

Identification of Abundantly Expressed Novel and Conserved Genes from the Infective Larval Stage of *Toxocara canis* by an Expressed Sequence Tag Strategy

KEVIN K. A. TETTEH,^{1†} ALEX LOUKAS,^{1‡} CINDY TRIPP,² AND RICK M. MAIZELS^{1*}

Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, Scotland, United Kingdom,¹ and Heska Corporation, Fort Collins, Colorado 80525²

Received 2 March 1999/Returned for modification 13 May 1999/Accepted 1 June 1999

Larvae of *Toxocara canis*, a nematode parasite of dogs, infect humans, causing visceral and ocular larva migrans. In noncanid hosts, larvae neither grow nor differentiate but endure in a state of arrested development. Reasoning that parasite protein production is orientated to immune evasion, we undertook a random sequencing project from a larval cDNA library to characterize the most highly expressed transcripts. In all, 266 clones were sequenced, most from both 3' and 5' ends, and similarity searches against GenBank protein and dbEST nucleotide databases were conducted. Cluster analyses showed that 128 distinct gene products had been found, all but 3 of which represented newly identified genes. Ninety-five genes were represented by a single clone, but seven transcripts were present at high frequencies, each composing >2% of all clones sequenced. These high-abundance transcripts include a mucin and a C-type lectin, which are both major excretory-secretory antigens released by parasites. Four highly expressed novel gene transcripts, termed *ant* (abundant novel transcript) genes, were found. Together, these four genes comprised 18% of all cDNA clones isolated, but no similar sequences occur in the *Caenorhabditis elegans* genome. While the coding regions of the four genes are dissimilar, their 3' untranslated tracts have significant homology in nucleotide sequence. The discovery of these abundant, parasite-specific genes of newly identified lectins and mucins, as well as a range of conserved and novel proteins, provides defined candidates for future analysis of the molecular basis of immune evasion by *T. canis*.

Toxocara canis is a common nematode parasite of dogs and related species. Adult worms live in the gastrointestinal tract, releasing eggs which enter the environment by the fecal route. Within the eggs, larval *T. canis* develop over a 2- to 3-week period (27). Embryonated eggs are then infective if ingested by a new host, as larvae hatch in the stomach and penetrate the epithelial layer. *T. canis* larvae show a remarkable lack of host specificity, infecting a wide range of taxa, including humans (27, 28). In the definitive (canid) host, larvae may mature by migrating to the intestine and developing to the adult stage; such maturation is favored in pups and in reproducing females (40). In most other dogs and in all paratenic hosts such as humans, development is arrested at the larval stage.

The arrested stage is remarkable for surviving for as long as 9 years in vivo (7), without reproduction or differentiation and without succumbing to attack by the host immune system. This diapausal state is mirrored in vitro, where larvae survive for many months in serum-free medium, secreting quantities of excretory-secretory antigens which have been characterized in biochemical (5, 50, 58) and immunological (44, 45, 56) terms.

We hypothesized that in the absence of cell division, tissue growth, or gametogenesis, a significant proportion of larval protein production, and therefore mRNA, is likely to be directed at immune evasion. To characterize abundant mRNAs, we conducted a small-scale expressed sequence tag (EST)

project. EST sequencing was pioneered for mammalian cells (1, 2) and *Caenorhabditis elegans* (49) and is an important component of major parasite sequencing projects (15, 18, 22, 23, 46, 64). By this means, we have identified a series of abundantly expressed genes from larval *T. canis*, among which are likely to be important mediators of parasite immune evasion.

MATERIALS AND METHODS

cDNA library. *T. canis* larvae were hatched and maintained in vitro as previously described (20, 50) for a period of 6 months, with weekly changes of serum-free medium. From 200,000 cultured *T. canis* larvae, using a single-step guanidine-phenol-chloroform extraction, 265 µg of total RNA was recovered, from which 6 µg of poly(A)⁺ RNA was isolated by oligo(dT) chromatography. cDNA synthesized from this mRNA was unidirectionally cloned into the Uni-Zap XR phage vector, using packaging extracts from Stratagene. The amplified library contained 1.9×10^9 PFU/ml with 91% recombinants. The possibility of host contamination was essentially eliminated because eggs were first incubated in vitro in formalin, and once hatched, larvae were cultured in serum-free RPMI 1640 medium.

Isolation of cDNA clones and in vivo excision of phagemids. Single clones were randomly picked from the plated out cDNA library. Phagemids were rescued in vivo by using ExAssist helper phage and nonsuppressing *Escherichia coli* SOLR (Stratagene).

Selection and maintenance of clones. Plasmids were prepared from overnight cultures by using a Qiaprep Spin Miniprep kit (Qiagen) and stored at -20°C. Insert sizes were determined by PCR using vector primers (M13 reverse and M13 forward primers or T3 and T7 primers). In the few cases where insert sizes could not be determined by PCR, restriction digestion of purified recombinant plasmids was performed with *Xba*I and *Xho*I (Promega). All clones are available to the research community on request.

Sequencing. Plasmid templates were sequenced by using dye terminator cycle sequencing chemistry with Amplitaq DNA polymerase (FS enzyme) on an ABI 377 automated sequencer (Applied Biosystems, Foster City, Calif.). Both 5' and 3' ends were sequenced in some cases where 5' sequences gave high-quality sequence through the poly(A) tail or where the 5' sequence showed unequivocal identity with a previously characterized *T. canis* clone.

Analysis. SeqEd version 1.0.3 (Applied Biosystems) was used to edit out vector sequence and flanking sequences. Sequences were aligned by using AssemblyLign and MacVector 6.0 software (Oxford Molecular). Nonredundant database searches used ungappped BLAST (basic local alignment search tool) (3) on

* Corresponding author. Mailing address: ICAPB, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, Scotland, United Kingdom. Phone: (44) 131 650 5511. Fax: (44) 131 650 5450. E-mail: r.maizels@ed.ac.uk.

† Present address: London School of Hygiene and Tropical Medicine, London WC1E 7HT, United Kingdom.

‡ Present address: Molecular Parasitology Unit, Queensland Institute of Medical Research, Queensland 4029, Australia.

the National Center for Biotechnology Information server. Nucleotide sequences were subjected to BLASTX searches against the GenBank nr (nonredundant) protein database. Nucleotide sequence searches used BLASTN (on both nr and dbEST databases), and deduced protein sequence from any unambiguous open reading frame was used to search with TBLASTN (against the nr nucleotide database).

Nomenclature. Gene naming follows the convention for nematodes (11, 14) of a three-letter name and a number, and a prefix indicating the source organism, in this case *Tc*. Genes are denoted in italics; proteins are capitalized. Three sets of interim gene names were used where no functional homology exists: *Tc-not* (novel transcript), *Tc-ant* (abundant novel transcript, where ≥ 5 clones of 263 are identical), and *Tc-huf* (homolog of unknown function, where similar sequences have been found in other nematode species). For interim gene designations, the number of the EST clone first sequenced is retained; for cDNAs assigned functional names, clones are generally numbered sequentially (e.g., *Tc-ctl-1* and *-2* for C-type lectins), except where numbering is significant in other organisms, principally for ribosomal proteins (e.g., *Tc-rps-4*, *-5*, and *-9* to conform with established nomenclature).

Database deposition. All sequences have been deposited in NCBI dbEST (17) with separate entries for 5' and 3' reads.

RESULTS AND DISCUSSION

Sequencing and similarity searches. A total of 261 clones, containing inserts ranging from 128 to 2,050 bp, were taken for analysis once nonrecombinant and chimeric clones were discarded. All were sequenced from the 5' end, and 218 were also sequenced from the 3' end. Most similarities were found, or were stronger, with the 5' sequences, but a significant minority of similarity relationships were found only with 3' sequence reads. In general, a probability value of 10^{-6} was sought as a minimum degree of similarity.

Clustering analysis. EST sequences were clustered on the basis of nucleotide identity. It was noted that in some of the larger clusters, identity of the 3' sequences was critical to correctly group differentially truncated clones with nonoverlapping 5' sequences.

As a result of these analyses, a total of 128 distinct gene products were identified. Of these, only three, mucin 1 (*Tc-muc-1* [25]), phosphatidylethanolamine-binding protein (*Tc-peb-1* [26]), and the 60S ribosomal protein L3 (*Tc-rpl-3* [51]), have been previously characterized. Approximately 50% (65/128) of the total number of genes have informative similarity to genes of known function from other species, a further 17 clones have database homologs of unknown function, and 47 genes (37%) show no similarity to known genes; among this last class, designated novel genes, 4 were classified as abundant transcripts (see below).

Abundant transcripts. A small number of transcripts are heavily represented in the library. Just 8 transcripts (6.2% of genes) account for 102 clones (39.1%) sequenced, while the 20 most abundant (all those isolated three or more times) accounted for 143 clones (54.8%). Table 1 presents the transcripts characterized in order of frequency, with the most highly sampled clone being cytochrome *c* oxidase subunit II ($25/261 = 9.6\%$ of clones), which is a mitochondrial DNA-encoded gene in other nematodes (36, 52). The second most common clone is a C-type lectin ($16/261 = 6.1\%$) which in work to be published elsewhere we demonstrate encodes the major surface and secreted glycoprotein of *T. canis* larvae, TES-32 (42).

Abundant novel transcripts. Unexpectedly, four more abundant clones are all novel, with no similarities in the database or, in the coding regions, to each other. These transcripts each represent between 2.7 and 5.7% of ESTs and together account for more than 18% of the library. They have consequently been designated *ant* genes and retain the number of the EST clone from which they were first isolated (*ant-003*, *-005*, *-030*, and *-034*).

The four *ant* genes differ in length and composition, but all

TABLE 1. Frequency of transcripts

Transcript (gene)	No. of clones sequenced	Frequency of transcripts (%)
Cytochrome <i>c</i> oxidase subunit II (<i>cox-2</i>)	25	9.6
C-type lectin, TES-32 (<i>ctl-1</i>)	16	6.1
Abundant novel transcript 003 (<i>ant-003</i>)	15	5.7
Abundant novel transcript 034 (<i>ant-034</i>)	13	5.0
Abundant novel transcript 030 (<i>ant-030</i>)	13	5.0
Abundant novel transcript 005 (<i>ant-005</i>)	7	2.7
Mucin 1, TES-120 (<i>muc-1</i>)	7	2.7
Phosphatidylethanolamine-binding protein, TES-26 (<i>peb-1</i>)	5	1.9
Novel transcript 095 (<i>not-095</i>)	4	1.5
Protein tyrosine phosphatase (<i>ptp-1</i>)	4	1.5
60S ribosomal protein L19 (<i>rpl-19</i>)	4	1.5
ADP/ATP translocase (<i>aat-1</i>)	3	1.1
ASP homolog 1/TC-CRISP (<i>vah-1</i>)	3	1.1
Mucin 2 (<i>muc-2</i>)	3	1.1
Novel transcript 018 (<i>not-018</i>)	3	1.1
Novel transcript 120 (<i>not-120</i>)	3	1.1
60S ribosomal protein P0 (<i>rpp-0</i>)	3	1.1
Superoxide dismutase 3 (<i>sod-3</i>)	3	1.1
Tubulin alpha-3 chain (<i>tua-3</i>)	3	1.1
14 transcripts present as two clones ^a	28	10.7
95 transcripts present as single clones	96	36.7
Total no. of clones sequenced	261	

^a The 14 genes present as two clones are *aki-1*, *ash-2*, *cap-1*, *huf-001*, *irp-1*, *muc-3*, *not-069*, *not-216*, *ofm-1*, *pam-1*, *pcc-1*, *rpl-7a*, *rpl-9*, and *sod-4*. Details for these, apart from *not* genes, are given in Table 3; details for *ash-2* and *huf-001* are in Table 2.

are 1.6 kb or more in length. The characterization of their full-length sequences, and of the protein products encoded by these genes, is currently under way, as none of the ESTs isolated include 5' methionine start codons. Figure 1 presents a map of the EST clones isolated for each *ant* gene. From this it can be seen that 3' sequencing proved essential in identifying all members of the cluster, because 5' sequences do not in all cases overlap. Because the multiple clones all have different 5' termini, the abundance of these transcripts appears not to be an artifactual amplification of single clones in the construction of the library.

Homology in 3' UTRs of the four *ant* genes. Although *ant-003*, *-005*, *-030*, and *-034* showed no similarities between coding regions, the 3' untranslated regions (UTRs) of all four genes bore significant levels of identity. It is notable also that none of these 3' ends contain consensus polyadenylation motifs such as AATAAA or similar sequence. This is not unprecedented among nematode genes: a recent survey of *C. elegans* 3' UTRs found that 7% of mRNAs bore no identifiable polyadenylation signal (16).

An alignment of the 3' coding regions and UTRs of these four transcripts shows little identity in coding region nucleotides (Fig. 2A) or amino acids (not shown) but similar sequences in all four genes immediately after the stop codon (Fig. 2B). No similar sequences were found in any other *T. canis* genes or in genes from other organisms such as *C. elegans*. In *C. elegans*, there are examples of 3' UTR motifs associated with repression of translation in genes such as *tra-2* and *fem-3* in germ line cell differentiation (4, 68). This is unlikely to be a useful parallel for the *T. canis ant* genes, as there is no similarity in 3' UTR sequence between the two species and there is no gonadal development in the larval stage of *T. canis*. Translational suppression of the *ant* genes would be a surprising event in light of their unusually high level of transcription.

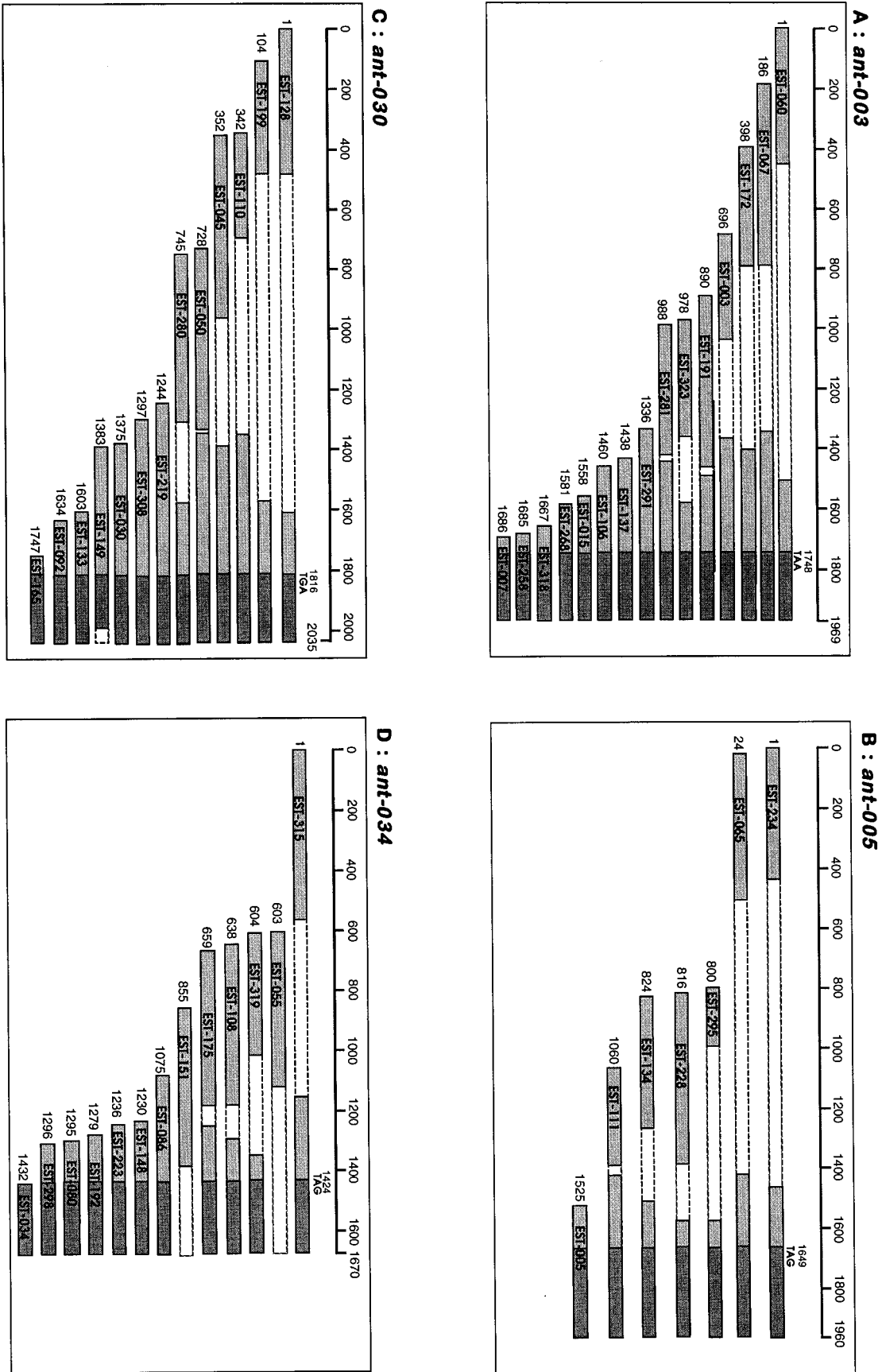
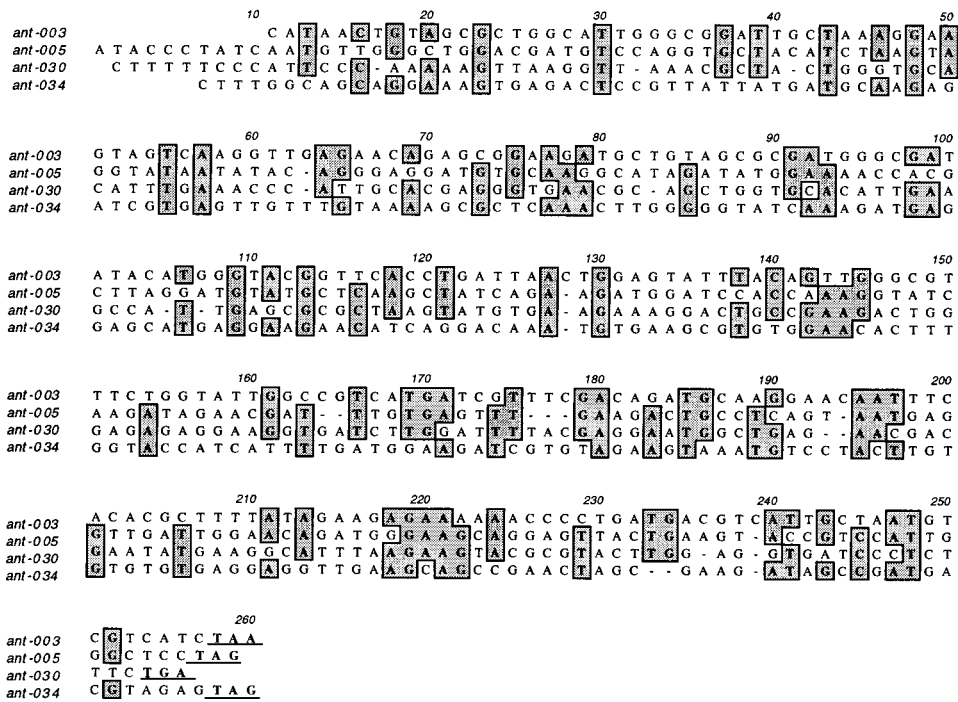


FIG. 1. Alignments of clones corresponding to *ant-003*, *-005*, *-030*, and *-034*. Segments sequenced in single-pass reactions are lightly shaded; unsequenced portions are shown unshaded in broken lines. Darker shading corresponds to 3' UTRs. The position and codon of the stop signal are indicated for each transcript. Numbers to the left of the bars indicate the approximate start of the cDNA relative to the longest clone. No start codons have yet been identified for any of these genes, and hence numbering is provisional. Note that without 3' sequence data, the longer EST clones belonging to *ant-005* and *ant-034* would not have been recognized.

A. 3' coding sequence (250 nt per transcript)



B. 3' untranslated sequence (220 - 281 nt per transcript)

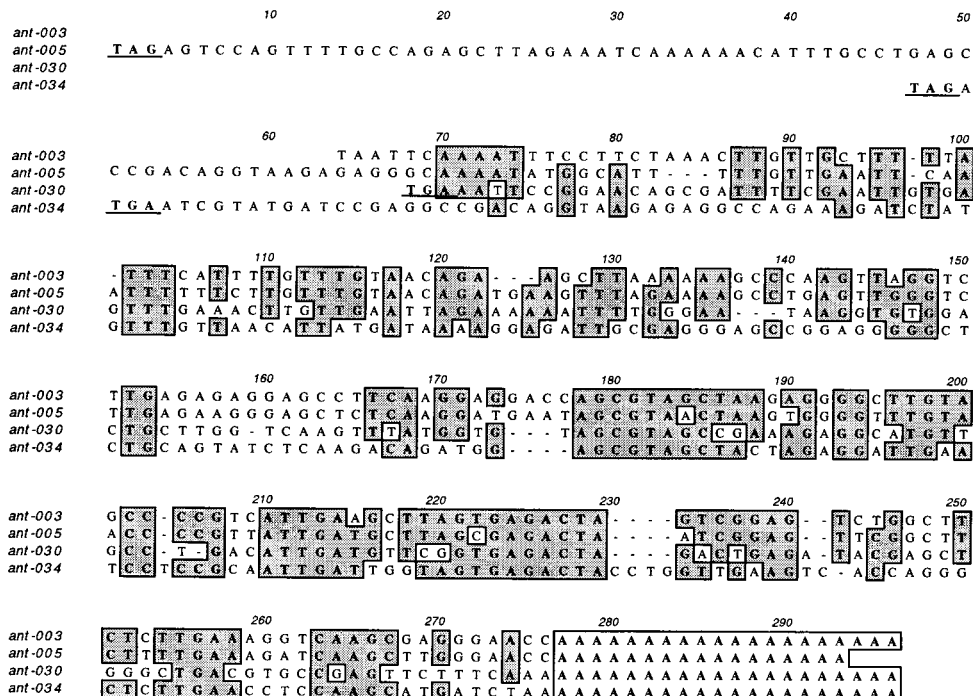


FIG. 2. Alignment of 250-nucleotide (nt) 3'-terminal coding sequences (A) and complete 3' UTRs (B) of the four *ant* genes; identical settings were used for both alignments in MacVector ClustalW alignment (gap penalty = 10.0, extend gap penalty = 5.0). The stop codons at the end of the coding sequence and the beginning of the UTR are underlined. Bases identical in three or four of the sequences are boxed and shaded. The polyadenylated tail is shown boxed without shading.

Homologs of genes of unknown function. Sixteen clones showed significant similarities to known nematode sequences with no assigned function (Table 2). These were all designated *huf* genes, retaining the number of the corresponding EST

clone. One clone, *Tc-huf-001*, contains a tandem array of four blocks of 36 amino acids each containing six cysteines in identical alignment. This motif, termed the NC6 (26) or SXC (12) domain, is found in *T. canis* and *C. elegans* proteins, particu-

TABLE 2. Homologs to proteins of unknown function in other nematodes^a

Gene	Similarity to:	P	EST name	Insert size	Read		GenBank accession no.	
					5'	3'	5'	3'
<i>Tc-huf-001</i>	NC6/SXC domain		Tc-EST-001	800	764	666	AA683457	AA683458
<i>Tc-huf-012</i>	<i>C. elegans</i> R53.5 gene	3.8E-06	Tc-EST-012	850	599	342	AA563573	AI374481
<i>Tc-huf-053</i>	<i>C. elegans</i> C13A10 gene	2.8E-38	Tc-EST-053	1,000	589	602	AA610971	AA610972
<i>Tc-huf-062</i>	<i>C. elegans</i> F41D9.1 gene	4.3E-42	Tc-EST-062	760	482	521	AA612566	AA907959
<i>Tc-huf-159</i>	<i>C. elegans</i> C56E6.5 gene	7.6E-15	Tc-EST-159	1,850	387	288	AA836705	AA836706
<i>Tc-huf-169</i>	<i>C. elegans</i> K02E10.6 gene	6.2E-17	Tc-EST-169	1,050	466	393	AA874757	AA874758
<i>Tc-huf-252</i>	<i>C. elegans</i> Y57G11C.12 gene	7.7E-30	Tc-EST-252	1,150	547	558	AA684533	AA684534
<i>Tc-huf-264</i>	<i>C. elegans</i> B0034.3 gene	1.4E-25	Tc-EST-264	1,420	589	444	AA684531	AA684532
<i>Tc-huf-287</i>	<i>C. elegans</i> K12H4.5 gene	5.3E-23	Tc-EST-287	452	429	416	AA875794	AA979711
<i>Tc-huf-296</i>	<i>C. elegans</i> T24F1.6 gene	1.9E-11	Tc-EST-296	2,300	259	491	AA875806	AA875807
<i>Tc-huf-302</i>	<i>B. malayi</i> MBAFCE2B09T3	2.3E-29	Tc-EST-302	760	283	621	AA879342	AA879343
<i>Tc-huf-316</i>	<i>C. elegans</i> T24D8.5 gene	3.7E-21	Tc-EST-316	720	542	578	AA874711	AA874712
<i>Tc-huf-325</i>	<i>O. volvulus</i> L3CAN05C08SK	7.4E-47	Tc-EST-325	1,250	370	351	AA879368	AA879369
<i>Tc-vah-1</i>	<i>C. elegans</i> T05A10.5 venom allergen/ASP homolog	4.6E-15	Tc-EST-269	1,870	459	590	AI 083044	AI 083045
<i>Tc-vah-2</i>	<i>B. malayi</i> venom allergen/ASP homolog	1.0E-13	Tc-EST-303	1,600	550	699	AI083048	AI083049
<i>Tc-vah-3</i>	<i>B. malayi</i> venom allergen/ASP homolog	4.7E-16	Tc-EST-249	650	507	512	AI083050	AI083051

^a All cDNAs were isolated as single transcripts except *vah-1* (three clones), *vah-2* (two clones), and *huf-001* (two clones).

larly those associated with nematode surfaces, but the function of the array is not known.

In addition, we found three genes which show similarity to *Ancylostoma* secreted protein (ASP), which is associated with larval activation and development in other nematode species (10, 30, 31, 61). However, the *T. canis* stage from which the cDNA library was made is developmentally arrested and is not analogous to activated hookworm larvae. This gene family shows similarity to hymenopteran venom allergens and, in common with members reported for *Brugia malayi* and *Onchocerca volvulus*, has been designated *vah* (venom allergen homolog). *Tc-vah-1*, previously reported as Tc-CRISP (cysteine-rich secreted protein) by Liu (39a), is remarkable for containing two NC6/SXC domains in tandem with a VAH homology unit. Further characterization of these clones, *Tc-vah-1*, -2, and -3 is in progress (51a).

Other novel genes. A total of 43 additional transcripts were found to have no significant similarities to any other sequences deposited in GenBank nr protein and dbEST databases. These were present at between one and four copies in the 261-member data set from *T. canis* and are designated *not* genes. Further studies on selected clones, such as *not-018*, for which antibodies to the protein product strongly recognize *T. canis* larval excretory-secretory antigens (61a) are ongoing.

Homologs of known genes. Sixty-five genes with database homologs were found; these are listed in Table 3. There are 23 metabolic and respiratory enzymes but remarkably few structural proteins (only actin, calponin, and α -tubulin) and no DNA replication proteins, consistent with the arrested state of the larval parasite. The various categories of genes are described below.

Mucins. A mucin gene, *Tc-muc-1*, has previously been reported to be abundantly expressed by *T. canis* larvae (25). Consistent with this, seven clones of *Tc-muc-1* were recorded (2.7% of the library). Three new mucin genes were found among the EST clones, each of which contains similar mucin domains and flanking NC6/SXC domains (26). The mucins differ in the composition of repeat motifs, and in the number and positions of NC6/SXC domains, and work in progress indicates that all are members of the TES-120 glycoprotein family associated with the parasite surface coat (40a).

C-type lectins. One of the most abundant transcripts (16/261) was found to correspond to peptide sequence obtained from TES-32, a prominent secreted glycoprotein. These ESTs

showed weak similarity to C-type lectins, but the homology was firmly established from full-length sequence. A detailed analysis of the functional lectin properties of TES-32/*Tc-ctl-1* has been submitted for publication (42). Two variants of this sequence, designated *Tc-ctl-2* and *Tc-ctl-3*, were also noted.

Proteases. Proteolytic enzymes have been prominent in most studies of parasitic helminths at the biochemical (32) and molecular (43) levels. Three transcripts, with strong similarities to cathepsin L (41), cathepsin Z (43), and asparaginyl endopeptidase (19), were each present as single copies. Full-length sequences of the cathepsins L (41) and Z (22a) have been determined. A protease inhibitor similar to the aspartyl protease inhibitor of *Ascaris* and *Brugia* (Bm33) (21) was also isolated as a single clone.

Transporters and receptors. One clone homologous to the acetylcholine receptors reported from other nematodes (24) has been isolated. A relatively frequent transcript (5/263) encodes *Tc-PEB-1*, a previously identified phosphatidylethanolamine-binding protein (26) which is present in parasite secretions as a 26-kDa protein (TES-26).

Structural proteins. Few structural proteins were identified in the EST data set, probably reflecting the arrested state of this parasite stage. Actin (67) and calponin (33) have been sequenced from other parasitic nematodes. For tubulins, most attention has focused on the β -tubulins (29), with which benzimidazole resistance is associated. *Tc-tua-1* has strong similarity to α -tubulins from *Haemonchus contortus* (37) and *C. elegans*.

Protein synthesis and modification. Two genes essential for protein synthesis (elongation factors 1a and 1b) were identified, as well as peptidyl-glycine alpha-hydroxylating monooxygenase, which modifies the C termini of peptides. Protein disulfide isomerase is a well-known requisite for protein folding and correct formation of disulfide bonds. In *C. elegans*, the *pdi* gene is found in an operon with one cyclophilin gene (55); a homolog from *O. volvulus* has also been characterized (65). Peptidyl-prolyl *cis-trans*-isomerase is similarly essential for correct protein conformation; the *Tc-ppi-1* gene product is related to the FK506-binding proteins of mammals and not to the multigene cyclophilin family of peptidyl-prolyl *cis-trans* isomerases described for *C. elegans* (55, 57).

Glycolysis, respiration, and other metabolic and citric acid cycle enzymes. Some 18 distinct metabolic enzymes had very high levels of similarity to *T. canis* ESTs and represent the

TABLE 3. Homology of *T. canis* ESTs with known genes^a

Gene	Gene product	Closest species	<i>P</i>	Clone no.	Insert size	Read		GenBank accession no.		Full-length verified sequence
						5'	3'	5'	3'	
Mucins										
<i>Tc-muc-1</i>	Mucin 1, surface coat TES-120	<i>T. canis</i>	6.3E-104	Tc-EST-087 ^b	810	557	539	AA583105	AA907963	U39815
<i>Tc-muc-2</i>	Mucin 2	<i>T. canis</i>	3.8E-49	Tc-EST-130 ^b	990	612	694	AA583106	AA583107	AF167707
<i>Tc-muc-3</i>	Mucin 3	<i>T. canis</i>	2.6E-36	Tc-EST-162 ^b	1,010	603	555	AA583111	AA583112	AF167708
<i>Tc-muc-4</i>	Mucin 4	<i>T. canis</i>	1.6E-45	Tc-EST-186	560	553	586	AA583108	AA583109	AF167709
Lectins										
<i>Tc-ctl-1</i>	C-type lectin/TES-32 glyco-protein			Tc-EST-209	830	615	393	AA874713	AA874714	AF041023
<i>Tc-ctl-2</i>	C-type lectin-2/variant of TES-32			Tc-EST-036	780	596	618	AI078880	AI078881	
<i>Tc-ctl-3</i>	C-type lectin-3/variant of TES-32			Tc-EST-155	610	590	737	AI080917	AI080918	
Proteases and protease inhibitors										
<i>Tc-aep-1</i>	Asparaginyl endopeptidase	<i>Homo sapiens</i>	3.5E-28	Tc-EST-262	1,300	520	648	AA825116	AA875769	
<i>Tc-api-1</i>	Aspartyl proteinase inhibitor	<i>C. elegans</i>	1.5E-09	Tc-EST-096	650	570	638	AA820019	AA820020	
<i>Tc-cpl-1</i>	Cathepsin L cysteine proteinase	<i>C. elegans</i>	2.3E-40	Tc-EST-013	1,150	604	390	AA563518	AI078878	U53172
<i>Tc-cpz-1</i>	Cathepsin Z proteinase precursor	<i>O. volvulus</i>	1.2E-73	Tc-EST-205	1,000	568	701	AI080921	AI080922	AF143817
Transporters and receptors										
<i>Tc-aat-1</i>	ADP/ATP translocase	<i>C. elegans</i>	1.9E-69	Tc-EST-041 ^b	750	515	663	AA603955	AA683466	
<i>Tc-acr-1</i>	Acetylcholine receptor (<i>tar-1</i>)	<i>Trichostrongylus colubriformis</i>	1.2E-35	Tc-EST-121	1,300	529	547	AA835586	AA835587	
<i>Tc-peb-1</i>	Phosphatidylethanolamine-binding protein (TES-26)	<i>T. canis</i>	6.1E-76	Tc-EST-164 ^b	1,200	406	459	AA864177	AA864178	U29761
Structural proteins										
<i>Tc-act-1</i>	Actin	<i>C. elegans</i>	1.3E-81	Tc-EST-097	1,450	524		AA820021		
<i>Tc-cap-1</i>	Calponin	<i>Meloidogyne javanica</i>	4.2E-21	Tc-EST-292 ^b	620	462	537	AA875799	AA907961	
<i>Tc-tua-1</i>	Tubulin alpha-3 chain	<i>Cricetulus griseus</i>	1.5E-69	Tc-EST-297 ^b	1,500	400	439	AA979712	AA907952	
Protein synthesis and modification factors										
<i>Tc-efa-1</i>	Elongation factor 1a	<i>C. elegans</i>	3.5E-84	Tc-EST-300	1,600	498	572	AA879374	AA879375	
<i>Tc-efb-1</i>	Elongation factor 1b	<i>C. elegans</i>	3.1E-42	Tc-EST-081	830	355	756	AA888759	AA888760	
<i>Tc-pam-1</i>	Peptidyl-glycine alpha-hydroxylating monooxygenase	<i>Drosophila melanogaster</i>	1.1E-07	Tc-EST-200 ^b	420	385	357	AA874705	AI080786	
<i>Tc-pdi-1</i>	Protein disulfide isomerase	<i>Gallus gallus</i>	5.9E-18	Tc-EST-235	474	445	468	AA874740	AA874741	
<i>Tc-ppi-1</i>	Peptidyl-prolyl <i>cis-trans</i> isomerase	<i>Mus musculus</i>	3.2E-15	Tc-EST-019	1,200	319		AA563572		
<i>Tc-ptp-1</i>	Protein tyrosine phosphatase	<i>M. musculus</i>	5.9E-49	Tc-EST-118	1,590	589	628	AA820013	AA820014	
Glycolysis, respiration, and other metabolic enzymes										
<i>Tc-aca-1</i>	Acetyl coenzyme A acetyltransferase	<i>C. elegans</i>	9.8E-33	Tc-EST-059	626	626	560	AA610965	AA907958	
<i>Tc-aki-1</i>	Adenylate kinase isoenzyme I	<i>G. gallus</i>	9.2E-39	Tc-EST-328	800	452	505	AA873889	AA873920	
<i>Tc-ald-1</i>	Aldehyde dehydrogenase	<i>Bos bovis</i>	1.3E-45	Tc-EST-299	1,700	559	360	AA979715	AA979716	
<i>Tc-ata-1</i>	Alanine aminotransferase	<i>C. elegans</i>	4.0E-42	Tc-EST-294	1,440	472	357	AA875802	AA875803	
<i>Tc-cgl-1</i>	Cystathionine gamma lyase	<i>C. elegans</i>	6.4E-54	Tc-EST-090	750	555	685	AA583117	AA583118	
<i>Tc-cyb-5</i>	Cytochrome <i>b</i> ₅	<i>H. sapiens</i>	4.5E-10	Tc-EST-032	406	406		AA569419		
<i>Tc-fba-1</i>	Fructose biphosphate aldolase	<i>O. volvulus</i>	2.1E-98	Tc-EST-183	1,010	480	533	AA915868	AA915869	
<i>Tc-fdh-1</i>	Fumarate dehydrogenase precursor	<i>H. sapiens</i>	1.7E-26	Tc-EST-237	317	317	314	AA285389	AA874746	
<i>Tc-frb-1</i>	Fumarate reductase, cytochrome <i>b</i> small subunit	<i>Ascaris suum</i>	2.9E-32	Tc-EST-083	1,560	452	480	AA888752	AA888753	
<i>Tc-gdh-1</i>	Glutamate dehydrogenase	<i>H. contortus</i>	4.7E-74	Tc-EST-026	1,740	584	331	AA568090	AI078879	
<i>Tc-pcc-1</i>	Propionyl coenzyme A carboxylase	<i>C. elegans</i>	8.6E-64	Tc-EST-074	1,560	320	399	AA888750	AA888751	
<i>Tc-pck-1</i>	Phosphoenolpyruvate carboxylase	<i>A. suum</i>	4.4E-69	Tc-EST-004	2,740	599	425	AA470327	AI374480	
<i>Tc-pgk-1</i>	Phosphoglycerate kinase	<i>Schistosoma mansoni</i>	1.8E-41	Tc-EST-109	1,400	576	805	AA825117	AA825118	
<i>Tc-sdi-1</i>	Succinate dehydrogenase iron-sulfur protein	<i>A. suum</i>	3.0E-73	Tc-EST-166	630	518	423	AA864181	AI374479	
<i>Tc-ubo-1</i>	NADH ubiquinone oxidoreductase 18-kDa subunit	<i>Bos taurus</i>	1.9E-39	Tc-EST-201	521	501	480	AA874708	AI080786	

Continued on following page

TABLE 3—Continued

Gene	Gene product	Closest species	<i>P</i>	Clone no.	Insert size	Read		GenBank accession no.		Full-length verified sequence	
						5'	3'	5'	3'		
Presumed mitochondrial proteins											
<i>Tc-atp-6</i>	ATP synthase A chain (protein 6)	<i>A. suum</i>	1.9E-68	Tc-EST-064	610	576	557	AA612565	AA979708		
<i>Tc-cox-2</i>	Cytochrome <i>c</i> oxidase II	<i>A. suum</i>	7.1E-75	Tc-EST-038	850	496	669	AA569422	AA569423		
<i>Tc-ubo-4</i>	NADH-ubiquinone oxidoreductase chain 4L	<i>A. suum</i>	2.6E-29	Tc-EST-163	220	213	216	AA864175	AA864176		
Antioxidants											
<i>Tc-sod-2</i>	Superoxide dismutase 2	<i>C. albicans</i>	1.5E-23	Tc-EST-084	850	401	340	AI078885	AI078886		
<i>Tc-sod-3</i>	Superoxide dismutase 3	<i>Zea mays</i>	2.1E-18	Tc-EST-100	900	449	608	AI080774	AI080775		
<i>Tc-sod-4</i>	Superoxide dismutase 4	<i>C. elegans</i>	3.6E-40	Tc-EST-143	804	630	520	AI080778	AI080779		
<i>Tc-sod-5</i>	Superoxide dismutase 5	<i>C. elegans</i>	2.4E-39	Tc-EST-266	900	731	730	AI080782	AI080783		
Ribosomal proteins											
<i>Tc-rpl-10</i>	60S ribosomal protein L10	<i>C. elegans</i>	8.6E-44	Tc-EST-271	800	321	590	AA875777	AA875778		
<i>Tc-rpl-19</i>	60S ribosomal protein L19	<i>M. musculus</i>	2.1E-60	Tc-EST-244 ^b	900	439	555	AA874749	AA874750	P49149	
<i>Tc-rpl-3</i>	60S ribosomal protein L3	<i>T. canis</i>	5.5E-60	Tc-EST-257	500	498	457	AA875762	AI080789		
<i>Tc-rpl-31</i>	60S ribosomal protein L31	<i>H. sapiens</i>	4.8E-49	Tc-EST-072	580	440	491	AA728649	AA728650		
<i>Tc-rpl-37a</i>	60S ribosomal protein L37	<i>S. mansoni</i>	3.6E-34	Tc-EST-267	400	372	381	AA875772	AA875773		
<i>Tc-rpl-7</i>	60S ribosomal protein L7A	<i>H. sapiens</i>	7.2E-29	Tc-EST-054 ^b	750	513	560	AA610961	AA907956		
<i>Tc-rpl-9</i>	60S ribosomal protein L9	<i>H. contortus</i>	1.1E-54	Tc-EST-182 ^b	630	396	480	AA873890	AI080784		
<i>Tc-rpp-0</i>	60S ribosomal protein p0	<i>C. elegans</i>	4.4E-35	Tc-EST-122 ^b	1,100	476	552	AA835588	AA835589		
<i>Tc-rpp-2</i>	60S ribosomal protein P2	<i>B. malayi</i>	3.6E-08	Tc-EST-088	288	288	275	AA907964	AA907965		
<i>Tc-rps-4</i>	40S ribosomal protein S4	<i>C. griseus</i>	4.1E-61	Tc-EST-006	750	582		AA563574			
<i>Tc-rps-5</i>	40S ribosomal protein S5	<i>C. elegans</i>	1.4E-70	Tc-EST-215	610	571	675	AA874719	AA874720		
<i>Tc-rps-8</i>	40S ribosomal protein S8	<i>H. sapiens</i>	7.4E-66	Tc-EST-277	640	481	600	AI080929	AI080930		
<i>Tc-rps-9</i>	40S ribosomal protein S9	<i>Rattus norvegicus</i>	8.6E-38	Tc-EST-270	410	381		AA875776			
Others											
<i>Tc-gep-1</i>	Granulin/epithelin precursor	<i>M. musculus</i>	8.4E-31	Tc-EST-141	1,000	654	625	AA836692	AA836693		
<i>Tc-hih-4</i>	Histone H4	<i>C. elegans</i>	3.5E-60	Tc-EST-174	250	444	459	AA557121	AA873915		
<i>Tc-lah-1</i>	Lupus autoantigen homolog	<i>C. elegans</i>	1.5E-22	Tc-EST-177	450	607	520	AI083048	AI083049		
<i>Tc-mps-1</i>	Metallopanstimulin (=rps27)	<i>Strongyloides ratti</i>	8.3E-41	Tc-EST-178	1,430	367	400	AA873918	AA979709		
<i>Tc-ofm-1</i>	Olfactomedin (F11C3.2)	<i>C. elegans</i>	1.6E-54	Tc-EST-079 ^b	1,870	479	628	AA728643	AA728644		
<i>Tc-ptb-1</i>	PC4/TIS7/"interferon-related protein"	<i>M. musculus</i>	4.1E-20	Tc-EST-075 ^b	1,440	616	299	AA618627	AI078882		
<i>Tc-tlp-1</i>	Tubby-like protein	<i>H. sapiens</i>	1.6E-42	Tc-EST-242	1,530	506	585	AA874759	AA874760		

^a Numerous sequences for which closest species listed is not a nematode show a higher BLASTX score with a *C. elegans* sequence for which there is either no assignment or which is described as similar to the nonnematode sequence given here. One ribosomal RNA gene (Tc-EST-052, 5' AA610968, 3' AA610969) was also identified.

^b Longest clone.

major metabolic pathways of glycolysis and aerobic respiration, as well as essential processes such as amino acid synthesis and degradation. Particularly prominent among these is the mitochondrial cytochrome *c* oxidase subunit II (25/261 = 9.7% of all clones). As mitochondrial mRNAs are not polyadenylated, the presence of mitochondrial DNA-encoded sequences in the cDNA library is difficult to interpret quantitatively.

Antioxidants. Oxidative stress is highly detrimental to both parasitic (54) and free-living (34) nematodes. In tissue-dwelling parasites, reactive oxygen intermediates from aggressive granulocytes may be countered by expression of antioxidants such as glutathione peroxidase and/or superoxide dismutase (SOD). Previous work characterized a SOD gene (*Tc-sod-1*) from *T. canis* larvae and showed that no glutathione peroxidase activity or gene sequence was detectable in this parasite (47). The EST data set contained seven clones encoding SOD isoforms, all quite distinct from *Tc-sod-1* (approximately 66% divergence in protein sequence) and more similar to *C. elegans* gene F55H2.1. The seven clones represent four distinct isoforms each showing 10 to 20% divergence in nucleotide and deduced amino acid sequence, including two-codon insertions/deletions. This level of divergence and the presence of triplet insertions/deletions led us to designate these four isoforms as

separate genes, *Tc-sod-2*, *-3*, *-4*, and *-5*. Confirmation of this assignment is under way.

Ribosomal proteins. Twenty ESTs (comprising 13 different genes; 8.4%) encode ribosomal proteins. This is close to the range found with EST projects for other nematodes; for example, 8.5% (1,339/15,811) of *B. malayi* ESTs deposited are ribosomal (12a), as are 5.0% (11/218) of *Necator americanus* ESTs (12a).

Other proteins. Some ESTs had homology to mammalian proteins for which a function in nematodes is not obvious. Five particularly interesting findings were noted.

(i) **Granulin/epithelin precursor (*Tc-gep-1*).** Granulins are synthesized as large (500- to 600-amino-acid) precursors from which are derived seven small (~60-amino-acid) 12-cysteine peptides with growth factor-like activity (8, 9, 69). Mammalian and fish kidney epithelial cells are rich sources of these peptides (8, 60), as are human and rodent leukocytes, suggesting that granulins may fulfill cytokine-like functions (6). If this is so, the synthesis of a granulin homolog by *T. canis* larvae may be important in the interaction between parasite and the host immune system.

(ii) **Lupus autoantigen homolog (*Tc-lah-1*).** The lupus autoantigen (also known as Sjögren syndrome type B antigen) is

a highly conserved ribonucleoprotein which is a target of autoantibodies in systemic lupus erythematosus (39). As auto-immune responses can be initiated by infectious agents (53), the expression of this homolog by *T. canis* may be significant. Similarly, *O. volvulus* expresses the RAL-1 product, which is homologous to the Sjögren syndrome type A antigen (48).

(iii) **Olfactomedin (*Tc-ofm-1*)**. A *T. canis* member of the olfactomedin gene family was found. The prototype gene encodes a 57-kDa glycoprotein in the extracellular mucus matrix of the olfactory neuroepithelium of frogs (66), but additional homologs are widely expressed in mammalian brain (35). *Tc-ofm-1* shows maximum similarity to a *C. elegans* homolog. As *T. canis* larvae produce high levels of mucins, including those constituting the surface coat (25, 59), *Tc-OFM-1* protein may be involved with mucus layers in this parasite.

(iv) **PC4/TIS7/“interferon-related protein” (*Tc-ptf-1*)**. A *T. canis* gene which is similar to a mammalian product described as PC4 or TIS7, induced in cell lines by activators such as nerve growth factor or phorbol esters (63), has been isolated. The same mammalian gene has also been designated interferon-related protein, due to an erroneous deposition of this sequence in 1985 as murine beta interferon (accession no. J00424). We have named the *T. canis* gene, which bears no similarity to mammalian interferons, *pth-1* (PC4/TIS7 homolog).

(v) **Tubby-like protein (*Tc-tilp-1*)**. Recently, a new multigene family related to the mouse gene *tubby* has been recognized (38). Mutations in *tubby* in the mouse result in obesity and degenerative changes in adult life; however, the existence of conserved homologs in nematodes (including *C. elegans*) and in plants indicates that this gene encodes a protein which fulfills a fundamental—and as yet unrecognized—function in all higher organisms.

Conclusion. The analysis of expressed genes that we present here has achieved its aims of identifying a number of abundant transcripts, one of which (*ctl-1*) corresponds to a major secreted product and four of which (*ant-003*, *-005*, *-030*, and *-034*) represent novel gene sequences. We have also made the first steps toward a comprehensive gene catalogue of a biologically intriguing and clinically significant parasite, providing a resource and springboard for future studies.

The outcome of this study supports the proposition that the EST strategy is highly applicable to many metazoan organisms which have relatively large genome sizes ($\geq 10^8$ bp) (13, 14), especially where interest is focused on genes expressed at moderate to high levels. In addition, genes restricted to a life cycle stage can be identified in this manner. Further evidence for the success of this approach is seen in the Filarial Genome Project, which in 3 years has deposited in dbEST over 15,000 sequences, which are estimated to represent some 5,000 separate gene transcripts, or around 33% of the total gene complement of *B. malayi* (15, 62). The availability of the full genome sequence of *C. elegans* lends an exceptional opportunity to compare free-living and parasite gene sequences and structure, with the potential to identify adaptations requisite for parasitism at the molecular level. Along their evolutionary path, parasitic species must have developed a myriad of immune evasion mechanisms, and we anticipate that our study and others like it will be instrumental in identifying the novel immune evasion genes upon which parasite survival depends.

ACKNOWLEDGMENTS

We thank the Medical Research Council for project grant support. We thank Mark Blaxter for detailed critical comments on the manuscript.

REFERENCES

- Adams, M. D., M. Dubnick, A. R. Kerlavage, R. Moreno, J. M. Kelley, T. R. Utterback, J. W. Nagle, C. Fields, and J. C. Ventner. 1992. Sequence identification of 2,375 human brain genes. *Nature* **355**:632–634.
- Adams, M. D., A. R. Kerlavage, R. D. Fleischmann, et al. 1995. Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. *Nature* **377**(Suppl.):3–17.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
- Anderson, P., and J. Kimble. 1997. mRNA and translation, p. 185–208. In D. L. Riddle, T. Blumenthal, B. J. Meyer, and J. R. Priess (ed.), *C. elegans* II. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Badley, J. E., R. B. Grieve, D. D. Bowman, L. T. Glickman, and J. H. Rockey. 1987. Analysis of *Toxocara canis* larval excretory-secretory antigens: physicochemical characterization and antibody recognition. *J. Parasitol.* **73**:593–600.
- Bateman, A., D. Belcourt, H. Bennet, C. Lazure, and S. Solomon. 1990. Granulins, a novel class of peptide from leukocytes. *Biochem. Biophys. Res. Commun.* **173**:1161–1168.
- Beaver, P. C. 1966. *Biology of parasites*, p. 215–227. Academic Press, New York, N.Y.
- Belcourt, D. R., C. Lazure, and H. P. J. Bennet. 1993. Isolation and primary structure of the three major forms of granulin-like peptides from hematopoietic tissues of a teleost fish (*Cyprinus carpio*). *J. Biol. Chem.* **268**:9230–9237.
- Bhandari, V., R. G. E. Palfree, and A. Bateman. 1992. Isolation and sequence of the granulin precursor cDNA from human bone marrow reveals tandem cysteine-rich granulin domains. *Proc. Natl. Acad. Sci. USA* **89**:1715–1719.
- Bin, Z., J. Hawdon, S. Qiang, R. Hainan, Q. Huiqing, H. Wei, X. Shu-Hua, L. Tiehua, G. Xing, F. Zheng, and P. Hotez. 1999. *Ancylostoma* secreted protein 1 (ASP-1) homologues in human hookworms. *Mol. Biochem. Parasitol.* **98**:143–149.
- Bird, D. M., and D. L. Riddle. 1994. A genetic nomenclature for parasitic nematodes. *J. Nematol.* **26**:138–143.
- Blaxter, M. L. 1998. *Caenorhabditis elegans* is a nematode. *Science* **282**:2041–2046.
- Blaxter, M. L. Personal communication.
- Blaxter, M. L., M. Aslett, J. Daub, D. Guiliano, and The Filarial Genome Project. Parasitic helminth genomics. *Parasitology*, in press.
- Blaxter, M. L., D. B. Guiliano, A. L. Scott, and S. A. Williams. 1997. A unified nomenclature for filarial genes. *Parasitol. Today* **13**:416–417.
- Blaxter, M. L., N. Raghavan, I. Ghosh, D. Guiliano, W. Lu, S. A. Williams, B. Slatko, and A. L. Scott. 1996. Genes expressed in *Brugia malayi* infective third stage larvae. *Mol. Biochem. Parasitol.* **77**:77–93.
- Blumenthal, T., and K. Steward. 1997. RNA processing and gene structure, p. 117–145. In D. L. Riddle, T. Blumenthal, B. J. Meyer, and J. R. Priess (ed.), *C. elegans* II. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Boguski, M. S., T. M. J. Lowe, and C. M. Tolstoshev. 1993. dbEST—database for “expressed sequence tags.” *Nat. Genet.* **4**:332–333.
- Chakrabati, D., G. R. Reddy, J. B. Dame, E. C. Almira, P. J. Laipis, R. J. Ferl, T. P. Yang, T. C. Rowe, and S. M. Schuster. 1994. Analysis of expressed sequence tags from *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **66**:97–104.
- Chen, J.-M., P. M. Dando, N. D. Rawlings, M. A. Brown, N. E. Young, R. A. Stevens, E. Hewitt, C. Watts, and A. J. Barrett. 1997. Cloning, isolation and characterization of mammalian legumain, an asparaginyl endopeptidase. *J. Biol. Chem.* **272**:8090–8098.
- de Savigny, D. H. 1975. In vitro maintenance of *Toxocara canis* larvae and a simple method for the production of *Toxocara* ES antigen for use in serodiagnosis test for visceral larva migrans. *J. Parasitol.* **61**:781–782.
- Dissanayake, S., M. Xu, C. Nkenfou, and W. F. Piessens. 1993. Molecular cloning and serological characterization of a *Brugia malayi* pepsin inhibitor homolog. *Mol. Biochem. Parasitol.* **62**:143–146.
- El-Sayed, N. M. A., C. M. Alarcon, J. C. Beck, V. C. Sheffield, and J. E. Donelson. 1995. cDNA expressed sequence tags of *Trypanosoma brucei rhodesiense* provide new insights into the biology of the parasite. *Mol. Biochem. Parasitol.* **73**:75–90.
- Falcone, F. H., et al. Unpublished data.
- Fan, J., D. J. Minchella, S. R. Day, D. P. McManus, W. U. Tiu, and P. J. Brindley. 1998. Generation, identification, and evaluation of expressed sequence tags from different developmental stages of the Asian blood fluke *Schistosoma japonicum*. *Biochem. Biophys. Res. Commun.* **252**:348–356.
- Fleming, J. T., H. A. Baylis, D. B. Sattelle, and J. A. Lewis. 1996. Molecular cloning and in vitro expression of *C. elegans* and parasitic nematode ionotropic receptors. *Parasitology* **113**:S175–S190.
- Gems, D., and R. M. Maizels. 1996. An abundantly expressed mucin-like protein from *Toxocara canis* infective larvae: the precursor of the larval surface coat glycoproteins. *Proc. Natl. Acad. Sci. USA* **93**:1665–1670.
- Gems, D. H., C. J. Ferguson, B. D. Robertson, A. P. Page, M. L. Blaxter, and R. M. Maizels. 1995. An abundant, *trans*-spliced mRNA from *Toxocara canis* infective larvae encodes a 26 kDa protein with homology to phosphatidyleth-

- anolamine binding proteins. *J. Biol. Chem.* **270**:18517–18522.
27. Gillespie, S. H. 1987. Human toxocarriasis. *J. Appl. Bacteriol.* **63**:473–479.
 28. Glickman, L. T., and P. M. Schantz. 1981. Epidemiology and pathogenesis of zoonotic toxocarriasis. *Epidemiol. Rev.* **3**:230–250.
 29. Guénette, S., R. K. Prichard, R. D. Klein, and G. Matlashewski. 1991. Characterization of a β -tubulin gene and β -tubulin gene products of *Brugia pahangi*. *Mol. Biochem. Parasitol.* **44**:153–164.
 30. Hawdon, J. M., B. F. Jones, D. R. Hoffman, and P. J. Hotez. 1996. Cloning and characterization of *Ancylostoma*-secreted protein. A novel protein associated with the transition to parasitism by infective hookworm larvae. *J. Biol. Chem.* **271**:6672–6678.
 31. Hawdon, J. M., S. Narasimhan, and P. J. Hotez. 1999. *Ancylostoma* secreted protein 2: cloning and characterization of a second member of a family of nematode secreted proteins from *Ancylostoma caninum*. *Mol. Biochem. Parasitol.* **99**:149–165.
 32. Healer, J., F. Ashall, and R. M. Maizels. 1991. Characterization of proteolytic enzymes from larval and adult *Nippostrongylus brasiliensis*. *Parasitology* **103**:305–314.
 33. Irvine, M., T. Huima, A. M. Prince, and S. Lustigman. 1994. Identification and characterization of an *Onchocera volvulus* cDNA clone encoding a highly immunogenic calponin-like protein. *Mol. Biochem. Parasitol.* **65**:135–146.
 34. Ishii, N., M. Fujii, P. S. Hartman, M. Tsuda, K. Yasuda, N. Senoo-Matsuda, S. Yanase, D. Ayusawa, and K. Suzuki. 1998. A mutation in succinate dehydrogenase cytochrome b causes oxidative stress and ageing in nematodes. *Nature* **394**:694–697.
 35. Karavanich, C. A., and R. R. H. Anholt. 1998. Molecular evolution of olfactomedin. *Mol. Biol. Evol.* **15**:718–726.
 36. Keddie, E. M., T. Higazi, and T. R. Unnasch. 1998. The mitochondrial genome of *Onchocerca volvulus*: sequence, structure and phylogenetic analysis. *Mol. Biochem. Parasitol.* **95**:111–127.
 37. Klein, R. D., S. C. Nulf, S. J. Alexander-Bowman, C. B. Mainone, C. A. Winterrowd, and T. G. Geary. 1992. Cloning of a cDNA encoding alpha-tubulin from *Haemonchus contortus*. *Mol. Biochem. Parasitol.* **56**:345–348.
 38. Kleyn, P. W., W. Fan, S. G. Kovats, J. J. Lee, J. C. Pulido, Y. Wu, L. R. Berkemeier, D. J. Misumi, L. Holmgren, O. Charlat, E. A. Woolf, O. Tayber, T. Brody, P. Shu, F. Hawkins, B. Kennedy, L. Baldini, C. Ebeling, G. D. Alperin, J. Deeds, N. D. Lakey, J. Culpepper, H. Chen, M. A. Glücksmann-Kuis, G. A. Carlson, G. M. Duyk, and K. J. Moore. 1996. Identification and characterization of the mouse obesity gene *tubby*: a member of a novel gene family. *Cell* **85**:281–290.
 39. Kotzin, B. L. 1996. Systemic lupus erythematosus. *Cell* **85**:303–306.
 - 39a. Liu, L. Personal communication.
 40. Lloyd, S. 1993. *Toxocara canis*: the dog, p. 11–24. In J. W. Lewis and R. M. Maizels (ed.), *Toxocara and toxocarriasis: epidemiological, clinical and molecular perspectives*. Institute of Biology, London, England.
 - 40a. Loukas, A., M. Hintz, K. Tetteh, and R. M. Maizels. Unpublished data.
 41. Loukas, A., P. M. Selzer, and R. M. Maizels. 1998. Characterisation of *Tc-cpl-1*, a cathepsin L-like cysteine protease from *Toxocara canis* infective larvae. *Mol. Biochem. Parasitol.* **92**:275–289.
 42. Loukas, A., N. P. Mullin, K. K. A. Tetteh, L. Moens, and R. M. Maizels. A novel C-type secreted by a tissue-dwelling parasitic nematode. *Curr. Biol.*, in press.
 43. Lustigman, S., J. H. McKerrow, K. Sha, J. Lui, T. Huima, M. Hough, and B. Brotman. 1996. Cloning of a cysteine protease required for the molting of *Onchocerca volvulus* third stage larvae. *J. Biol. Chem.* **271**:30181–30189.
 44. Maizels, R. M., D. H. Gems, and A. P. Page. 1993. Synthesis and secretion of TES antigens from *Toxocara canis* infective larvae, p. 141–150. In J. W. Lewis and R. M. Maizels (ed.), *Toxocara and toxocarriasis: epidemiological, clinical and molecular perspectives*. Institute of Biology, London, England.
 45. Maizels, R. M., M. W. Kennedy, M. Meghji, B. D. Robertson, and H. V. Smith. 1987. Shared carbohydrate epitopes on distinct surface and secreted antigens of the parasitic nematode *Toxocara canis*. *J. Immunol.* **139**:207–214.
 46. Manger, I. D., A. Hehl, S. Parmley, L. D. Sibley, M. Marra, L. Hillier, R. Waterston, and J. C. Boothroyd. 1998. Expressed sequence tag analysis of the bradyzoite stage of *Toxoplasma gondii*: identification of developmentally regulated genes. *Infect. Immun.* **66**:1632–1637.
 47. Matzilevich, D. A., C. Tripp, L. Tang, X. Ou, L. Matzilevich, M. K. Shaw, and R. M. Maizels. Cloning, expression and localization of *Tc-sod-1*, an extracellular superoxide dismutase from infective stage larvae of *Toxocara canis*. Submitted for publication.
 48. McCauliffe, D. P., E. Zappi, T.-S. Lieu, M. Michalak, R. D. Sontheimer, and J. D. Capra. 1990. A human Ro/SS-A autoantigen is the homologue of calreticulin and is highly homologous with onchocercal RAL-1 antigen and an aplysia “memory molecule.” *J. Clin. Investig.* **86**:332–335.
 49. McCombie, W. R., M. D. Adams, J. M. Kelley, M. G. FitzGerald, T. R. Utterback, M. Khan, M. Dubnick, A. P. Kerlavage, J. C. Ventner, and C. Fields. 1992. *Caenorhabditis elegans* expressed sequence tags identify gene families and potential disease gene homologues. *Nat. Genet.* **1**:124–131.
 50. Meghji, M., and R. M. Maizels. 1986. Biochemical properties of larval excretory-secretory glycoproteins of the parasitic nematode *Toxocara canis*. *Mol. Biochem. Parasitol.* **18**:155–170.
 51. Moore, J., V. Todorova, and M. W. Kennedy. 1995. A cDNA encoding ribosomal protein L3 from the parasitic nematode *Toxocara canis*. *Gene* **165**:239–242.
 - 51a. Murray, J., et al. Unpublished data.
 52. Okimoto, R., J. L. Macfarlane, D. O. Clary, and D. R. Wolstenholme. 1992. The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*. *Genetics* **130**:471–498.
 53. Oldstone, M. B. A. 1987. Molecular mimicry and autoimmune disease. *Cell* **50**:819–820.
 54. Ou, X., R. Thomas, M. R. Chacón, L. Tang, and M. E. Selkirk. 1995. *Brugia malayi*: differential susceptibility to and metabolism of hydrogen peroxide in adults and microfilariae. *Exp. Parasitol.* **80**:530–540.
 55. Page, A. P. 1997. Cyclophilin and protein disulfide isomerase genes are co-transcribed in a functionally related manner in *Caenorhabditis elegans*. *DNA Cell Biol.* **16**:1335–1343.
 56. Page, A. P., A. J. Hamilton, and R. M. Maizels. 1992. *Toxocara canis*: monoclonal antibodies to carbohydrate epitopes of secreted (TES) antigens localize to different secretion-related structures in infective larvae. *Exp. Parasitol.* **75**:56–71.
 57. Page, A. P., K. MacNiven, and M. O. Hengartner. 1996. Cloning and biochemical characterization of the cyclophilin homologues from the free-living nematode *Caenorhabditis elegans*. *Biochem. J.* **317**:179–185.
 58. Page, A. P., and R. M. Maizels. 1992. Biosynthesis and glycosylation of serine/threonine-rich secreted proteins from *Toxocara canis* larvae. *Parasitology* **105**:297–308.
 59. Page, A. P., W. Rudin, E. Fluri, M. L. Blaxter, and R. M. Maizels. 1992. *Toxocara canis*: a labile antigenic coat overlying the epicuticle of infective larvae. *Exp. Parasitol.* **75**:72–86.
 60. Plowman, G. D., J. M. Green, M. G. Neubauer, S. D. Buckley, V. L. McDonald, G. J. Todaro, and M. Shoyab. 1992. The epithelin precursor encodes two proteins with opposing activities on epithelial cell growth. *J. Biol. Chem.* **267**:13073–13078.
 61. Schallig, H. D. F. H., M. A. W. van Leeuwen, B. E. Verstrepen, and A. W. C. A. Cornelissen. 1997. Molecular characterization and expression of two putative protective excretory secretory proteins of *Haemonchus contortus*. *Mol. Biochem. Parasitol.* **88**:203–213.
 - 61a. Tetteh, K. Unpublished data.
 62. The Filarial Genome Project. Deep within the filarial genome: an update on progress in the Filarial Genome Project. *Parasitol. Today*, in press.
 63. Tirone, F., and E. M. Shooter. 1989. Early gene regulation by nerve growth factor in PC12 cells: induction of an interferon-related gene. *Proc. Natl. Acad. Sci. USA* **86**:2088–2092.
 64. Verdun, R. E., N. Di Paolo, T. P. Urmenyi, E. Rondinelli, A. C. C. Frasc, and D. O. Sanchez. 1998. Gene discovery through expressed sequence tag sequencing in *Trypanosoma cruzi*. *Infect. Immun.* **66**:5393–5398.
 65. Wilson, W. R., R. S. Tuan, K. J. Shepley, D. O. Freedman, B. M. Greene, K. Awadzi, and T. R. Unnasch. 1994. The *Onchocerca volvulus* homologue of the multifunctional polypeptide protein disulfide isomerase. *Mol. Biochem. Parasitol.* **68**:103–117.
 66. Yokoe, H., and R. R. H. Anholt. 1993. Molecular cloning of olfactomedin, an extracellular matrix protein specific to olfactory neuroepithelium. *Proc. Natl. Acad. Sci. USA* **90**:4655–4659.
 67. Zeng, W., and J. E. Donelson. 1992. The actin genes of *Onchocerca volvulus*. *Mol. Biochem. Parasitol.* **55**:207–216.
 68. Zhang, B., M. Gallegos, A. Puoti, E. Durkin, S. Fields, J. Kimble, and M. P. Wickens. 1997. A conserved RNA-binding protein that regulates sexual fates in the *C. elegans* hermaphrodite germ line. *Nature* **390**:477–484.
 69. Zhou, J., G. Gao, J. W. Crabb, and G. Serrero. 1993. Purification of an autocrine growth factor homologous with mouse epithelin precursor from a highly tumorigenic cell line. *J. Biol. Chem.* **268**:10863–10869.