*]

[a] *School of Biology, Biomedical Sciences Research Complex, University of St Andrews, North Haugh, St Andrews, KY16 9ST, UK*
[b] *Cell & Molecular Sciences Department, The James Hutton Institute, Invergowrie, Dundee, DD2 5DA, UK*
[c] *INRAE Institute Sophia Agrobiotech, 400 route des Chappes, BP 167, 06903, Sophia Antipolis, France*
[d] *The Crop Science Centre, Department of Plant Sciences, University of Cambridge, CB2 3EA, UK*

A B S T R A C T

Plant parasitic nematodes need to overcome the barrier presented by the plant cell wall in order to invade their host. A variety of plant cell wall degrading enzymes are present in endoparasitic nematodes including enzymes that degrade cellulose (beta 1,4 endoglucanases) and various pectin components. We describe the cloning and functional analysis of genes encoding GH53 arabinogalactan endo-1,4-beta-galactosidases from three related plant parasitic nematodes *Globodera rostochiensis, Globodera pallida* and *Rotylenchulus reniformis*. Phylogenetic and structural analyses strongly indicate that these genes have been acquired by horizontal gene transfer from bacteria. We show that the genes are expressed at invasive stages of the parasites in the secretory gland cells. We also demonstrate that the enzymes from these species are biochemically active, showing the expected hydrolytic enzymatic activity when galactan was used as a substrate. This work further demonstrates the importance of cell wall degradation to the success of the parasitic process and the extensive role that horizontal gene transfer has played in the evolution of plant parasitism by nematodes.

1. Introduction

The plant cell wall is made up of an interconnected network of proteins and polysaccharides with the most abundant of these being cellulose, hemi-cellulose, and pectin (*e.g.* Ref. [1]. This network provides a strong physical barrier due to the properties of the components, including covalent bonding between polysaccharides [2]. Cellulose is the most abundant polysaccharide and consists of large, unbranched strands of β-1,4-linked glucose molecules. Hemicelluloses are also abundant in the cell wall with the most common being xylan, arabinoxylan, and mannan. The hemicelluloses are intertwined with the cellulose fibres via hydrogen bonds and together these are enmeshed within layers of pectin. The composition of plant cell walls may vary among species and tissues, each cell type possessing distinct and dynamic cell wall compositions and organisation [3]. For example, the cell walls of monocots contain relatively low levels of pectic polysaccharides [4].

All plant pathogens, including plant parasitic nematodes, need to break down the plant cell wall in order to gain access to their hosts. The first stage of the life cycle of cyst nematodes requires them to invade host roots and migrate intracellularly (destructively) through the host tissues until they reach the inner cortex layer [5]. The nematode uses a combination of physical disruption of the plant cell wall using the stylet and biochemical degradation of the cell wall components during this migration. Since the initial discovery of genes encoding cellulases (β 1,4 endoglucanases) in cyst nematodes (Smant et al., 1998), a wide range of cell wall degrading and modifying enzymes have been identified from plant-parasitic nematodes [6]. The first plant parasitic nematode genome-wide analysis, of the root-knot nematode *Meloidogyne incognita,* has revealed that over 60 genes encoding cell wall degrading or modifying proteins from six different families are present, including cellulases, xylanases, polygalacturonases, pectate lyases and arabinases [7]. Likewise, analysis of potato cyst nematode genomes has revealed a similar repertoire of enzymes [8,9]. Phylogenetic analysis of the sequences from a range of root-knot and cyst nematodes suggests that these sequences have been obtained by horizontal gene transfer from a variety of bacterial species [10]. In addition, *Bursaphelenchus* spp. contain GHF45 cellulases that were very likely to have been acquired by horizontal gene transfer from fungi [11], while *Xiphinema index* contains GHF12 cellulases most likely acquired from bacteria [12]. This suggests

that independent horizontal gene transfer events have occurred in each of the clades of nematodes that can parasitise plants. The complement of cell wall degrading enzymes present in plant-parasitic nematodes may vary to reflect the cell wall composition of the host [13] and expression levels of cell wall degrading enzymes can change in response to perceived cues relating to the composition of the plant cell wall in different hosts [14].

Pectin is a key component of the plant cell wall and is composed of three main types of carbohydrates: homogalacturonan (HG), and rhamnogalacturonan I and II (RGI and RGII) with HG being the most abundant. Pectin can be categorised into smooth or hairy regions based on the level of branching sidechains that are present. HG comprises chains of 1,4-linked α-D-galacturonic acid and as there are no branching chains extending from the HG backbone, it is classified as smooth [15]. RGI contains repeating units of L-rhamnose and D-galacturonic acid in its backbone. Branching chains such as galactan, arabinan, and arabino-galactan are attached to the L-rhamnose molecules of the RGI backbone [16]. RGII has a backbone of galacturonic acid units (monosaccharides), like HG, but also has branching side chains like RGI. These side chains are often complex and made up of multiple different types of carbohydrates including apiose, fucose, and aceric acid (3-C-carboxy-5-deoxy-γ-L-xylofuranose) [17]. The presence of branching side chains on RGI and RGII means that these are referred to as hairy regions of pectin.

The breakdown of pectin requires the concerted action of several different degrading enzymes and many of these have been identified in plant-parasitic nematodes. Pectate lyases cleave the glycosidic bond between α-1,4-polygalacturonic acid units in homogalacturonan through a β-elimination reaction [18]. These enzymes were first described in plant-parasitic nematodes from the potato cyst nematode *Globodera rostochiensis* [19] and have subsequently been identified in many other species including root-knot nematodes and migratory endoparasites (reviewed by [20]). Polygalacturonases, which act on pectate and on other galacturonans by hydrolysing 1,4-alpha-D-galactosiduronic linkages, have been identified in root-knot nematodes [21] as well as in a transcriptome dataset for the closely related migratory endoparasitic *Pratylenchus coffeae* [22], although no functional analysis of this predicted protein has been reported. In terms of metabolising pectin side chains, arabinases have been identified in genome sequences for root-knot nematodes [7,23]. In addition, a candidate GH53 arabinogalactan endo-1,4-beta- galactanase was identified in the genome sequences for *G. rostochiensis* and *Globodera pallida* [8,9]. A similar cDNA sequence was previously identified in the transcriptome of the beet cyst nematode *Heterodera schachtii* [24]. This study reported the cloning of a full-length cDNA that could encode a GH53 enzyme and showed that it was expressed at the J2 stage and localised to the subventral gland cells. However, no functional characterisation of any nematode GH53 protein has been conducted to date.

Here we report cloning and functional analysis of GH53 encoding genes from three related plant parasitic nematodes *G. rostochiensis, G. pallida* and *Rotylenchulus reniformis*. We examine spatial and temporal expression patterns and demonstrate that the enzymes from these species are biochemically active.

2. Materials and methods

2.1. Biological material

Populations of *G. pallida* and *G. rostochiensis* were grown on the susceptible potato (*Solanum tuberosum*) cultivar Désirée in a glasshouse. Cysts were extracted using standard protocols and stored at 4 °C for at least 6 months before use. Second stage juveniles (J2) were hatched in tomato root diffusate prepared as previously described [25]. Fixed samples of *Rotylenchulus reniformis* were provided by Dr Catherine Lilley, University of Leeds and *Xanthomonas campestris* pv. *Campestris* (Xcc) wild-type strain 8004 was obtained from Dr. John Maxwell DOW, BIOMERIT Research Centre, Department of Microbiology, University

College, Cork, Ireland.

2.2. Cloning

Genes encoding sequences similar to GH53 Arabinogalactan endo-β-1,4-galactanases were identified from transcriptomic and genomic studies for *G. pallida* [26], *G. rostochiensis* [9] and *R. reniformis* [27]. The coding regions of the putative GH53 genes from the three nematode species (GROS_g08150 (GrGAL1), GPLIN_000142900 (GpGAL1), RrGAL1) and the bacterial control *Xanthomonas campestris* (XC_0587 (GalA_Xc)) were amplified by PCR from cDNA (or gDNA for the *Xanthomonas* control) using the proof-reading KOD Hot Start DNA polymerase (Merck) using primer sets shown in Supplementary Table S1. Messenger RNAs were isolated from nematode material using a Dynabeads mRNA Direct Micro kit (Invitrogen) and treated with RQ1 DNase (Promega). cDNA was synthesised from approximately 400 ng purified mRNA using the Superscript III system (Invitrogen) with poly(dT) primers following the manufacturer's instructions. The open reading frame of each of the genes was cloned from the start of the predicted mature peptide to the stop codon (excluding the endogenous signal peptide). PCR products were separated on 1.5% agarose gels, excised, and purified using a QIAquick gel extraction kit (Qiagen)before cloning into the Gateway-compatible TOPO entry vector pCR8/GW/TOPO (Invitrogen), following the manufacturer's instructions. The cloned genes were sequenced at The James Hutton Institute sequencing facility.

GrGAL1, GpGAL1, RrGAL1 and XC_0587 (GalA_Xc) were subsequently cloned into the protein expression vector pOPIN_S3C [28]. The pOPIN_S3C vector contains a 6x Histidine (His) tag, a SUMO domain, and a 3C protease cleavage site. The green fluorescent protein (GFP) gene used as negative control, that was previously cloned into the bacterial expression vector pKC026 from a TOPO donor clone [29], was also transferred into the pOPIN_S3C vector. In-fusion cloning into the pOPIN_S3C vector was carried out using the NEBuilder HiFi DNA assembly kit following the manufacturer's protocol (New England Biolabs). Primer sets for pOPIN_S3C cloning are shown in Supplementary Table S1.

2.3. Phylogenetic analysis and protein structure modelling

BLAST similarity searches were conducted using the tBLASTn function against the non-redundant nucleotide database with the *Globodera* and *Rotylenchulus* GH53s as query sequences [30]. The top 100 results from each of the BLAST searches were combined and filtered to remove duplicates and low confidence hits, resulting in a unique list of 78 sequences. This list was subsequently found to have a very large proportion of very similar sequences from *Xanthomonas* species. The list was therefore manually filtered to remove the majority of these, keeping only the highest percentage identity hits in cases where there were multiple sequences included from one species. Additional fungal sequences were subsequently added to the list, from *Aspergillus aculeatus, Aspergillus nidulans, Humicola insolens*, and *Thermothelomyces thermophilus* species to ensure a diverse phylogenetic representation. These were identified through BLAST searches of the Protein Data Bank (PDB) database. Pairwise alignments of all sequences were created using Muscle and visualised using Jalview [31]. Model selection (LG (general matrix) + F (empirical base frequency) + G4 (rate heterogeneity gamma model) model as determined by ModelFinder [32]) and phylogenetic tree inference were carried out using IQ-TREE [33] using default parameters. The phylogenetic tree derived from 1000 bootstrap replicates [34] was then annotated using FigTree v1.4.3.

Predicted structures of the GrGAL1, GpGAL1 and RrGAL1 proteins were produced using a 1-to-1 thread model based on the amino acid sequence of β-1,4-galactanase from *Bacteroides thetaiotaomicron*. The sequence and structure from *B. thetaiotaomicron* were identified using BLAST similarity searches with mask low complexity settings applied. Searches were completed using BLAST and the Research Collaboratory
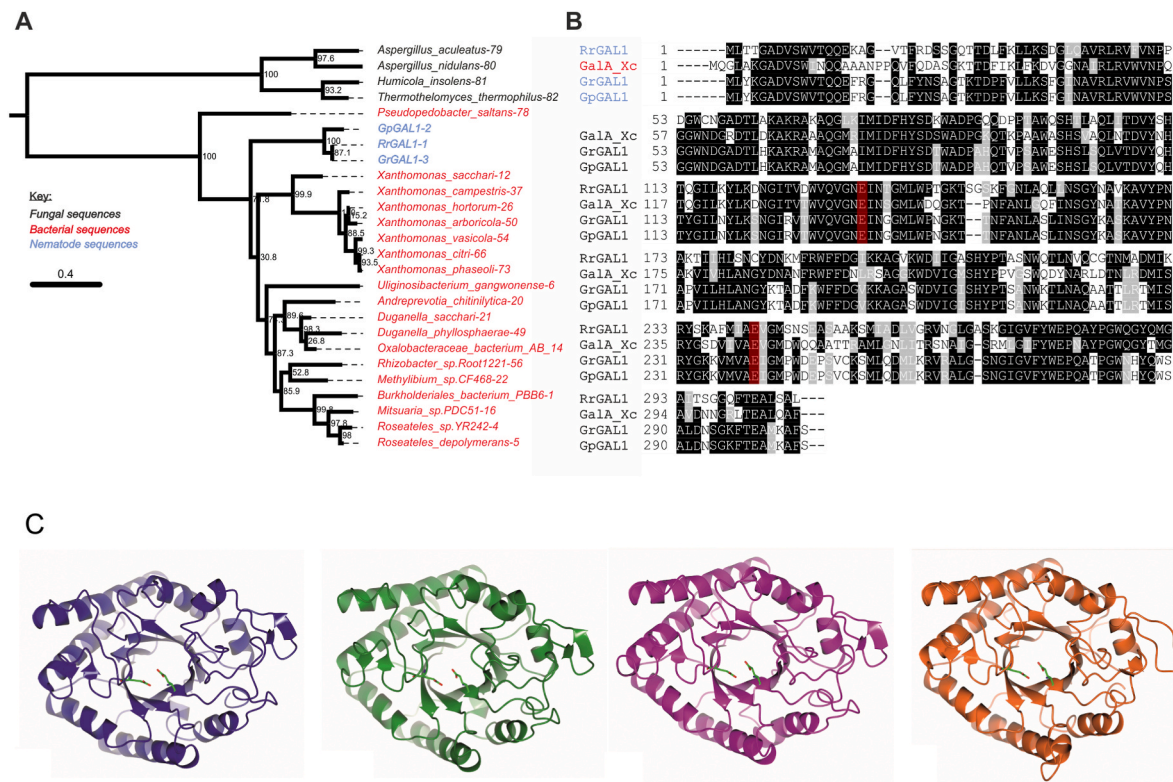
Fig. 1. Nematode GH53 sequences. A: Phylogenetic analysis of nematode, bacterial and fungal GH53 sequences. The tree is midpoint re-rooted and based on 1000 bootstraps. Each entry is followed by "-XX" which is the identification number attributed in Supplementary Table S2. Bar, 0.4%, sequence dissimilarity (evolutionary distance). **B:** Alignment of mature protein sequences GrGAL1 (*G. rostochiensis*), GpGAL1 (*G. pallida*), RrGAL1 (*R. reniformis*) and GalA_Xc (*X. campestris*). Catalytic glutamate residue (E) are highlighted in red. **C:** Predicted structures of nematode GH53 proteins using a 1-to-1 threaded model. Predicted structures shown in the order (left to right) GrGAL1, GpGAL1, RrGAL1, confirmed structure of BTGH53 from *Bacteroides thetaiotaomicron* (6gp5_A). Catalytic glutamates in the active site are present in all four structures, in very similar positions (highlighted green/red, central).

for Structural Bioinformatics Protein Data Bank (RCSB PDB) [35] (www. rcsb.org). The 1-to-1 thread model was achieved using Protein Homology/analogy Recognition Engine V 2.0 (PHYRE2) [36]. Rendered images of predicted protein structures were produced using CCP4 molecular graphics (CCP4mg) (V2.10.10) [37].

2.4. Gene expression profiles

In situ hybridisation was used to determine the spatial expression patterns of the *GrGAL1, GpGAL1* and *RrGAL1* sequences in J2s of *G. rostochiensis, G. pallida* and *R. reinforms,* respectively, using the protocol described in Jones et al. [38]. Normalised gene expression data for various life stages of *G. rostochiensis* [9], *G. pallida* [8] and *R. reniformis* [27] were used to determine temporal expression patterns of each of the genes.

2.5. Protein expression and purification

Ten millilitres of LB media were inoculated with a single colony from GrGAL1, GpGAL1, RrGAL1, GalA_Xc, or the GFP control in Shuffle *Escherichia coli* cells (New England Biolabs). Cultures were grown overnight at 37 °C with shaking. One hundred microlitres of the overnight cultures were added to 100 mL of fresh LB media which were incubated at 30 °C with shaking until an OD_{600} of 0.7 was reached. Cultures were cooled to 18 °C. A sample of each culture was taken and stored as a pre-induction control. Expression of protein in the remaining cultures was induced by the addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM, before incubation overnight at 18 °C with shaking. Cultures were centrifuged at 4000 *g* for 10 min to pellet the remaining cells. The supernatant was

discarded and the pellet was resuspended in 1 mL of ice-cold lysis and wash buffer (50 mM Tris-HCL, 500 mM NaCl, 50 mM glycine, 5% glycerol, 20 mM imidazole, EDTA-free protease inhibitor tablets (Complete Mini; Roche Diagnostics), pH 8.0). Cells were sonicated for 30 s followed by 30 s cooling on ice (6 times) and the lysate was centrifuged at 13,000 *g* for 2 min. The supernatant was transferred to a new Eppendorf tube with 100 μl of Ni-NTA Superflow resin (Qiagen) and incubated for 1 h at room temperature on a rotator. The sample was centrifuged twice at 13,000 *g* for 1 min and the supernatant was removed. Beads were washed in 1 ml lysis and wash buffer, centrifuged at 13000 g × 1min before discarding the supernatant. Elution buffer (2.5 mL; 50 mM Tris-HCL, 500 mM NaCl, 50 mM glycine, 5% glycerol, 500 mM imidazole, EDTA-free protease inhibitor tablets, pH 8.0) was added to the beads in a 15 ml Falcon tube, which were then incubated at room temperature for 10 min and then centrifuged twice at 13,000 *g* for 1 min. The supernatant, containing the purified protein of interest, was stored at −20 °C. Presence and purity of proteins was assessed by Coomassie blue staining following separation on NuPAGE precast gels as previously described [39] (Supplementary Fig. 1).

2.6. Enzyme activity assays

Twenty-five microlitres of each protein solution (GrGAL1, GpGAL1, RrGAL1 GalA_Xc or the GFP control at a concentration of 500 μg/mL), was incubated with 25 μL of 0.1% (w/v) galactan substrate solution in 50 mM sodium acetate (pH 5) at room temperature for 1 h. A blank sample was also set up containing 25 μL sterile distilled water and 25 μL of galactan solution. Fifty microlitres of DNS reagent (60 mM 3,5-dinitrosalicylic acid, 500 mM NaOH, 150 mM potassium sodium tartrate tetrahydrate) was added to each sample. Samples were boiled for 15 min
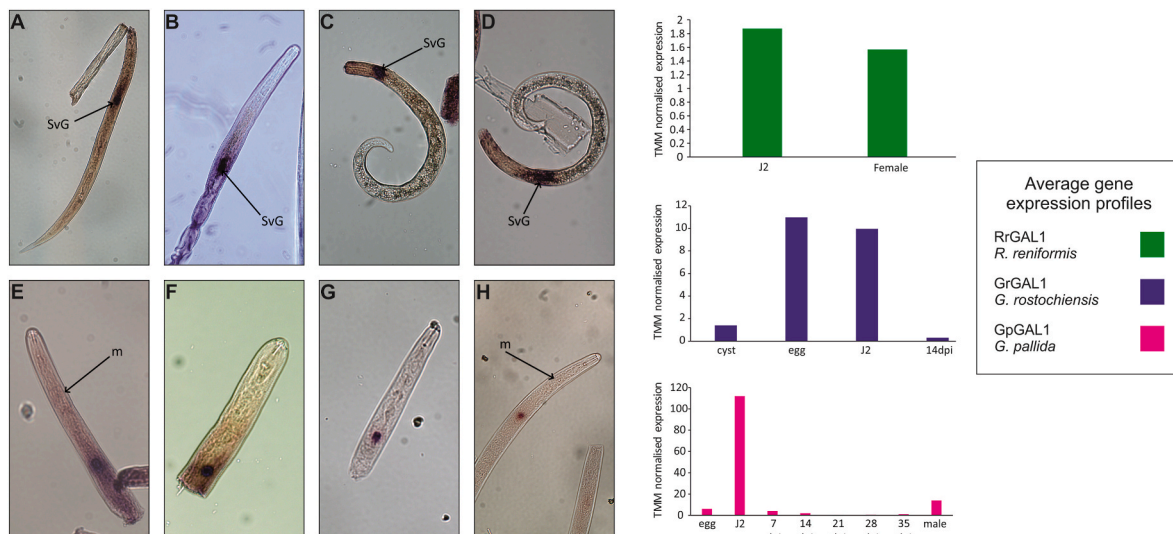
Fig. 2. *In situ* **hybridisation and temporal expression profiles of nematode GH53 mRNAs. A–B:** *GrGAL1* showing subventral gland localisation; **C–D:** *RrGAL1* showing subventral gland localisation, **E – H:** *GpGAL1* showing localisation in a circular structure posterior to the m16etacorpal bulb; **I:** Average gene expression profiles for GH53 transcripts across the life cycle. SvG – subventral gland, m -m16etacorpal bulb.

and cooled on ice for 3 min before adding 200 μL of water. Absorbance readings for each sample were taken at 540 nm using a Promega GloMax multi + plate reader.

The additional substrates xylan (from beechwood, Sigma), saccharose (VWR chemicals), pectin (from apple, Sigma), arabinogalactan (AG-II) (from Larchwood, Sigma) and polygalacturonic acid (Sigma) were tested using the same protocol as described above. All substrates were used at 0.1% (w/v) solutions in 50 mM sodium acetate pH 5.0. A standard curve was produced using 1 mg/mL (0.1% w/v) galactose (the reducing sugar produced on hydrolysis of galactan polymer). Seven samples were used containing 0.05, 0.2, 0.4, 0.6, 0.8, 1, and 2 mg galactose plus DNS reagent in a total volume of 300 μl. A blank sample containing water and DNS reagent only was also tested.

Data accessibility: All sequence data used in this research are available through previously deposited genome or transcriptome resources for *G. pallida* [8], *G. rostochiensis* [9] and *R. reniformis* [27].

3. Results

3.1. Identification of GH53 sequences

Analysis of the *G. rostochiensis* genome [9] allowed a gene (GROS_g08150) similar to GH53s from a variety of species to be identified. This was subsequently renamed *GrGAL1* and the coding sequence used as a query for BLAST similarity searches to identify putative GH53 proteins in related nematodes. The sequence identified in the *G. pallida* genome (GPLIN_000142900) was truncated. A search against the *G. pallida* transcriptome returned two incomplete but overlapping sequences: comp4850_c0_seq1 and comp4850_c0_seq4. These were computationally recapitulated, alongside the partial genomic fragment, and renamed *GpGAL1* after successful cloning of the full-length gene. Similarly, two partial sequences were identified in *R. reniformis* (transcripts comp30258_c0_seq1, comp30258_c1_seq1) that were used to generate a putative full-length sequence (*RrGAL1*) that was confirmed by cloning.

GrGAL1 and GpGAL1 shared high sequence percentage identity with a GH53 protein from *Duganella sacchari* (67.11%) (NCBI seqID: WP_072787792.1). *D. sacchari* is a Gram-negative, soil dwelling bacteria, strains of which have been isolated from the rhizosphere of sugar cane plants [40]. GrGAL1 also shared high percentage identity with GalA, a GH53 protein from *Xanthomonas campestris* (55.37%) (NCBI

seqID: WP_011038708.1). *X. campestris* is a Gram-negative bacterial pathogen of Solanaceous plant species such as tomato and peppers (Potnis et al. 2015). RrGAL1 had the highest percentage identity with an arabinogalactan endo-1,4-β-galactosidase from the strain YR242 of a *Roseateles* sp. Bacteria (65.68%) (NCBI seqID: WP_092947600.1). As GH53s are not usually present in animals, a phylogenetic analysis of the nematode sequences and GH53 proteins from bacteria and fungi was undertaken in order to explore the likely origin of these genes (Fig. 1A). This analysis showed that the nematode sequences clustered with the bacterial sequences, while the fungal sequences formed an outgroup. This, coupled with a previous analysis [9], which showed that the *G. rostochiensis* sequence has an extremely high Alien Index score, suggests that like other genes encoding plant cell wall degrading enzymes in plant-parasitic nematodes, the GH53 sequences are present in the nematode genomes as a result of horizontal gene transfer from bacteria. Alignment of the nematode sequences against bacterial GH53 mature protein sequences showed that the key catalytic glutamate (E) residues are conserved across PPN and bacterial species (Fig. 1B).

The predicted protein sequences of GrGAL1, GpGAL1 and RrGAL1 were used for searches against the RCSB PDB to identify similar proteins for which a crystal structure is available. For all three sequences, the highest identity hit (40%ID for *Globodera* sequences and 39% for the *Rotylenchulus* sequence) was with the β-1,4-galactanase BTGH53 from *Bacteroides thetaiotaomicron* (NCBI seqID: 6GP5_A) [41]. A 1-to-1 thread model for the nematode sequences was produced using the protein structure and sequence alignment with the solved structure from *B. thetaiotaomicron* using Protein Homology/analogy Recognition Engine V 2.0 (PHYRE2). The predicted structures suggest that all three nematode proteins follow the same folding pattern: (β/α)$_8$ barrel. Furthermore, the conserved catalytic glutamates of the GH53 proteins are in the centre of the β-barrel in each structural prediction which matches with the position in bacterial protein structures (Fig. 1C).

3.2. Expression profiles of the GH53 sequences

In situ hybridisation was conducted for the genes encoding GrGAL1, GpGAL1 and RrGLA1 to identify localisation of the mRNA transcripts at the J2 life stage of these three species. *GrGAL1* and RrGAL1 (Fig. 2A–D) were localised in the subventral gland cells. However, *GpGAL1* consistently produced a condensed spherical staining pattern in the region of the oesophageal glands (Fig. 2E–H). This structure is too small to be
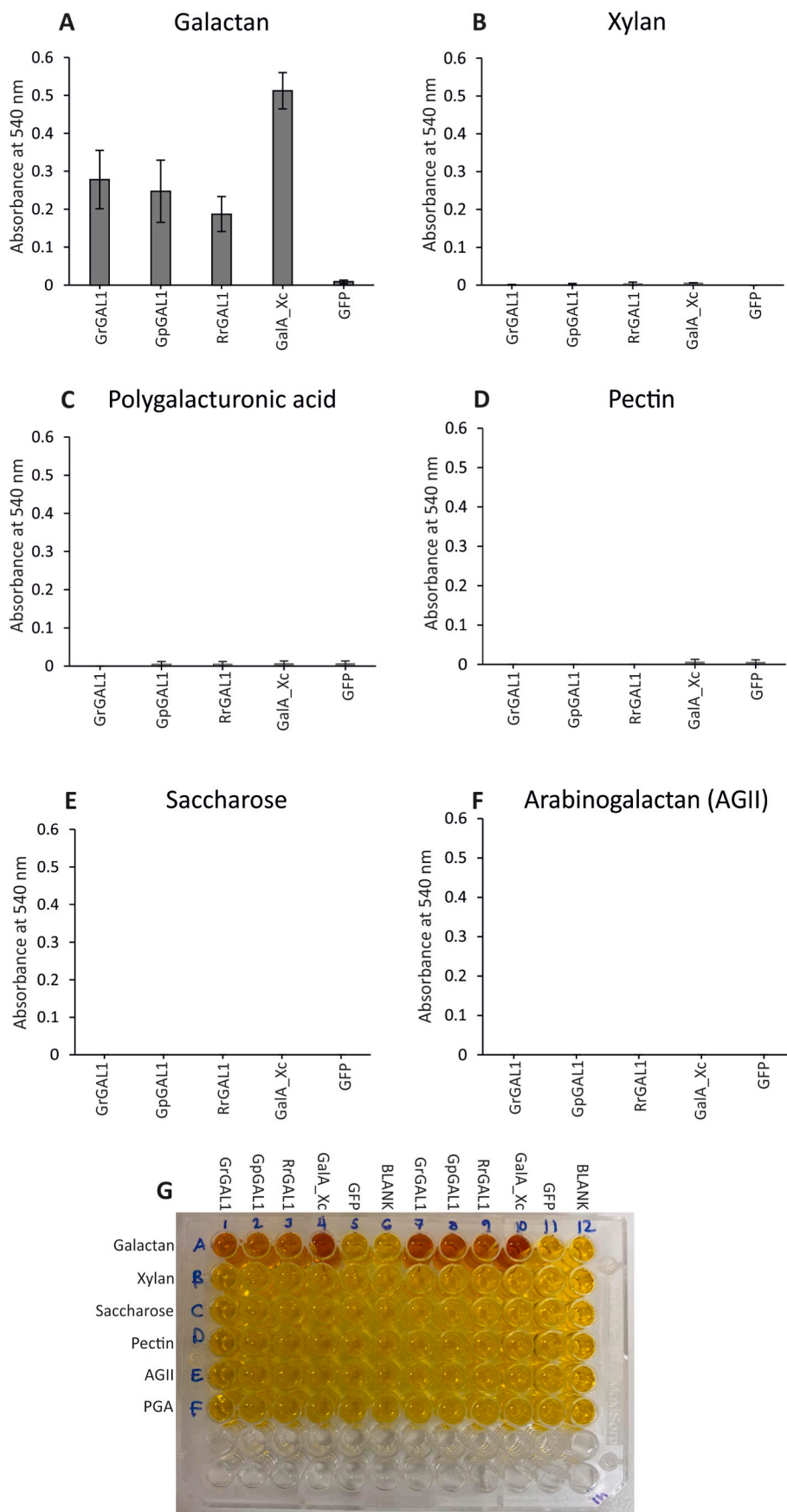
Fig. 3. DNS assay of recombinant nematode GH53 enzymes and positive (GalA_Xc) and negative (GFP) controls against plant cell wall polysaccharide component substrates. A: GrGAL1, GpGAL1, RrGAL1 and GalA_Xc produce detectable reducing sugars as monitored by absorbance at 540 nm while no sugars are released in the GFP sample. **B–H:** No reducing sugars are produced by any of the recombinant proteins using xylan (B), polygalacturonic acid (C), pectin (D), saccharose (E) or arabinogalactan type II (AGII) as a substrate. Bars represent the mean ± SE for $n = 3$. **G:** Image of 96-well plate assay depicting the colour change observed when GrGAL1, GpGAL1, RrGAL1, GalA_Xc, and GFP were tested using different substrates.

either of the pharyngeal gland cells, although it is possible that this structure corresponded to the nucleus of a gland cell. It is unclear why the transcript appears to localise only at the nucleus and not across the full gland cell. ISH negative controls using the appropriate sense primer probes displayed no specific signals with minimal background staining around the cut site (not shown). Analysis of RNAseq data showed that the expression of the GH53 gene was restricted to egg and J2 (*G. rostochiensis*) or J2 alone (*G. pallida*) while the *RrGAL1* gene was expressed in both J2 and female nematodes (Fig. 2 I).

3.3. Biochemical function of the nematode GH53 proteins

Recombinant proteins were produced from the three nematode GH53 sequences as well as the *X. campestris* bacterial GH53 protein (GalA_Xc), which was used as a positive control, and GFP, which was used as a negative control. A 3,5-dinitrosalicylic acid (DNS) assay was used to detect reducing sugars released due to the activity of the enzymes on various polysaccharide substrates. The presence of reducing sugars was indicated by a colour change (yellow to red/brown) measurable at 540 nm in a spectrophotometer. It was anticipated that the GH53 enzymes would hydrolyse the galactan substrate, releasing galactan oligosaccharides with reducing ends. The GH53s were also tested against other common polysaccharides found in the plant cell wall; xylan, saccharose, pectin, type II arabinogalactan (AG-II), and polygalacturonic acid (pectic acid). The DNS assay showed that recombinant GrGAL1, GpGAL1, and RrGAL1 all showed the anticipated hydrolytic enzymatic activity when galactan was used as the substrate at room temperature and pH 4 (Fig. 3A and G). No activity was detected using the GFP negative control while the positive control (GalA_Xc) showed enzymatic activity as expected (Fig. 3A and G). No enzymatic activity was detected when using xylan, polygalacturonic acid, pectin saccharose or arabinogalactan (AG-II) as substrate with any of the recombinant proteins, including the bacterial control (Fig. 3B–G), suggesting that the hydrolytic enzymatic activity of all the GH53 tested here maybe substrate specific.

4. Discussion

We describe the identification and functional characterisation of GH53 proteins from three related nematodes, *G rostochiensis, G. pallida* and *R. reniformis*. A similar protein was previously identified from *Heterodera schachtii* [42], although no confirmation of the biochemical activity of the predicted protein was included in this study. This therefore represents the first functional analysis of GH53 proteins from nematodes. The absence of similar genes in other (non plant-parasitic) nematodes, coupled with the phylogenetic and structural analysis presented here, strongly suggest that the gene encoding the GH53 enzymes was acquired from a bacterial donor by horizontal gene transfer. Analysis of the transcriptome of cyst nematodes that parasitise monocots [13] suggests that the GH53 gene has subsequently been lost in cyst nematodes that infect cereals, possibly as a result of the different composition of the cell wall in monocots as compared to dicots [43]. Expression analysis showed that spatial and temporal expression profiles of the GH53 sequences are consistent with a role in invasion of the host. For both species of PCN the J2 is the invasive stage and expression of the GH53 encoding genes is limited to these life stages. By contrast, expression of the *RrGAL-1* gene was detected at both life stages of *R. reniformis* sampled (juveniles and adult females). In this species, the adult female is the invasive stage while the juveniles are migratory [44]. Constitutive expression of the *RrGAL-1* gene is therefore consistent with this functional role, given that later juvenile stages will show changes in gene expression that reflect preparation for invasion of the host.

Biochemical analysis confirmed that the GH53 proteins specifically hydrolysed galactan, releasing shortened galactan oligosaccharides with reducing ends. None of the other substrates tested were hydrolysed, confirming the specificity of the nematode and bacterial enzymes.

Although this confirms that the nematode GH53s are functional enzymes it should be noted that the most likely natural substrate for these enzymes (type 1 arabinogalactan) is not currently available to purchase. A more detailed functional analysis of the biochemical properties was not, therefore, possible. However, we were able to show that the enzymes were active at between pH4 and pH5 in the assay used here, a similar pH to that in the plant apoplast (*e.g.* Grignon & Sentenac, 1991) where the enzymes are likely to be secreted and function *in vivo*. Plant-parasitic nematodes have been shown to produce a very wide range of plant cell wall degrading enzymes that, acting in concert, are able to degrade all main components of the plant cell wall. All of these enzymes identified to date have been acquired by horizontal gene transfer, mainly from bacteria [10,45]. The presence of at least one such cell wall degrading enzyme in all four clades of the Nematoda in which plant parasitism has evolved demonstrates the critical importance of the ability to degrade, or soften, the plant cell wall for plant-parasitic nematodes. Indeed, although such experiments were not performed here, RNA interference has previously been used to demonstrate that knock down of genes encoding cell wall degrading enzymes results in reduced infection (*e.g.* Ref. [46]; Hu et al., Vanholme et al., 2007; [47, 48]. This represents a further demonstration of the importance of cell wall degradation to the success of the parasitic process.

Author statement

Kerry Leslie: Conceptualization, Methodology, Investigation, Writing – Review and Editing. **Sophie Mantelin:** Conceptualization, Resources, Supervision, Writing – Review and Editing. **Sebastian Eves-van den Akker:** Conceptualization, Supervision, Writing – Review and Editing, Funding acquisition. **John T. Jones:** Conceptualization, Supervision, Investigation, Writing – Original Draft, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pmpp.2022.101930.

References

[1] L.D. Talbott, P.W. Ray, Molecular size and separability features of pea cell wall polysaccharides: implication for models of primary wall structure, Plant Physiol. 98 (1992) 357–368.

[2] K. Iiyama, Lam Thi Bach Tuyet, B.A. Stone, Covalent cross-links in the cell wall, Plant Physiol. 104 (1994) 315–320.

[3] C.T. Anderson, J.J. Kieber, Dynamic construction, perception, and remodeling of plant cell walls, Annu. Rev. Plant Biol. 71 (2020) 39–69, https://doi.org/10.1146/annurev-arplant-081519-035846.

[4] J. Vogel, Unique aspects of the grass cell wall, Curr. Opin. Plant Biol. 11 (2008) 301–307.

[5] M. Sobczak, W. Golinowski, Cyst nematodes and syncytia, in: J.T. Jones, G. Gheysen, C. Fenoll (Eds.), Genomics and Molecular Genetics of Plant-Nematode Interactions, Springer Academic Publishers, Dordrecht, The Netherlands, 2011, pp. 61–82, https://doi.org/10.1007/978-94-007-0434-3_4.

[6] E.L. Davis, A. Haegeman, T. Kikuchi, Degradation of the plant cell wall by nematodes, in: Genomics and Molecular Genetics of Plant-Nematode Interactions, Springer Netherlands, Dordrecht, 2011, pp. 255–272, https://doi.org/10.1007/978-94-007-0434-3_12. Jones J., Gheysen G. & Fenoll C.

[7] P. Abad, J. Gouzy, J.-M. Aury, P. Castagnone-Sereno, E.G.J. Danchin, E. Deleury, L. Perfus-Barbeoch, V. Anthouard, F. Artiguenave, V.C. Blok, M.-C. Caillaud, P. M. Coutinho, C. Dasilva, F. De Luca, F. Deau, M. Esquibet, T. Flutre, J.V. Goldstone, N. Hamamouch, T. Hewezi, O. Jaillon, C. Jubin, P. Leonetti, M. Magliano, T. R. Maier, G.V. Markov, P. McVeigh, G. Pesole, J. Poulain, M. Robinson-Rechavi, E. Sallet, B. Segurens, D. Steinbach, T. Tytgat, E. Ugarte, C. van Ghelder, P. Veronico, T.J. Baum, M. Blaxter, T. Bleve-Zacheo, E.L. Davis, J.J. Ewbank, B. Favery, E. Grenier, B. Henrissat, J.T. Jones, V. Laudet, A.G. Maule, H. Quesneville, M.N. Rosso, T. Schiex, G. Smant, J. Weissenbach, P. Wincke, Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*, Nat. Biotechnol. 26 (2008) 909–915.

[8] J.A. Cotton, C.J. Lilley, L.M. Jones, T. Kikuchi, A.J. Reid, P. Thorpe, I.J. Tsai, H. Beasley, V.C. Blok, P.J.A. Cock, S. Eves-van den Akker, N. Holroyd, M. Hunt, S. Mantelin, A. Naghra, A. Pain, J.E. Palomares-Rius, M. Zarowiecki, M. Berriman, J.T. Jones, P.E. Urwin, The genome and life-stage specific transcriptomes of *Globodera pallida* elucidate key aspects of plant parasitism by a cyst nematode, Genome Biol. 15 (2014) R43.

[9] S. Eves-van den Akker, D.R. Laetsch, P. Thorpe, C.J. Lilley, E.G.J. Danchin, M. Da Rocha, C. Rancurel, N.E. Holroyd, J.A. Cotton, A. Szitenberg, E. Grenier, J. Montarry, B. Mimee, M.-O. Duceppe, I. Boyes, J.M.C. Marvin, L.M. Jones, H. B. Yusup, J. Lafond-Lapalme, M. Esquibet, M. Sabeh, M. Rott, H. Overmars, A. Finkers-Tomczak, G. Smant, G. Koutsovoulos, V.C. Blok, S. Mantelin, P.J. A. Cock, W. Phillips, B. Henrissat, P.E. Urwin, M. Blaxter, J.T. Jones, The genome of the yellow potato cyst nematode, *Globodera rostochiensis*, reveals insights into the bases of parasitism and virulence, Genome Biol. 17 (2016) 124.

[10] E.G.J. Danchin, M.-N. Rosso, P. Vieira, J. de Almeida-Engeler, P.M. Coutinho, B. Henrissat, P. Abad, Multiple lateral gene transfers and duplications have promoted plant parasitism ability yin nematodes, Proc. Natl. Acad. Sci. U.S.A. 107 (2010) 17651–17656.

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

[12] E.G.J. Danchin, P. Thorpe, L. Perfus-Barbeoch, C. Rancurel, M. Da Rocha, K. Labadie, J. Guy, C. Da Silva, S. Bajew, R. Neilsen, E. Guzeeva, D. Esmenjaud, J. Helder, J.T. Jones, S. Eves-van den Akker, The transcriptomes of *Xiphinema index* and *Longidorus elongatus* reveal independent acquisition of parasitism genes by horizontal gene transfer in early-branching plant-parasitic nematodes, Genes 8 (2017) 287, https://doi.org/10.3390/genes8100287.

[13] S.S. Pokhare, P. Thorpe, P. Hedley, J. Morris, S.S. Habash, A. Elashry, S. Eves-van den Akker, F.M.W. Grundler, J.T. Jones, Signatures of adaptation to a monocot host in the plant-parasitic cyst nematode *Heterodera sacchari*, Plant J. 103 (2020) 1263–1274, https://doi.org/10.1111/tpj.14910, tpj.14910.

[14] C.A. Bell, C.J. Lilley, J. McCarthy, H.J. Atkinson, P.E. Urwin, Plant-parasitic nematodes respond to root exudate signals with host-specific gene expression patterns, PLoS Pathog. 15 (2) (2019), e100750, https://doi.org/10.1371/journal.ppat.1007503.

[15] M. Ochoa-Villarreal, E. Aispuro-Hernndez, I. Vargas-Arispuro, M. ngel, Plant cell wall polymers: function, structure and biological activity of their derivatives, in: Polymerization, InTech, 2012, https://doi.org/10.5772/46094.

[16] B.L. Ridley, M.A. O'Neill, D. Mohnen, Pectins: structure, biosynthesis, and oligogalacturonide-related signaling, Phytochemistry 57 (2001) 929–967.

[17] S. Pérez, M.A. Rodríguez-Carvajal, T. Doco, A complex plant cell wall polysaccharide: rhamnogalacturonan II. A structure in quest of a function, Biochimie 85 (2003) 109–121.

[18] S.R. Herron, J.A. Benen, R.D. Scavetta, J. Visser, F. Jurnak, Structure and function of pectic enzymes: virulence factors of plant pathogens, Proc. Natl. Acad. Sci. U.S. A. 97 (2000) 8762–8769.

[19] H. Popeijus, H. Overmars, J.T. Jones, V.C. Blok, A. Goverse, J. Helder, J. Bakker, A. Schots, G. Smant, Non-symbiotic degradation of plant cell walls by animals, Nature 406 (2000) 36–37.

[20] A. Haegeman, S. Mantelin, J.T. Jones, G. Gheysen, Secretions of plant parasitic nematodes, Gene 492 (2012) 19–31.

[21] S. Jaubert, J.B. Laffaire, P. Abad, M.N. Rosso, A polygalacturonase of animal origin isolated from the root-knot nematode *Meloidogyne incognita*, FEBS (Fed. Eur. Biochem. Soc.) Lett. 522 (2002) 109–112.

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

[23] C.H. Opperman, D. McK. Bird, V.M. Williamson, D.S. Rokhsar, M. Burke, J. Cohn, J. Cromer, S. Diener, S. Gajan, T.D. Houfek, Q. Liu, T. Mitros, J. Schaff, R. Schaffer, E. Scholl, B.R. Sosinski, V.P. Thomas, E. Windham, Sequence and genetic map of *Meloidogyne hapla:* a compact nematode genome for plant parasitism, Proc. Natl. Acad. Sci. U.S.A. 105 (2008) 14802–14807, https://doi.org/10.1073/pnas.0805946105.

[24] B. Vanholme, W. van Thuyne, K. Vanhouteghem, J. De Meutter, B. Cannoot, G. Gheysen, Molecular characterisation and functional importance of pectate lyases secreted by the cyst nematode Heterodera schachtii, Mol. Plant Pathol. 8 (2009) 267–278, https://doi.org/10.1111/j.1364-3703.2007.00392.x.

[25] J.T. Jones, R.H. Curtis, P.J. Wightman, P.R. Burrows, Isolation and characterisation of a putative collagen gene from the potato cyst nematode *Globodera pallida*, Parasitology 113 (1996) 581–588.

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

[27] S. Eves-van den Akker, C.J. Lilley, H.B. Yusup, J.T. Jones, P.E. Urwin, Functional C-terminally encoded plant peptide (CEP) hormone domains evolved de novo in the plant parasite *Rotylenchus reniformis*, Mol. Plant Pathol. 17 (2016) 1265–1275, https://doi.org/10.1111/mpp.12402.

[28] L.E. Bird, H. Rada, J. Flanagan, J.M. Diprose, R.J. Gilbert, R.J. Owens, Application of In-Fusion™ cloning for the parallel construction of *E. coli* expression vectors, in: Chapter 15 in Methods in Molecular Biology, 1116, 2014, pp. 209–234, https://doi.org/10.1007/978-1-62703-764-8_15.

[29] Y. Mei, P. Thorpe, A. Guzha, A. Haegeman, V.C. Blok, K. MacKenzie, G. Gheysen, J. T. Jones, S. Mantelin, 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 (2015) 409–424.

[30] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.

[31] A.M. Waterhouse, J.B. Procter, D.M.A. Martin, M. Clamp, G.J. Barton, Jalview Version 2-A multiple sequence alignment editor and analysis workbench, Bioinformatics 25 (2009) 1189–1191.

[32] S. Kalyaanamoorthy, B.Q. Minh, T.K.F. Wong, A. Von Haeseler, L.S. Jermiin, ModelFinder: fast model selection for accurate phylogenetic estimates, Nat. Methods 14 (2017) 587–589.

[33] L.T. Nguyen, H.A. Schmidt, A. Von Haeseler, B.Q. Minh, IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies, Mol. Biol. Evol. 32 (2015) 268–274.

[34] D.T. Hoang, O. Chernomor, A. Von Haeseler, B.Q. Minh, L.S. Vinh, UFBoot2: improving the ultrafast bootstrap approximation, Mol. Biol. Evol. 35 (2018) 518–522.

[35] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I. N. Shindyalov, P.E. Bourne, The protein Data Bank, Nucleic Acids Res. 28 (2000) 235–242.

[36] L.A. Kelley, S. Mezulis, C.M. Yates, M.N. Wass, M.J.E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis, Nat. Protoc. 10 (2015) 845–858.

[37] S. McNicholas, E. Potterton, K.S. Wilson, M.E.M. Noble, IUCr, Presenting your structures: the CCP4mg molecular-graphics software, Acta Crystallogr. Sect. D Biol. Crystallogr. 67 (2011) 386–394.

[38] J.T. Jones, C. Furlanetto, E. Bakker, B. Banks, V.C. Blok, Q. Chen, A. Prior, Characterisation of a chorismate mutase from the potato cyst nematode *Globodera pallida*, Mol. Plant Pathol. 4 (2003) 43–50.

[39] Y. Mei, K.M. Wright, A. Haegeman, L. Bauters, A. Diaz-Granados, A. Goverse, G. Gheysen, J.T. Jones, S. Mantelin, The *Globodera pallida* SPRYSEC effector *Gp*SPRY-414-2 that suppresses plant defenses targets a regulatory component of the dynamic microtubule network, Front. Plant Sci. 9 (2018) 1019.

[40] M. Madhaiyan, S. Poonguzhali, V.S. Saravanan, K. Hari, K.C. Lee, J.S. Lee, *Duganella sacchari* sp. nov. and *Duganella radicis* sp. nov., two novel species isolated from rhizosphere of field-grown sugar cane, Int. J. Syst. Evol. Microbiol. 63 (2013) 1126–1131, https://doi.org/10.1099/ijs.0.040584-0.

[41] M. Böger, J. Hekelaar, S.S. van Leeuwen, L. Dijkhuizen, A. Lammerts van Bueren, Structural and functional characterization of a family GH53 β-1,4-galactanase from *Bacteroides thetaiotaomicron* that facilitates degradation of prebiotic galactooligosaccharides, J. Struct. Biol. 205 (2019) 1–10, https://doi.org/10.1016/j.jsb.2018.12.002.

[42] B. Vanholme, A. Haegeman, J. Jacob, B. Cannoot, G. Gheysen, Arabinogalactan endo-1,4-β-galactosidase: a putative plant cell wall-degrading enzyme of plant-parasitic nematodes, Nematology 11 (2009) 739–747, https://doi.org/10.1163/156854109X404599.

[43] Z.A. Popper, Evolution and diversity of green plant cell walls, Curr. Opin. Plant Biol. 11 (2008) 286–292, https://doi.org/10.1016/j.pbi.2008.02.012.

[44] H.S. Gaur, R.N. Perry, The biology and control of the plant parasitic nematode, *Rotylenchulus reniformis*, Agric. Zool. Rev. 4 (1991) 177–212.

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

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

[47] L. Hu, R. Cui, L. Sun, B. Lin, K. Zhuo, J. Liao, Molecular and biochemical characterization of the beta-1,4-endoglucanase gene *Mj-eng-3* in the root-knot nematode *Meloidogyne javanica*, Exp. Parasitol. 135 (2013) 15–23.

[48] H. Peng, J. Cui, H. Long, W. Huang, L. Kong, S. Liu, H. Wenting, X. Xu, D. Peng, Novel pectate lyase genes of *Heterodera glycines* play key roles in the early stage of parasitism, PLoS One 11 (3) (2016), e0149959, https://doi.org/10.1371/journal.pone.0149959.

Supplementary Table 1: Primers used in the study.

Species	Primer name	Sequence	Length (bp)	Function
G. rostochiensis	08150_F	ACCATGCTGTACAAAGGTGC	20	Cloning into TOPO vector
G. rostochiensis	08150R_nostop	TTGGTAATTGAACGCTGTCATC	22	Cloning (absence of stop codon) into TOPO vector
G. rostochiensis	08150R_stop	TTATTGGTAATTGAACGCTGTC	22	Cloning (presence of stop codon) into TOPO vector
G. rostochiensis	Rosg08150_ISHF4	GGAGTGAAGAAGGCCGGTG	19	*In situ* hybridisation
G. rostochiensis	Rosg08150_ISHR4	CTCACGCGTTTCAGCATGTC	20	*In situ* hybridisation
G. pallida	FOR_4850	ACCATGCTGTACAAAGGTGCCGATGTC	27	Cloning
G. pallida	REV_4850_nostop	ACTGAACGCTTTCATCGCCT	20	Cloning (absence of stop codon)
G. pallida	REV_4850_stop	ATGTTAACTGAACGCTTTCATCG	23	Cloning (presence of stop codon)
G. pallida	Pal4850_ISHF3	GCACACCTACGGCATTTTGA	20	*In situ* hybridisation
G. pallida	Pal4850_ISHR3	TGTTGATCAGACTCGCCAGG	20	*In situ* hybridisation
R. reniformis	FOR_830247817	ACCATGCTCACAACGGGTGCCG	22	Cloning
R. reniformis	830247817_REV_nostop	TAGCGCACTCAATGCCTC	18	Cloning (absence of stop codon)
R. reniformis	830247817_REV_stop	TCTGATCATAGCGCACTCAA	20	Cloning (presence of stop codon)
R. reniformis	Ren30258_ISUHF	TTGTGAACCCGCCAGATG	18	*In situ* hybridisation
R. reniformis	Ren30258_ISUHR	AACCTGCACCCAATCCAC	18	*In situ* hybridisation

Supplementary Table 2: Sequences used for phylogenetic analysis.

# in tree	Sequence name in tree	Species name	Species type
1	RrGAL1	*R. reniformis*	Nematode
2	GpGAL1	*G. pallida*	Nematode
3	GrGAL1	*G. rostochiensis*	Nematode
4	WP_092947600.1	*Roseateles sp.* YR242	Bacteria
5	WP_058935312.1	*Roseateles depolymerans*	Bacteria
6	ALV07139.1	*Roseateles depolymerans*	Bacteria
12	WP_043094721.1	*Xanthomonas sacchari*	Bacteria
13	OYT87800.1	*Burkholderiales bacterium* PBB6	Bacteria
14	WP_020701442.1	*Oxalobacteraceae bacterium* AB_14	Bacteria
16	SFR90348.1	*Mitsuaria sp.* PDC51	Bacteria
20	WP_035053143.1	*Andreprevotia chitinilytica*	Bacteria
21	WP_072787792.1	*Duganella sacchari*	Bacteria
22	WP_047507752.1	*Methylibium sp.* CF468	Bacteria
26	WP_095575148.1	*Xanthomonas hortorum*	Bacteria
37	WP_076053575.1	*Xanthomonas campestris*	Bacteria
49	WP_070249534.1	*Duganella phyllosphaerae*	Bacteria
50	WP_047126143.1	*Xanthomonas arboricola*	Bacteria
54	WP_054393996.1	*Xanthomonas vasicola*	Bacteria
56	WP_082569018.1	*Rhizobacter sp.* Root1221	Bacteria
66	WP_007962641.1	*Xanthomonas citri*	Bacteria
73	WP_017158928.1	*Xanthomonas phaseoli*	Bacteria
78	WP_013634183.1	*Pseudopedobacter saltans*	Bacteria
79	AAA32692.1	*Aspergillus aculeatus*	Fungus
80	4BF7_A	*Aspergillus nidulans*	Fungus
81	1HJQ_A	*Humicola insolens*	Fungus
82	1HJS_A	*Thermothelomyces thermophilus*	Fungus

Supplementary Figure 1: Expression and purification of the recombinant GH53 enzymes and positive (GalA_Xc) and negative (GFP) controls. **[M]** PageRuler Plus molecular mass marker (Invitrogen); **[1]** pre-induction sample; **[2]** post-induction (IPTG) sample; **[3]** total soluble protein; **[4]** resin wash; **[5]** purified protein.