Mol Gen Genomics (2004) 272: 138–148
DOI 10.1007/s00438-004-1054-0

ORIGINAL PAPER

M. Mitreva · A. A. Elling · M. Dante · A. P. Kloek
A. Kalyanaraman · S. Aluru · S W. Clifton
D. McK. Bird · T. J. Baum · J. P. McCarter

A survey of SL1-spliced transcripts from the root-lesion nematode *Pratylenchus penetrans*

Received: 30 March 2004 / Accepted: 6 August 2004 / Published online: 28 August 2004

Abstract Plant-parasitic nematodes are important and cosmopolitan pathogens of crops. Here, we describe the generation and analysis of 1928 expressed sequence tags (ESTs) of a splice-leader 1 (SL1) library from mixed life stages of the root-lesion nematode *Pratylenchus penetrans*. The ESTs were grouped into 420 clusters and classified by function using the Gene Ontology (GO) hierarchy and the Kyoto KEGG database. Approximately 80% of all translated clusters show homology to *Caenorhabditis elegans* proteins, and 37% of the *C. elegans* gene homologs had confirmed phenotypes as assessed by RNA interference tests. Use of an SL1-PCR approach, while ensuring the cloning of the 5′ ends of mRNAs, has demonstrated bias toward short transcripts. Putative nematode-specific and *Pratylenchus* -specific genes were identified, and their implications for nematode control strategies are discussed.

Keywords *Pratylenchus* · Expressed sequence tags (ESTs) · Comparative genomics · Gene expression · Parasite

Communicated by W. R. McCombie

Electronic Supplementary Material Supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s00438-004-1054-0

M. Mitreva (✉) · M. Dante · S W. Clifton · J. P. McCarter
Genome Sequencing Center, Department of Genetics,
Washington University School of Medicine, St. Louis,
MO 63108, USA
E-mail: mmitreva@watson.wustl.edu
Tel.: +1-314-2861118
Fax: +1-314-2861810

A. A. Elling · T. J. Baum
Interdepartmental Genetics Program,
Iowa State University, Ames, IA 50011, USA

A. A. Elling · T. J. Baum
Department of Plant Pathology, Iowa State University,
Ames, IA 50011, USA

A. P. Kloek · J. P. McCarter
Divergence Inc., 893 North Warson Road, St. Louis,
MO 63141, USA

A. Kalyanaraman · S. Aluru
Department of Electrical and Computer Engineering,
Iowa State University, Ames, IA 50011, USA

D. McK. Bird
The Center for the Biology of Nematode Parasitism,
North Carolina State University, Raleigh,
NC 27695, USA

Introduction

Nematodes are the most prevalent animals in the world (Boucher and Lambshead 1994). Although the roundworm body plan is broadly conserved (Bird and Bird 1991), different species exhibit unique life cycle adaptations, including parasitism of plants, insects and vertebrates (Blaxter and Bird 1997). Phylogenetic analysis (Blaxter et al. 1998) has shown parasitism to be phylum-wide, with multiple independent evolutionary origins (Blaxter 2003). To date, plant parasitism has been observed for members of three nematode clades (Blaxter et al. 1998). Species of Clade IV, which includes the tylenchid nematodes, are especially devastating to agricultural production (Koenning et al. 1999) and a significant effort has been made to generate expressed sequence tags (ESTs) from the Tylenchida. To date, approximately 100,000 sequences have been deposited in NCBI's GenBank, mainly from the root-knot (*Meloidogyne* spp.) and cyst (*Heterodera* and *Globodera* spp.) nematodes (McCarter et al. 2000, 2003a, 2003b; Popeijus et al. 2000a; Dautova et al. 2001b; Parkinson et al. 2003; Wylie et al. 2003). These genera establish elaborate long-term feeding sites within host roots. By contrast, root-lesion nematodes (genus *Pratylenchus*) are migratory plant endoparasites (Zunke 1990b). They have a wide host range, suggesting that *Pratylenchus* is a less specialized (i.e., more primitive) form of plant-parasite, possibly representing an evolutionary intermediate between the very specialized sedentary endoparasites and

free-living forms. The evolutionary position of *Pratylenchus,* as a member of the Tylenchida*,* makes this an excellent candidate for investigation of the evolution of plant parasitism by nematodes. Furthermore, root-lesion nematodes are important plant pathogens, and a survey of the *Pratylenchus* genome will help to elucidate the interactions between *P. penetrans* and its host plants.

Pratylenchus penetrans has a cosmopolitan distribution, and attacks underground plant organs including roots, tubers and rhizomes, where populations of 10,000–35,000 individuals per 10 g of root can develop (Nickle 1991). It has been reported as a pathogen of many crops, including legumes (Townshend 1978; Elliott and Bird 1985), vegetables (Townshend 1963a; Hung and Rohde 1973), strawberries (Townshend 1963b), fruits (Mountain and Patrick 1959; Pitcher et al. 1960; Wyss 1970), corn and potato (Dickerson et al. 1964) and turfgrasses (Troll and Rohde 1966). The life cycle of the nematode is completed in 45–65 days, so that several generations can develop within one growing season. The nematodes remain vermiform and motile throughout their larval and adult stages, and all stages from L2 on can infect plants by penetrating the root and invading the cortex. The main entry points are the regions of root hair development (Zunke 1990b) and the elongation zone (Troll and Rohde 1966; Townshend 1978). *P. penetrans* migrates intracellularly by puncturing cell walls with its stylet and may feed briefly before invading a cell. In addition to mechanical force generated through its stylet and body (Zunke 1990b), *P. penetrans* also appears to degrade cell walls by secreting cellulolytic enzymes (Krusberg 1960; Uehara et al. 2001). During typical feeding periods (5–10 min) a salivation zone develops at the stylet tip and cytoplasmic streaming increases in the plant cell, but the affected cell does not die (Zunke 1990b). However, long-term feeding (hours) results in profound morphological plant cellular changes, including shrinking of the feeding cell tonoplast, nuclear hypertrophy, vacuole formation, and eventual death with condensed cytoplasm and degenerated organelles (Townshend et al. 1989; Zunke 1990b). Ectoparasitic feeding on root hairs has been reported (Zunke 1990a) but is controversial (Kurppa and Vrain 1985). Following *P. penetrans* invasion lesions develop first as water-soaked areas and become discolored later on (Townshend and Stobbs 1981). Since *P. penetrans* is able to hydrolyze amygdalin in vitro (Mountain and Patrick 1959), it has been hypothesized that a potential mechanism for lesion development is the production of phytotoxins through hydrolysis of the plant β-glucoside amygdalin. In agreement with this view, *P. penetrans* - infected cells accumulate large amounts of phenols, which could be produced by damaged plant cells in the lesion or cleaved from glycosides by nematode enzymes (Acedo and Rohde 1971). Oxidation of these phenols may convert lesions into necrotic areas.

Here we describe the generation and analysis of 1928 ESTs from a *P. penetrans* mixed-stage library generated by splice-leader 1 (SL1) amplification. This is the first genomic survey approach to any migratory endoparasitic nematode of plants. We used semi-automated bioinformatics tools to exploit and interpret the data. This first analysis may set the stage for future genomic studies of *P. penetrans* and improved understanding of the biology of this important parasite.

Materials and methods

Preparation of mixed-stage RNA and construction of a cDNA library

A monoxenic *P. penetrans* population was maintained on corn-roots grown on Gamborg's plates (Sigma), and harvested by washing the worms free of the roots with M9 buffer. Worm pellets were frozen in liquid nitrogen, and pulverized using an Alloy Tool Steel Set (Fisher Scientific International). Total parasite RNA was prepared using TRIzol reagent (GibcoBRL, Life Technologies), and mRNA was extracted from 2 μg of total RNA using a Dynabeads mRNA purification kit (Dynal Biotech) and eluted into 10 μl of 10 mM TRIS-HCl. First strand synthesis was performed using a linker primer (5′-GAGAGAGAGAGAGAGAGAACTAGTCTCGAGTTTTTTTTTTTTT-3′) and Superscript II RT (Invitrogen, Life Technologies). Amplification with *Taq* polymerase used the SL1 (5′-GGGTTTAATTACCCAAGTTTGA-3′) and Xhop (5′-GAGAGAGAACTAGTCTCGA-3′) primers and 5 μl of the first-strand reaction. Cycling parameters were: 95°C for 5 min, 30 cycles of 95°C for 1 min, 47°C for 1 min and 72°C for 3 min, followed by 5 min at 72°C. cDNA fragments >1 kb long were selected after electrophoresis on a 0.8% TAE agarose gel, and cloned into the pCRII-TOPO vector following the TOPO TA protocol (Invitrogen). The ligation mix was then introduced into chemically competent *Escherichia coli* DH10B cells (GibcoBRL).

Sequencing and clustering

Sequencing, EST processing and clustering were performed as described previously (McCarter et al. 2003a; Mitreva et al. 2004). Submissions were deposited in GenBank in June and July 2002. Information for clone requests and sequence trace files is available at http://www.nematode.net. The completed cluster assembly, NemaGene *Pratylenchus penetrans* v 1.0, was used as the basis for all subsequent analysis and is available for searching and acquisition by FTP at http://www.nematode.net. ''Fragmentation'', defined as the representation of a single gene by multiple non-overlapping clusters, was estimated by examining *P. penetrans* clusters with homology to *C. elegans* (Mitreva et al. 2004). Overall representation of *P. penetrans* genes is based on a theoretical gene number of 21,712, comparable to that of *C. elegans* (Wormpep v. 104).

Analysis and functional assignments

WU-BLAST sequence comparisons (Altschul et al. 1990; Gish 1996–2002; http://blast.wustl.edu) were performed using 488 contig sequences grouped into 420 clusters. Clusters were used to search multiple databases, including the non-redundant GenBank (27 June 2003) and Wormpep v.104 *C. elegans* (Welcome Trust Sanger Institute, unpublished) protein databases. Internally constructed databases based on intersections of data from GenBank (Wheeler et al. 2001) allowed examination of sequences in specific phylogenetic distributions. Homologies were reported for e-value scores of $\leq$ 1e-05.

To identify cases in which *P. penetrans* homologs in *C. elegans* have been surveyed for a phenotype using RNA interference, Wormpep BLAST matches were cross-referenced to a list of 17,913 *C. elegans* genes for which information on RNAi-induced phenotypes was available (as of 28 March 2003; http://www.wormbase.org). For each *P. penetrans* cluster, only the best *C. elegans* match was considered.

Clusters were assigned to putative functional categories using two methods. First, InterProScan v.3.1 (ftp://ftp.ebi.ac.uk/pub/software/unix/iprscan) was used to search contig translations versus InterPro domains (8 November 2002) (Apweiler et al. 2001; Zdobnov and Apweiler 2001). Using InterPro, clusters were mapped to the three organizing principles of Gene Ontology (categorization scheme version http://go_200211_assocdb.sql; The Gene Ontology Consortium 2000). Mappings are stored in a MySQL database, displayed using AmiGo (25 November 2002) (http://www.godatabase.org/cgi-bin/go.cgi), and are available at http://www.nematode.net. Second, clusters were assigned by Enzyme Commission Number (IUBMB 1992) to metabolic pathways using the Kyoto Encyclopedia of Genes and Genomes database (KEGG; IUBMB 1992; Bono et al. 1998; Kanehisa and Goto 2000). Complete KEGG mappings are available at http://www.nematode.net.

Results

Sequence generation and cluster analysis

A total of 1928 5′ ESTs were obtained from an SL1-PCR-based cDNA library prepared from a mix of all motile life-cycle stages of the root-lesion nematode *P. penetrans*, and submitted to GenBank. The mean length of the ESTs submitted was 580 nt and their average GC content was 44%. Clustering, implemented to reduce data redundancy and improve sequence quality, discarded two problematic ESTs and grouped the remaining 1926 ESTs by sequence identity into 488 contigs, which were in turn organized into 420 clusters. ESTs within a contig are derived from nearly identical transcripts, whereas contigs within a cluster may represent different alleles, splice isoforms of a gene, or closely related gene family members that cannot be definitively classified as separate genes (McCarter et al. 2003a). In 172 cases, clusters consist of a single EST (9% of all ESTs analyzed), whereas the largest single cluster contains 121 ESTs (Fig. 1). The majority of clusters (53%) had 10 or fewer ESTs. Based on the identified clusters, this project generated sequence data for 420 genes, corresponding to a new gene discovery rate of 22% (420/1928). Fragmentation, whereby one gene is represented by multiple non-overlapping clusters (Mitreva et al. 2004), was not observed for this *P. penetrans* dataset, most probably because the use of an SL1 primer in

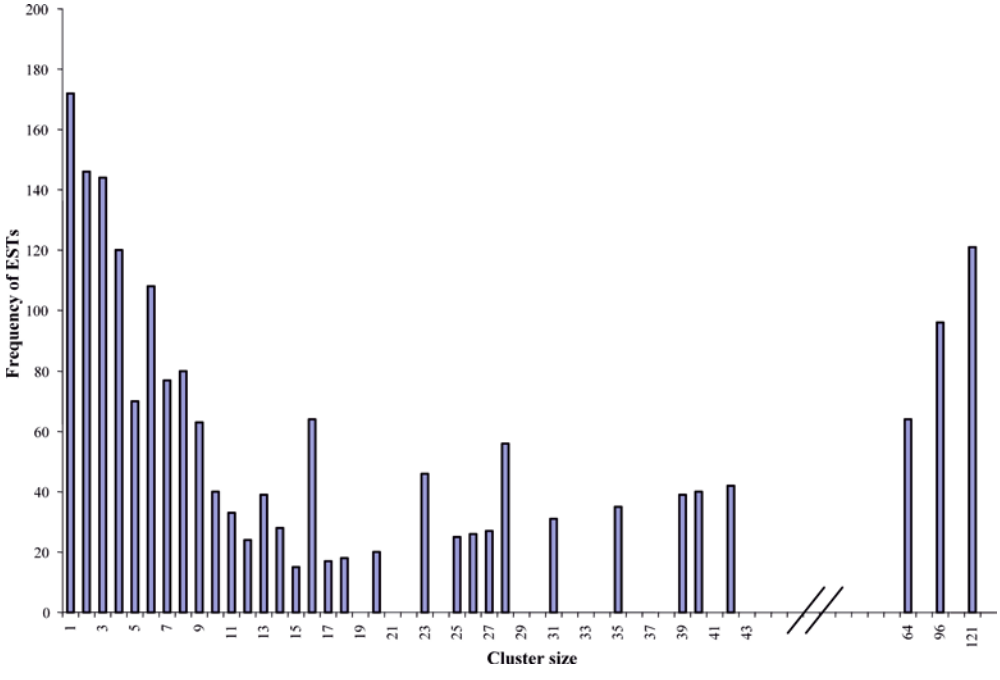

Fig. 1 Histogram showing the distribution of *P. penetrans* ESTs by cluster size

library construction ensured that the sequenced clones extended to the 5′ ends of the mature transcripts. Contig building reduced the number of nucleotides for further analysis from 1,158,654 to 277,622. Sampling of another 1928 ESTs from the same source is estimated to result in the discovery of only 144 new clusters, a new gene discovery rate of only 7.5% (ESTFreq; W. Gish, personal communication). Further sampling from this species will therefore have to await the construction of a new cDNA library.

Transcript abundance

Generally, high cDNA abundance in a library correlates with a high transcript level in the organism under study at the point of sampling (Audic and Claverie 1997). The 15 most abundant clusters (Table 1) made up 34% of all ESTs, and included genes encoding a putative GTP cyclohydrolase I (PP00075.cl), a member of the MIP family of pore-forming proteins (PP00272.cl), a collagen (PP00267.cl), a neuropeptide-like protein NLP-21 (PP00253.cl), and a voltage-dependent anion channel-related protein (PP00258.cl). In addition to clusters with identifiable homology, we found three highly abundant clusters with no known sequence similarities: PP00266.cl, PP000263.cl and PP000262.cl. These three putatively novel genes encode predicted ORFs of 273, 218 and 236 amino acid residues respectively; it is possible that they represent *Pratylenchus* -specific genes.

Characteristics of libraries generated by SL1-PCR

For many nematode genes, *trans* -splicing of a short leader sequence to the 5′ end of the mRNA is a feature of transcript maturation. The most common *trans* -splice leader is SL1, the sequence of which is highly conserved across the phylum (Krause and Hirsh 1987; Conrad et al. 1991; Blaxter and Liu 1996). It is estimated that 80% of *Ascaris suum* transcripts (Nilsen 1993), 70% of *C. elegans* transcripts (Blumenthal and Steward 1997) and 60% of *Globodera rostochiensis* transcripts are (Ling Qin, personal communication) SL1 *trans*-spliced. While the extent to which each nematode species uses SL1 is not known, we have been able to successfully make SL1-PCR libraries from 18 nematode species, including seven plant parasites (http://www.nematode.net). SL1-PCR libraries are useful as they ensure that full-length cDNAs are cloned, and they can be produced from very small amounts of starting RNA. Also, since they use an SL1-specific 22mer for amplification, they ensure that the cDNAs cloned are extremely likely to be of nematode origin and not a product of host contamination.

Since this analysis of *P. penetrans* is one of the first to examine nematode ESTs generated from an SL1 splice leader PCR library (Hoekstra et al. 2000), we investigated whether the use of such a library resulted in any bias in the characteristics of the identified transcripts. For instance, since reverse transcriptase is known to have limited processivity, often terminating before reaching the 5′ end of longer transcripts (Skalka and Goff 1993), and PCR amplification with an SL1 primer requires that the first-strand cDNA should extend to the 5′ end of the mRNA template, we suspected that our EST collection would be biased toward shorter transcripts. In addition, shorter clones will have an amplification advantage during PCR cycles. A bias toward shorter transcripts was indeed observed. Using *C. elegans* Wormpep as a reference set of putatively full-length nematode peptide sequences, we compared sets of Wormpep proteins that were BLAST matched (most significant score) by the SL1-derived *P. penetrans* clusters with proteins matched by EST clusters from three other parasitic nematode species made from conventional (non-splice leader) libraries (Table 2). The complete set of over 21,000 *C. elegans* proteins in Wormpep has a mean length of 441 amino acids and the set with homology to clusters in *Meloidogyne incognita, Strongyloides stercoralis* and *Trichinella spiralis* had a mean length of 513 residues. In contrast, the profile of Wormpep proteins matching *P. penetrans* SL1

Table 1 The 15 most abundantly represented transcripts in the *P. penetrans* SL1 cDNA library

Cluster	ESTs per cluster	E-value	Descriptor	GenBank Accession No.
PP00075.cl	121	1e-131	*Ostertagia ostertagi* GTP-cyclohydrolase I	AAC06296
PP00266.cl	96	-	Novel	-
PP00272.cl	64	3e-06	*Drosophila virilis* Pore-forming protein MIP family	AAC38845
PP00271.cl	42	7e-05	*C. elegans* Putative nuclear protein	NP_490809
PP00270.cl	40	4e-20	*C. elegans* Hypothetical 41.7 kDa protein T20F5.3	P91478
PP00269.cl	39	1e-21	*Plasmodium falciparum* Glutamic acid-rich protein	P13816
PP00267.cl	35	4e-134	*Ascaris suum* Putative cuticular collagen	CAB85465
PP00265.cl	31	3e-42	*Mus musculus* Unnamed protein product	BAC40503
PP00263.cl	28	-	Novel	-
PP00264.cl	28	7e-15	*Betula pendula* Embryonic protein BP8	CAA79329
PP00240.cl	27	6e-117	*C. elegans* Hypothetical protein K04D7.1	CAA93514
PP00262.cl	26	-	Novel	-
PP00260.cl	25	1e-05	*Drosophila melanogaster* Mam CG8118-PA	AAF58299
PP00259.cl	23	3e-06	*C. elegans* Neuropeptide-Like Protein NLP-21	NP_499466
PP00258.cl	23	6e-78	*C. elegans* Voltage-dependent anion channel 2	NP_501211

Table 2 Comparison of the length of Wormpep proteins with homology to other nematode species

	Species						
					Mi, *Ss* , *Ts*	*C. elegans*	
	P. penetrans	*M. incognita*	*S. stercoralis*	*T. spiralis*	pooled	Wormpep104	Ribosomal
Wormpep protein length (aa)	243/420	970/1625	2090/3311	1391/3454	4451/8390	21342	156
Mean	263	511	555	451	513	441	307
SD	252	513	599	367	520	409	374
Median	189	390	419	369	397	345	194

[a]number of clusters with homology to *C. elegans* proteins; [b]total number of clusters (genes)

transcripts had a mean length of only 263 amino acids, or about half that seen for the other species, a highly significant difference according to a two-tailed T-test ($P < 0.001$; Snedecor and Cochran 1967). The median lengths also showed about a two-fold difference.

Plotting the lengths of proteins in the *P. penetrans* SL1 matching set revealed a sharp cut-off towards larger sizes, with 95% of polypeptides being less than 640 amino acids long (Supplementary Figure S1). This indicates that, while portions of long transcripts encoding high molecular weight proteins are included in conventional cDNA libraries, SL1-PCR libraries, even with gel selection for inserts of >1 kb (see Materials and methods), included a majority of transcripts of less than 600 nt, with an upper limit of <2000 nt (excluding the 3′ and 5′ UTRs). Enrichment for short transcripts is also observed in other SL1 based libraries that we have generated (M. Mitreva and J. McCarter, unpublished). In addition to showing bias towards short transcripts, the *P. penetrans* SL1-PCR based library obviously represents only mRNA species that are preceded by an SL1 sequence at their 5′ ends. One mRNA fragment carrying a SL1 sequence has been isolated from *P. penetrans* (GenBank Accession No. AAF77031, 85% identity with PP00140.cl, 49/57); however, the fraction of *P. penetrans* transcripts that actually have an SL1 leader is not known.

Functional classification based on Gene Ontology and KEGG assignments

To categorize transcripts by putative function, we used the Gene Ontology (GO) classification hierarchy (http://www.geneontology.org; The Gene Ontology Consortium 2000). GO provides a classification scheme based on a dynamic controlled vocabulary and hierarchy that encompasses descriptions of molecular, cellular and biological functions across genomes. Initially, the *P. penetrans* clusters were matched to InterPro protein domains by InterProScan (http://www.ebi.ac.uk/interpro), then existing mappings allowed organization into the GO hierarchy. The GO classification is presented for each of the main organizing principles and their respective subcategories (Fig. 2, Supplementary Table S1). Of the 420 clusters analyzed, 223 (53%) aligned to 264 unique InterPro domains, and 147 (35%) mapped to GO domains. Some 75% of all the *P. penetrans* clusters with mappings for biological processes are involved in metabolic activities, among which protein metabolism is the most prominent (30%), followed by nucleic acid metabolism (11%). One-third of molecular function mappings are involved in binding activities, with nucleic acid (17%) and nucleotide binding (15%) being the most common. Mappings for catalytic activities are common (16%), including GTPase hydrolase (3%) and GTP cyclohydrolase (2%). We compared the GO mappings of *P. penetrans* to those available for four other nematode species, including *C. elegans* (Supplementary Table S2). The most striking differences were an over-representation of 'structural proteins' (29% vs. 2–17%) and an under-representation of enzymes (16% vs. 31–39%). Both of these findings are consistent with the bias in our SL1-PCR library toward shorter transcripts, and are not likely to reflect actual differences in the biology of *P. penetrans* relative to other nematodes. Almost all of the structural proteins observed in *P. penetrans* were ribosomal proteins. Many ribosomal proteins are quite small, with a median length in *C. elegans* of just 194 amino acids (Table 2).

As an alternative classification scheme, clusters were assigned to metabolic pathways using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (http://www.genome.ad.jp/kegg). Of the 420 clusters, 41 (10%) encode 30 putative enzymes mapped to 36 metabolic pathways (Table 3). Pathways well represented by *P. penetrans* clusters included carbohydrate and nucleotide metabolism. Complete listings of the KEGG mappings, including graphical representations, are available at http://www.nematode.net.

Homology matches to other organisms

BLAST comparisons were made using the *P. penetrans* clusters to search three databases: predicted *C. elegans* proteins, other translated nematode sequences, and translated non-nematode sequences (Fig. 3). The great majority of clusters matched public database entries (80.2%, 337/420) meeting a cut-off value of $<$1e-05 (Fig. 3). Of all *P. penetrans* clusters with homologies, the majority (62.6%) matched sequences in all three

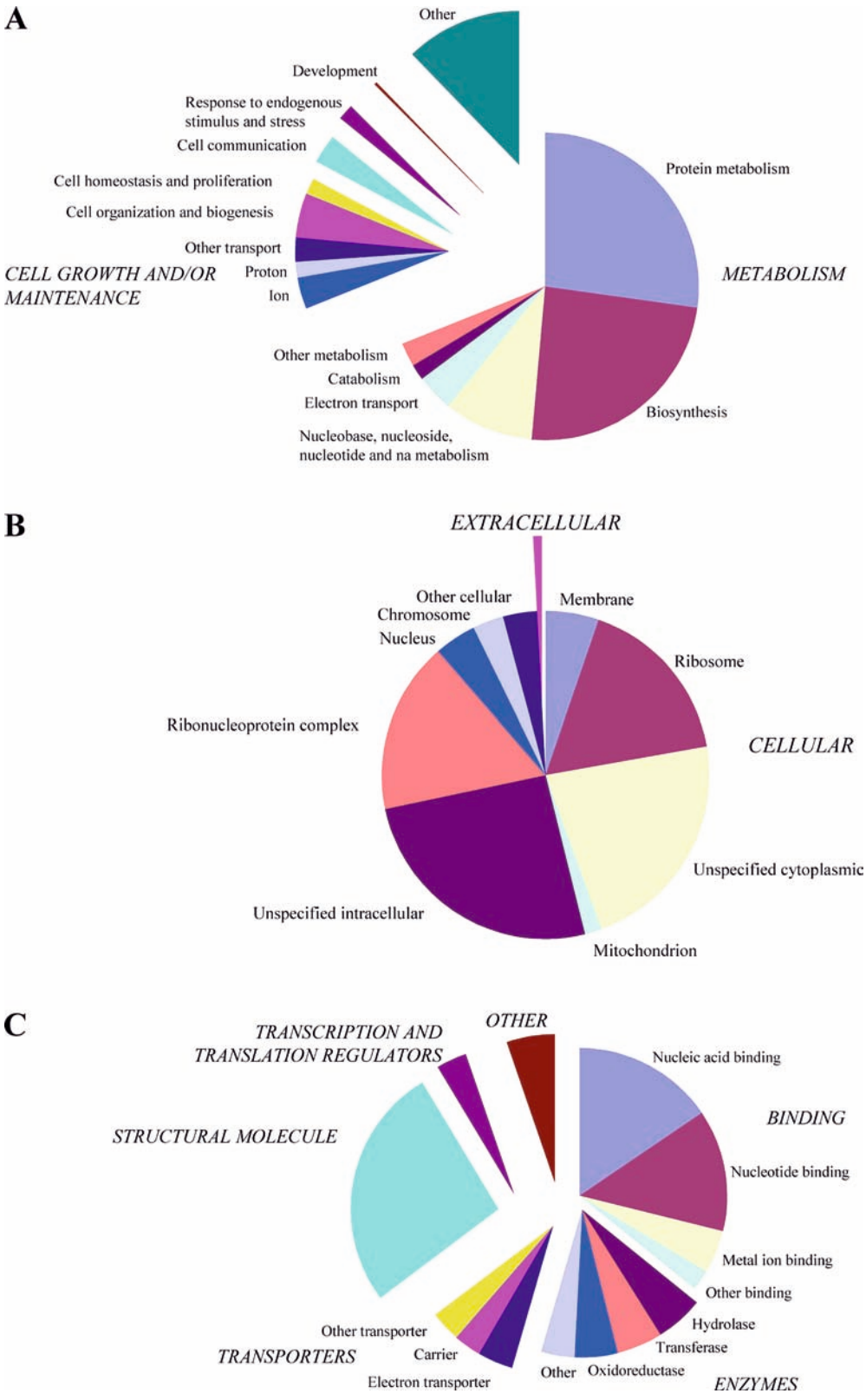

Fig. 2A-C Gene Onthology mappings. **A** Biological process. **B** Cellular component. **C** Molecular function**.** The actual mappings are more complicated than one-to-one, since individual GO categories can have multiple mappings. For instance, GO:0015662: P-type ATPase (PP00115.cl, InterPro domain IPR004014) is a nucleotide-binding protein, a hydrolase and a transporter. Hence, some values in Supplementary Table S1 add up to more than 100%

databases, as has been observed in previous studies (McCarter et al. 2003a; Mitreva et al. 2004). Many gene products in this category are conserved across metazoans, and are involved in core biological processes. In contrast to our prior studies of *M. incognita* (McCarter et al. 2003a) and *S. stercoralis* (Mitreva et al. 2004), the *C. elegans* genome alone was no longer the best source for *P. penetrans* sequence homologies (78.9%), as 95.5% of the clusters matched nematode sequences from species other than *C. elegans*. The prominence of matches to

Table 3 KEGG biochemical pathway mappings for *P. penetrans* clusters

KEGG pathway[a]	*P. penetrans* clusters	Enzymes
1.2 Citrate cycle (TCA cycle)	1	1
1.4 Pentose and glucuronate interconversions	2	2
1.5 Fructose and mannose metabolism	2	3
1.6 Galactose metabolism	2	2
1.7 Ascorbate and aldarate metabolism	1	1
1.10 Propanoate metabolism	1	1
1.11 Butanoate metabolism	2	3
2.1 Oxidative phosphorylation	6	3
2.2 ATP synthesis	4	1
2.6 Methane metabolism	1	1
2.7 Nitrogen metabolism	2	1
3.1 Fatty acid biosynthesis (path 1)	1	1
3.2 Fatty acid biosynthesis (path 2)	1	1
3.3 Fatty acid metabolism	1	1
3.6 Bile acid biosynthesis	1	1
3.8 Androgen and estrogen metabolism	1	1
4.1 Purine metabolism	5	3
4.2 Pyrimidine metabolism	4	4
4.3 Nucleotide sugars metabolism	1	1
5.3 Glycine, serine and threonine metabolism	1	1
5.4 Methionine metabolism	1	1
5.6 Valine, leucine and isoleucine degradation	1	1
5.9 Lysine degradation	2	2
5.11 Prostaglandin and leukotriene metabolism	1	1
5.13 Phenylalanine metabolism	1	1
5.14 Tryptophan metabolism	1	2
6.9 Glutathione metabolism	3	2
8.1 Glycerolipid metabolism	3	3
8.6 Sphingoglycolipid metabolism	1	1
8.9 Globoside metabolism	1	2
9.3 Vitamin B6 metabolism	1	1
9.7 Folate biosynthesis	2	3
9.10 Porphyrin and chlorophyll metabolism	2	2
10.3 Flavonoids, stilbene and lignin biosynthesis	1	1
11.9 Tetrachloroethene degradation	1	1
11.16 Benzoate degradation via CoA ligation	3	3

[a] 65 total and 60 unique mappings; 36 metabolic pathways are represented out of 88 possible

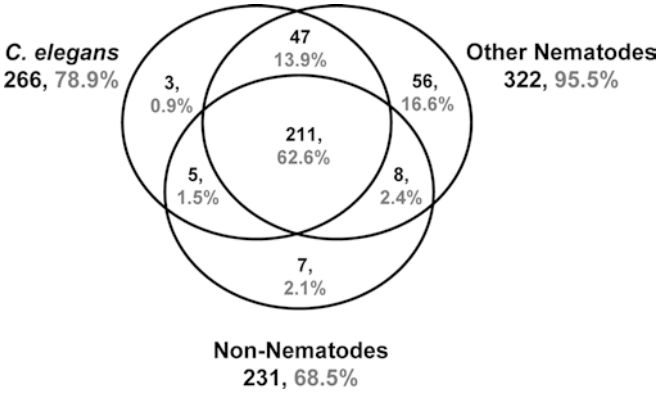

Fig. 3 Distribution of *P. penetrans* BLAST matches by database

'other nematodes' is most likely to be due to the continuous increase in the number of sequences from Tylenchida and other Clade IV spcies in public databases. There are now nearly 100,000 Tylenchida ESTs and 34,000 other Clade IV ESTs in dbEST, most of them submitted by the Genome Sequencing Center at Washington University (GSC) since our prior analyses. We found that 106 *P. penetrans* clusters (i.e., 25% of the clusters with matches to known sequences) had homologies only to nematodes, of which the most conserved match was to the M60.4b gene of *C. elegans*, with a BLAST e-value of 2e-72 (PE00128.cl). Table 4 lists the 10 most conserved nematode-specific clusters. The majority of these gene products are hypothetical proteins without known function.

For one-fifth of the *P. penetrans* clusters no significant sequence similarities were detected to any protein in the public databases. Some of these clusters may represent genes that are specific to the genus *Pratylenchus* or even to *P. penetrans*. Other clusters may lack homology either because they lack an ORF, i.e., are mainly untranslated regions or, because they span a portion of an ORF that is too short to generate a significant match. To assess the real number of novel or diverged amino acid sequences, the ORF length for contigs with and without BLAST homology was compared (Supplementary Figure S2). The mean ORF length is shorter in contigs without homology (124 amino acids) than for

Table 4 The 10 most conserved nematode-specific *P. penetrans* clusters

P. penetrans cluster	ESTs per cluster	Descriptor	Accession GenBank	E-value
PP00128.cl	3	*C. elegans* hypothetical protein M60.4b	AAP40521	2e - 72
PP00279.cl	1	*C. elegans* hypothetical protein M05B5.2	CAA95834	6e - 66
PP00254.cl	18	*C. elegans* hypothetical protein C39E9.8	T19850	2e - 64
PP00595.cl	1	*H. glycines* esophageal gland cell secretory protein 12	AAF76926	2e - 61
PP00069.cl	2	*C. elegans* hypothetical protein Jc8.8	CAB05225	2e - 56
PP00252.cl	16	*C. elegans* hypothetical protein M60.4b	AAP40521	2e - 52
PP00182.cl	5	*C. elegans* putative protein	NP_508619	4e - 48
PP00107.cl	3	*C. elegans* hypothetical protein F22A3.2	AAA83193	3e - 45
PP00041.cl	2	*C. elegans* hypothetical protein C40H1.5	CAA79556	9e - 45
PP00553.cl	1	*C. elegans* hypothetical protein F58A3.5	CAB02670	6e - 42

those with homology (161 aa), a significant difference at >99% confidence (two-tailed t-test with unequal variance; Snedecor and Cochran 1967). Characterization of novel *Pratylenchus* genes should begin with those containing longer, and therefore more probably real, ORFs.

Comparison to *C. elegans* genes with RNAi phenotypes

We examined *C. elegans* genes for which RNAi information is available and which have *P. penetrans* homologs, because the targeted knock-out of individual genes in *C. elegans* provides a substantial amount of new data which may be transferable to orthologous genes in parasitic nematodes and thus allow inferences as to their functions in the latter. Eighty percent (266/337) of all *P. penetrans* clusters with homology matched *C. elegans* genes. These clusters were compared to a list of 17,913 *C. elegans* genes for which RNAi data are available (Fraser et al. 2000; Gonczy et al. 2000; Maeda et al. 2001; Kamath et al. 2003). Of the 231 cases for which *C. elegans* RNAi information was available for the best BLAST match to the *P. penetrans* cluster, a phenotype has been described for 97 (36%; Supplementary Table S3)—a more than threefold enrichment in phenotypes versus the entire *C. elegans* dataset (Supplementary Figure S3). *C. elegans* genes with expressed *P. penetrans* homologs were also enriched for genes with severe RNAi phenotypes (Supplemental Figure S3, Supplementary Table S4). Previous studies showed a correlation between sequence conservation and severe RNAi phenotypes (Fraser et al. 2000; Gonczy et al. 2000; McCarter et al. 2003a). Hence, certain genes in the *P. penetrans* dataset may encode proteins which, if disabled, could compromise survival of the parasite. One group of particular interest includes proteins that are required for nematode survival and lack homology outside the phylum Nematoda, since these are potential targets for nematode control without toxicity to the host or other non-target organisms. Of the 47 *P. penetrans* clusters with homologs only to *C. elegans* and other nematodes (Fig. 3), 7 had matches to *C. elegans* genes with RNAi phenotypes. For example, PP00047.cl matched T05H4.12 (9e-16), a gene with no close mammalian homolog, which nevertheless contains a coupling factor 6 domain (Pfam05511) which is a component of mitochondrial ATP synthase and is required for the interactions of the catalytic and proton-translocating segments (Javed et al. 1991). The RNAi knockout caused embryonic, larval and adult lethality, sterility and slow growth in *C. elegans*. Homologs are found in *C. briggsae*, nematode parasites of animals (*Haemonchus contortus*, *Ascaris suum*, *Strongyloides stercoralis*, *Necator americanus*) and plants (*Meloidogyne hapla*, *M. chitwoodi*). Therefore, it is expected that further analysis will identify genes, in addition to those already mentioned, which warrant more detailed investigation.

Horizontal gene transfer candidates

Plant-parasitic nematodes are known to possess prokaryotic-like genes for cell wall degrading enzymes and other functions. Most studies have focused on cyst and root-knot nematodes, where cellulases (de Meutter et al. 1998; Smant et al. 1998; Rosso et al. 1999; Goellner et al. 2000), xylanase (Dautova et al. 2001a), pectate lyase (Popeijus et al. 2000b; Doyle and Lambert 2002), polygalacturonase (Jaubert et al. 2002), and NodL acetyltransferase (McCarter et al. 2003a; Scholl et al. 2003) have been identified. Such genes may have been acquired by horizontal gene transfer. To our knowledge, the only known prokaryotic-like *P. penetrans* gene product isolated so far is a β-1,4-endoglucanase homolog (Uehara et al. 2001). A search of our 420 clusters for sequences with prokaryotic signatures failed to identify any new candidates, perhaps because the dataset is of moderate size or because migratory endoparasitic nematodes contain fewer such sequences than their sedentary endoparasitic counterparts.

Discussion

P. penetrans is an important plant parasite worldwide, having a very wide host range. However, most studies of plant-parasitic nematodes focus on root-knot and cyst nematodes (*Meloidogyne* spp., *Heterodera* spp. and *Globodera* spp.). As recently as July 2002, NCBI's GenBank contained only 12 nucleotide entries for *P. penetrans*, and just 85 entries for the complete genus, representing just five loci: for small and large subunit ribosomal RNAs (Al-Banna et al. 1997), a major sperm protein, a heat shock protein, and a β-1,4-endoglucanase (Uehara et al. 2001). We have generated and analyzed 1928 ESTs for *P. penetrans*, significantly increasing the available sequence information for this nematode.

Strikingly, there was almost no overlap between the most abundant transcripts found in the *P. penetrans* mixed-stage SL1 library and the previously studied *M. incognita* J2 library (McCarter et al. 2003a). A homolog of the *C. elegans* neuropeptide-like protein NLP-21 was the only exception, with 36 ESTs in *M. incognita* (MI00775.cl) and 23 ESTs in *P. penetrans* (PP00259.cl). The two clusters shared 41% homology. In *C. elegans*, 32 *nlp* genes have been identified, which can be classified into at least 11 distinct families (Bargmann 1998; Nathoo et al. 2001); *nlp* genes have been described for other plant-parasitic nematode species, including *H. glycines*, *G. rostochiensis* , *G. pallida,* and *M. javanica* (Nathoo et al. 2001), but this is the first identification of a neuropeptide-like sequence in *Pratylenchus*.

Fully 6% (121) of all *P. penetrans* ESTs have homology with GTP cyclohydrolase I (1e-131). Previous studies showed that GTP cyclohydrolase I consists of four identical subunits, which in turn consist of two

identical polypeptides with one GTP-binding site each (Yim and Brown 1976; Katzenmeier et al. 1991). GTP cyclohydrolase I catalyzes the conversion of GTP into dihydroneopterin-3′-triphosphate, which is a precursor for tetrahydrofolate and tetrahydrobiopterin (H_4B) (Schramek et al. 2002). H_4B is an important cofactor for the catalytic action of monooxygenases, including nitric oxide synthases (Gross et al. 2000), glyceryl-ether monooxygenase (Taguchi and Armarego 1998) and various aromatic amino acids (Kaufman 1993), as well as neurotransmitter-releasing factor (reviewed in Kaufman 1993). H_4B has been shown to be a virulence factor and life-cycle regulator in *Leishmania major* (Cunningham et al. 2001) and hypothetically could be a virulence factor or life-cycle regulator in *P. penetrans*. It would be interesting to clarify whether this transcript is abundant throughout all stages or whether it is highly expressed in some stages but not in others.

The presence of 5′ *trans* -spliced leader sequences in nematodes allows the use of an SL-PCR strategy to clone full-length cDNAs from very small amounts of RNA. The SL1 sequence is widely conserved across the phylum and can be used to amplify transcripts from most species including *P. penetrans*. While the SL1-PCR cDNA library approach has clear advantages, when compared to conventional cDNA libraries, clones were found to derive from shorter transcripts. Possible improvements to the protocol to allow the capture of longer transcripts could include testing of various reverse transcriptases with higher processivity, the use of longer PCR extension times, and a further increase in the stringency of electrophoretic size selection.

The generation and analysis of ESTs with semi-automated bioinformatics tools is an effective and valuable method for analyzing gene expression in parasitic nematodes (McCarter et al. 2000, 2003a; Mitreva et al. 2004). While this approach cannot provide an overview of the entire genome, it does allow a comparison of gene expression profiles among species, and is amenable to large-scale analyses. The rich knowledge about sedentary cyst and root-knot nematodes might be of limited value with regard to control strategies for *P. penetrans*, which is a migratory species and does not develop a long-term feeding site. Thus, the information gained here is also important for the potential development of new strategies to generate plants that are resistant to the root-lesion nematodes.

Acknowledgements Plant parasitic nematode genome research at Washington University was supported by NSF Plant Genome award 0077503 to D.M.B. and S.W.C. J.P.M. was supported by a Helen Hay Whitney/Merck postdoctoral fellowship. A.A.E. was supported by the Iowa Agriculture and Home Economics Experiment Station, Ames, IA and by Hatch Act and State of Iowa funds as well as by a Storkan-Hanes-McCaslin Foundation fellowship. We thank Merry McLaird, Brandi Chiapelli, Claire Murphy, and the GSC EST Group for their role in library generation and sequencing, Todd Wylie and John Martin for bioinformatics analysis, and Robert H. Waterston for project guidance. J.P.M. and A.P.K. are employees and equity holders of Divergence Inc.

References

Acedo JR, Rohde RA (1971) Histochemical root pathology of *Brassica oleracea capitata L.* infected by *Pratylenchus penetrans* (Cobb) Filipjev and Schuurmans Stekhoven (Nematoda: Tylenchidae). J Nematol 3:62–68

Al-Banna L, Williamson V, Gardner SL (1997) Phylogenetic analysis of nematodes of the genus *Pratylenchus* using 26S rDNA. Mol Phylogenet Evol 7:94–102

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Apweiler R, et al (2001) The InterPro database, an integrated documentation resource for protein families, domains and functional sites. Nucleic Acids Res 29:37–40

Audic S, Claverie JM (1997) The significance of digital gene expression profiles. Genome Res 7:986–995

Bargmann CI (1998) Neurobiology of the *Caenorhabditis elegans* genome. Science 282:2028–2033

Bird AF, Bird J (1991) The structure of nematodes (2nd edn) Academic Press, San Diego

Blaxter ML (2003) Nematoda: genes, genomes and the evolution of parasitism. Adv Parasitol 54:101–195

Blaxter M, Bird DM (1997) Parasitic nematodes. In: Riddle DL, Blumenthal T, Meyer B, Priess J (eds) *C. elegans* II. Cold Spring Harbor Press, Cold Spring Harbor, N.Y.

Blaxter M, Liu L (1996) Nematode spliced leaders – ubiquity, evolution and utility. Int J Parasitol 26:1025–1033

Blaxter ML, De Ley P, Garey JR, Liu LX, Scheldeman P, Vierstraete A, Vanfleteren JR, Mackey LY, Dorris M, Frisse LM, Vida JT, Thomas WK (1998) A molecular evolutionary framework for the phylum Nematoda. Nature 392:71–75

Blumenthal T, Steward K (1997) RNA processing and gene structure. In: Riddle DL, Blumenstiel B, Meyer BJ, Priess JR (eds) *C. elegans*. Vol. II. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 117–145

Bono H, Ogata H, Goto S, Kanehisa M (1998) Reconstruction of amino acid biosynthesis pathways from the complete genome sequence. Genome Res 8:203–210

Boucher G, Lambshead PJD (1994) Ecological biodiversity of marine nematodes in samples from temperate, tropical, and deep-sea regions. Conserv Biol 9:1594–1604

Conrad R, Thomas J, Spieth J, Blumenthal T (1991) Insertion of part of an intron into the 5′ untranslated region of a *Caenorhabditis elegans* gene converts it into a trans-spliced gene. Mol Cell Biol 11:1921–1926

Cunningham ML, Titus RG, Turco SJ, Beverley SM (2001) Regulation of differentiation to the infective stage of the protozoan parasite *Leishmania major* by tetrahydrobiopterin. Science 292:285–287

Dautova M, Overmars HA, Schots A, Gommers FJ, Bakker J, Smant G (2001a) A novel cell wall degrading enzyme in the plant parasitic nematode *Meloidogyne incognita*. Phytopathology 91:S21

Dautova M, Rosso MN, Abad P, Gommers FJ, Bakker J, Smant G (2001b) Single pass cDNA sequencing—a powerful tool to analyse gene expression in preparasitic juveniles of the Southern root-knot nematode *Meloidogyne incognita*. Nematology 3:129–139

De Meutter J, Tytgat T, van der Schueren E, Smant G, Schots A, Coomans A, van Montagu M, Gheysen G (1998) Cloning of two endoglucanase genes from *Heterodera schachtii*. Med Fac Landbouww Univ Gent 63/3a:619–623

Dickerson OJ, Darling HM, Griffin GD (1964) Pathogenicity and population trends of *Pratylenchus penetrans* on potato and corn. Phytopathology 54:317–322

Doyle EA, Lambert KN (2002) Cloning and characterization of an esophageal-gland-specific pectate lyase from the root-knot nematode *Meloidogyne javanica*. Mol Plant-Microbe Interacts 15:549–556

Elliott AP, Bird GW (1985) Pathogenicity of *Pratylenchus penetrans* to navy bean (*Phaseolus vulgaris L.*). J Nematol 17:81–85

Fraser AG, Kamath RS, Zipperlen P, Martinez-Campos M, Sohrmann M, Ahringer J (2000) Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. Nature 408:325–330

Goellner M, Smant G, de Boer JM, Baum TJ, Davis EL (2000) Isolation of beta-1,4-endoglucanase genes from *Globodera tabacum* and their expression during parasitism. J Nematol 32:154–165

Gonczy P, et al (2000) Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. Nature 408:331–336

Gross SS, Jones CL, Hattori Y, Raman CS (2000) Tetrahydrobiopterin: an essential cofactor of nitric oxide synthase with an elusive role. In: Ignarro LJ (ed) Nitric oxide: biology and pathobiology. Academic Press, San Diego, pp 167–185

Hoekstra R, Visser A, Otsen M, Tibben J, Lenstra JA, Roos MH (2000) EST sequencing of the parasitic nematode *Haemonchus contortus* suggests a shift in gene expression during transition to the parasitic stages. Mol Biochem Parasitol 110:53–68

Hung CL, Rohde RA (1973) Phenol accumulation related to resistance in tomato to infection by root-knot and lesion nematodes. J Nematol 5:253–258

IUBMB (1992) Enzyme Nomenclature: Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology. Academic Press, San Diego

Jaubert S, Laffaire JB, Abad P, Rosso MN (2002) A polygalacturonase of animal origin isolated from the root-knot nematode *Meloidogyne incognita*. FEBS Lett 522:109–112

Javed AA, Ogata K, Sanadi DR (1991) Human mitochondrial ATP synthase: cloning cDNA for the nuclear-encoded precursor of coupling factor 6. Gene 97:307–310

Kamath RS, Fraser AG, Dong Y, Poulin G, Durbin R, Gotta M, Kanapin A, Le Bot N, Moreno S, Sohrmann M, Welchman DP, Zipperlen P, Ahringer J (2003) Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. Nature 421:231–237

Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 28:27–30

Katzenmeier G, Schmid C, Kellermann J, Lottspeich F, Bacher A (1991) Biosynthesis of tetrahydrofolate. Sequence of GTP cyclohydrolase I from *Escherichia coli*. Biol Chem Hoppe Seyler 372:991–997

Kaufman S (1993) New tetrahydrobiopterin-dependent systems. Annu Rev Nutr 13:261–286

Koenning SR, Overstreet C, Noling JW, Donald PA, Becker JO, Fortnum BA (1999) Survey of crop losses in response to phytoparasitic nematodes in the United States for 1994. J Nematol 31:587–618

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

Krusberg LR (1960) Hydrolytic and respiratory enzymes of species of *Ditylenchus* and *Pratylenchus*. Phytopathology 50:9–22

Kurppa S, Vrain TC (1985) Penetration and feeding behavior of *Pratylenchus penetrans* in strawberry roots. Revue de Nematologie 8:273–276

Maeda I, Kohara Y, Yamamoto M, Sugimoto A (2001) Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi. Curr Biol 11:171–176

McCarter J, Abad P, Jones JT, Bird D (2000) Rapid gene discovery in plant parasitic nematodes via Expressed Sequence Tags. Nematology 2:719–731

McCarter J, Dautova Mitreva M, Martin J, Dante M, Wylie T, Rao U, Pape D, Bowers Y, Theising B, Murphy CV, Kloek AP, Chiapelli BJ, Clifton SW, Bird MD, Waterston R (2003a) Analysis and functional classification of transcripts from the nematode *Meloidogyne incognita*. Genome Biol 4:R26

McCarter JP, Mitreva M, Clifton SW, Bird DM, Waterston R (2003b) Nematode gene sequences: update for December 2003. J Nematol 35:465–469

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

Mountain WB, Patrick ZA (1959) The peach replant problem in Ontario. Can J Bot 37:459–470

Nathoo AN, Moeller RA, Westlund BA, Hart AC (2001) Identification of neuropeptide-like protein gene families in *Caenorhabditis elegans* and other species. Proc Natl Acad Sci USA 98:14000–14005

Nickle WR (1991) Manual of agricultural nematology. Marcel Dekker, New York

Nilsen TW (1993) *Trans* -splicing of nematode premessenger RNA. Annu Rev Microbiol 47:413–440

Parkinson J, Mitreva M, Hall N, Blaxter M, McCarter JP (2003) 400000 nematode ESTs on the Net. Trends Parasitol 4922:132–136

Pitcher RS, Patrick ZA, Mountain WB (1960) Studies on the host-parasite relations of *Pratylenchus penetrans* (Cobb) to apple seedlings. Nematologica 5:309–314

Popeijus H, Blok VC, Cardle L, Bakker E, Phillips MS, Helder J, Smant G, Jones JT (2000a) Analysis of genes expressed in second stage juveniles of the potato cyst nematodes *Globodera rostochiensis* and *G. pallida* using the expressed sequence tag approach. Nematology 2:567–574

Popeijus H, Overmars HA, Jones JT, Blok VC, Goverse A, Helder J, Schots A, Bakker J, Smant G (2000b) Degradation of plant cell walls by a nematode. Nature 406:36–37

Rosso MN, Favery B, Piotte C, Arthaud L, De Boer JM, Hussey RS, Bakker J, Baum TJ, Abad P (1999) Isolation of a cDNA encoding a beta-1,4-endoglucanase in the root-knot nematode *Meloidogyne incognita* and expression analysis during plant parasitism. Mol Plant Microbe Interact 12:585–591

Scholl EH, Thorne JL, McCarter JP, Bird DM (2003) Horizontally transferred genes in plant-parasitic nematodes: a high-throughput genomic approach. Genome Biol 4:R39

Schramek N, Bracher A, Fischer M, Auerbach G, Nar H, Huber R, Bacher A (2002) Reaction mechanism of GTP cyclohydrolase I: single turnover experiments using a kinetically competent reaction intermediate. J Mol Biol 316:829–837

Skalka AM, Goff SP (1993) Reverse transcriptase. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Smant G, Stokkermans JP, Yan Y, de Boer JM, Baum TJ, Wang X, Hussey RS, Gommers FJ, Henrissat B, Davis EL, Helder J, Schots A, 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 95:4906–4911.

Snedecor WG, Cochran GW (1967) Statistical methods. The Iowa State University Press, Ames, Iowa

Taguchi H, Armarego WL (1998) Glyceryl-ether monooxygenase [EC 1.14.16.5]. A microsomal enzyme of ether lipid metabolism. Med Res Rev 18:43–89

The Gene Ontology Consortium (2000) Gene ontology: a tool for the unification of biology. Nat Genet 25:25–29

Townshend JL (1963a) The pathogenicity of *Pratylenchus penetrans* to celery. Can J Plant Sci 43:70–74

Townshend JL (1963b) The pathogenicity of *Pratylenchus penetrans* to strawberry. Can J Plant Sci 43:75–78

Townshend JL (1978) Infectivity of *Pratylenchus penetrans* on alfalfa. J Nematol 10:318–323

Townshend JL, Stobbs L (1981) Histopathology and histochemistry of lesions caused by *Pratylenchus penetrans* in roots of forage legumes. Can J Plant Pathol 3:123–128

Townshend JL, Stobbs L, Carter R (1989) Ultrastructural pathology of cells affected by *Pratylenchus penetrans* in alfalfa roots. J Nematol 21:530–539

Troll J, Rohde RA (1966) Pathogenicity of *Pratylenchus penetrans* and *Tylenchorhynchus claytoni* on turfgrasses. Phytopathology 56:995–998

Uehara T, Kushida A, Momota Y (2001) PCR-based cloning of two beta-1,4-endoglucanases from the root-lesion nematode *Pratylenchus penetrans*. Nematology 3:335–341

Wheeler DL, Church DM, Lash AE, Leipe DD, Madden TL, Pontius JU, Schuler GD, Schriml LM, Tatusova TA, Wagner L, Rapp BA (2001) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 29:11–16

Wylie T, Martin J, Dante M, Mitreva M, Clifton SW, Chinwalla A, Waterston RH, Wilson RK, McCarter JP (2003) Nematode.net: a tool for navigating sequences from parasitic and free-living nematodes. Nucleic Acids Res 32:D423-D426

Wyss U (1970) Parasitierungsvorgang und Pathogenität wandernder Wurzelnematoden an *Fragaria vesca* var. *semperflorens*. Nematologica 16:55–62

Yim JJ, Brown GM (1976) Characteristics of guanosine triphosphate cyclohydrolase I purified from *Escherichia coli*. J Biol Chem 251:5087–5094

Zdobnov EM, Apweiler R (2001) InterProScan—an integration platform for the signature-recognition methods in InterPro. Bioinformatics 17:847–848

Zunke U (1990a) Ectoparasitic feeding behaviour of the root lesion nematode, *Pratylenchus penetrans* , on root hairs of different host plants. Revue Nématol 13:331–337

Zunke U (1990b) Observations on the invasion and endoparasitic behavior of the root lesion nematode *Pratylenchus penetrans*. J Nematol 22:309–320