

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

International Journal for Parasitology 36 (2006) 829–839

www.elsevier.com/locate/ijpara

Identification and analysis of genes expressed in the adult filarial parasitic nematode *Dirofilaria immitis*[☆]

Yong Yin^{a,*}, John Martin^a, James P. McCarter^{a,b}, Sandra W. Clifton^a, Richard K. Wilson^a, Makedonka Mitreva^a

^a Department of Genetics, Genome Sequencing Center, Washington University School of Medicine, Box 8501, St Louis, MO 63108, USA

^b Divergence Inc., 893 North Warson Road, St Louis, MO 63141, USA

Received 23 November 2005; received in revised form 28 February 2006; accepted 2 March 2006

Abstract

The heartworm *Dirofilaria immitis* is a filarial parasitic nematode infecting dogs and other mammals worldwide causing fatal complications. Here, we present the first large-scale survey of the adult heartworm transcriptome by generation and analysis of 4005 expressed sequence tags, identifying about 1800 genes and expanding the available sequence information for the parasite significantly. *Brugia malayi* genomic data offered the most valuable information to interpret heartworm genes, with about 70% of *D. immitis* genes showing significant similarities to the assembly. Comparative genomic analyses revealed both genes common to metazoans or nematodes and genes specific to filarial parasites that may relate to parasitism. Characterization of abundant transcripts suggested important roles for genes involved in energy generation and antioxidant defense in adults. In particular, we proposed that adult heartworm likely adopted an anaerobic electron transfer-based energy generation system distinct from the aerobic pathway utilized by its mammalian host, making it a promising target in developing next generation macrofilaricides and other treatments. Our survey provided novel insights into the *D. immitis* transcriptome and laid a foundation for further comparative studies on biology, parasitism and evolution within the phylum Nematoda.

© 2006 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.

Keywords: *Dirofilaria immitis*; Heartworm; Parasite; Expressed sequence tags (EST's); Comparative genomics; Filarial nematode; *Caenorhabditis elegans*; Energy generation

1. Introduction

The dog heartworm *Dirofilaria immitis* belongs to the clade III filarial parasitic nematodes (Blaxter et al., 1998), slender roundworms that are parasitic in blood and tissue of vertebrates and have life cycle stages within intermediate insect hosts. Heartworm has a worldwide distribution in temperate and tropical climates, infecting dogs, its definitive host, as well as cats, ferrets and other mammals including humans (Knight, 1977, 1987). L1 larvae (microfilariae, ~300 µm long), which are able to survive 1–3 years circulating in the host bloodstream, develop into the infective L3 stage (L3i) after entering

a mosquito during its feeding. After entering another host when the mosquito feeds again, the L3i larvae molt and develop further into adults (Knight, 1977, 1987). Adult females (up to 27 cm long) and males (~17 cm) normally reside in the host pulmonary arteries and right ventricles. Their life span in dogs is at least 5–7 years, with the number of adults reaching up to 250 in a single dog, resulting in various lung and heart diseases such as severe pulmonary arterial inflammation and congestive heart failure that can be fatal (Knight, 1977, 1987). In the United States, heartworm infection has been found in dogs native to all 50 states. A 2001 survey of over 18,000 veterinary clinics reported heartworm positive tests for more than 240,000 dogs and 3000 cats (McCall, 2005), with the actual incidence likely much higher. Heartworm test kits and macrocyclic lactone anthelmintic drugs used to prevent the infection are among the best-selling products in companion animal veterinary medicine, even though only half of dogs are estimated to receive any preventive treatments. Currently, the arsenical melarsomine dihydrochloride is the only Federal

[☆] Note: EST sequences are available from GenBank, EMBL, and DDJB under the accession numbers BQ454296–BQ457429, BQ481967–BQ482740, BU587123–BU587163, and CD285915–CD285970. The sequences are also available at www.nemnode.net.

* Corresponding author. Tel.: +1 314 286 1118; fax: +1 314 286 1810.

E-mail address: yyin@watson.wustl.edu (Y. Yin).

Department of Agriculture-approved drug to eliminate adult heartworms in infected dogs (McTier et al., 1994). Both new preventive strategies such as vaccines and safer curative macrofilaricides are needed.

To date, studies of the *D. immitis* transcriptome have been limited to individual genes, with only 112 *D. immitis* gene sequences available in public databases before the current study (April 2005). A comprehensive expressed sequence tag (EST) study would greatly accelerate the identification of heartworm genes and prediction of their putative functions based on primary sequence similarities, helping us to better understand its biology and eventually develop next generation drugs and vaccines. In addition, various human endemic diseases are caused by related filarial nematodes. Lymphatic filariasis, a disease carried by 120 million people worldwide and recognized as the second leading cause of permanent and long-term disability, is caused by the infection with *Wuchereria bancrofti*, *Brugia malayi* and *Brugia timori* (Molyneux et al., 2003). Another filarial nematode *Onchocerca volvulus* is the causative agent of river blindness (onchocerciasis) infecting 18 million individuals worldwide (Molyneux et al., 2003). A large-scale survey and analysis of the *D. immitis* transcriptome will complement the data being generated from the human parasites in various EST and genome sequencing projects (Blaxter et al., 1999, 2002; Williams et al., 2000, 2002; Ghedin et al., 2004), setting a stage for comparative genomic studies among veterinary and human filarial parasites.

Here we report, for the first time, a large-scale survey of the adult *D. immitis* transcribed genome, identifying about 1800 genes by sequencing and analyzing 4005 ESTs. Comparative genomic approaches were used to interpret heartworm genes using sequence data from other species, especially those generated in the *B. malayi* Genome Project (Ghedin et al., 2004). In addition, those comparisons identified putative proteins specific to nematodes or individual nematode clades. Abundant transcripts were examined into more detail, highlighting special aspects of adult heartworm physiology.

2. Materials and methods

2.1. Source material, library construction, and EST sequencing

Adult heartworms were collected at necropsy of a euthanized adult dog obtained from an animal shelter near College of Veterinary Medicine at North Carolina State University. The live worms, including both males and females, were rinsed in $1 \times$ sterile PBS, frozen in liquid N_2 immediately upon recovery and stored at $-80^\circ C$. Pulverization of four to five whole worms of mixed sexes was performed using an alloy steel mortar and pestle set (Fisher Scientific). The absolute and relative numbers of males and females used are not known. No special measures were taken to remove developing embryos and larvae from the sample; therefore a small number of embryonic transcripts were expected to be included in the EST sampling. Poly(A)+RNA was isolated, two cDNA libraries were constructed (using either the splice leader 1 (SL1)-based or SMART (Clontech Laboratories) cDNA library construction

systems) from which 815 and 3190 ESTs were generated by 5'-end sequencing, respectively, as described previously (McCarter et al., 2003; Mitreva et al., 2004a,b).

2.2. Sequence analysis and functional assignment

Examination, processing, and clustering of the EST sequences were performed as described previously (McCarter et al., 2003; Mitreva et al., 2004a). A maximum likelihood-based program ESTFREQ (W. Gish, unpublished) was used to estimate the complexity of the EST libraries; TRANSLATE (S. Eddy, unpublished) was used for predicting open reading frames (ORFs) with default parameters. Databases used for sequence comparison were: *B. malayi* genome draft assembly (October 2005), WORMPEP (v140), GenBank (15 April 2005), non-redundant protein database NR (15 April 2005) and genome of *Wolbachia* strain TRS of *B. malayi* (15 April 2005). WU-BLASTX ($S=100$ $M=PAM120$ $V=0$ $W=4$ $T=17$) was used to query translated nucleotide sequences against protein databases and WU-TBLASTX ($Q=10$ $R=2$ $gapw=10$ $filter=seg+xnu$ $hspsepsmax=5000$ $gapsepsmax=5000$) for querying translated nucleotide sequences versus translated nucleotide databases in all six reading frames. Expectation value (E -value) of $1.0 e^{-5}$ was used as the cut-off to accept sequence similarities in both BLAST searches. Default parameters for InterProScan (Quevillon et al., 2005) were used to query against InterPro (Release 9.0) (Mulder et al., 2005) in assigning Gene Ontology terms (Release 2,00,509) (The Gene Ontology Consortium, 2000). An E -value cut-off of $1.0 e^{-10}$ reported by WU-BLASTX against the GeneDB (Release 34.0) from Kyoto Encyclopedia of Genes and Genomes (KEGG) was used for metabolic pathway association, the top match and all the matches within a range of 30% of the top score, if meeting the cut-off, were accepted for KEGG association (Bono et al., 1998; Kanehisa and Goto, 2000; Kanehisa et al., 2004). TargetP (v1.1) was used to identify *D. immitis* clusters with putative signal peptides for secretion and predictions within reliability class 1 were accepted (Emanuelsson et al., 2000); such clusters were further screened to identify transmembrane (TM) regions by TMHMM (v2.0) (Krogh et al., 2001). A *D. immitis* protein was assigned as secreted if either no TM region was identified or only one TM domain was predicted and its First60 value was between 10 and 30.

3. Results and discussion

3.1. Properties and clustering of ESTs

As part of a larger effort to examine ESTs from about 30 parasitic nematodes (Mitreva et al., 2005), we constructed two *D. immitis* adult cDNA libraries using different protocols and sequenced in total 4005 ESTs from the 5' end, of which 3999 passed our automated screen and manual inspection. They were submitted to the dbEST division of GenBank immediately upon generation in 2002 and 2003. Their average length is 405 nucleotides, totaling 1.62 million bases, which represents

97% of total sequence for heartworm currently available. By applying our semi-automated bioinformatics tools (McCarter et al., 2003; Mitreva et al., 2004a), the 3999 ESTs were grouped into 2534 contigs, each containing members derived from identical transcripts. This step reduced the number of nucleotides for subsequent analysis to 1.09 million. The average contig length was 431 nucleotides; the median length increased from 441 to 532 nucleotides. The longest contig sequence contains 1082 bases, while the longest EST read is 719 bases.

To further determine unique gene representation, contigs were conservatively assigned to 1793 NemaGene clusters, most containing transcripts derived from a single genomic locus but containing variations generated by, for example, polymorphism, allelic difference, or alternative splicing, while some may also represent members of highly conserved gene families (Mitreva et al., 2004a). In addition, multiple clusters may represent a single gene if they cover non-overlapping portions of the same transcript (i.e. fragmentation). Using *Caenorhabditis elegans* as a reference, we found 834 *D. immitis* clusters contained similarities to 816 *C. elegans* genes. The majority of these (800/834, 95.9%) had a one-to-one relationship with their *C. elegans* counterparts, leaving only 20 clusters with similarities to multiple regions of the *C. elegans* genes (fragmentation rate of 2.2%). Discounting for fragmentation, we estimated that 1754 unique *D. immitis* genes were identified, suggesting a new gene discovery rate of 44% (1754/3999). Assuming *D. immitis* has a similar number of loci as *C. elegans* (19,735, WORMPEP v140), we likely identified about 9% of all *D. immitis* genes.

The size of the NemaGene clusters ranged from a single EST to 144 (Fig. 1). The majority of the clusters (70%) have 10 or fewer EST members, with 32% containing a single EST and

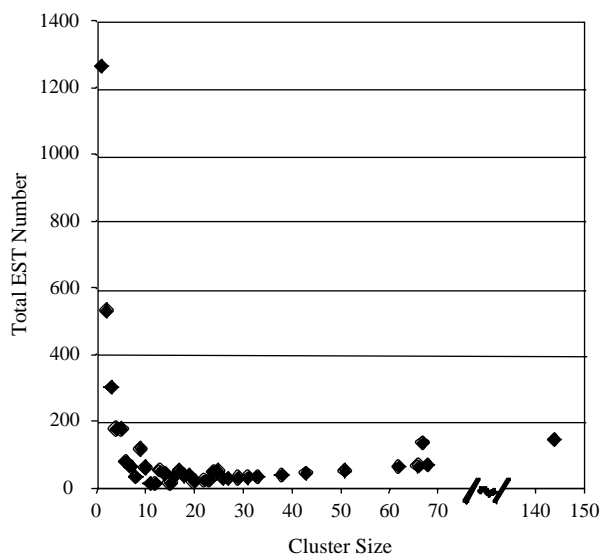


Fig. 1. Distribution of *Diriofilaria immitis* expressed sequence tags (ESTs) by cluster sizes. The number of total ESTs included by all the clusters with the same cluster size was summed up and shown. None but the largest cluster (with 144 ESTs) contains more than 70 EST members and 1264 clusters contain only one EST.

a significant number of clusters (302, 17%) contain multiple contig members. Statistical analysis showed that a second EST sampling of the same size would provide only 18% novel sequences, suggesting cDNA library normalization would be desirable for further sampling. Furthermore, only 19 clusters (1%) represented previously available heartworm genes in public databases, indicating that our study has generated valuable novel sequence information for *D. immitis*.

3.2. Similarities to genes from other species

The TIGR *B. malayi* Genome Project (Ghedini et al., 2004), which aims to generate the first sequenced genome of a parasitic nematode, has recently been completed. Being also a clade III filarial parasite, the genomic sequence of *B. malayi* can be expected to provide the best source of information in interpreting *D. immitis* genes. Indeed, 70% (1254/1793) of the *D. immitis* clusters showed similarities to the *B. malayi* draft genome assembly (October 2005). The actual number of matches to the genome is likely higher because the current *B. malayi* assembly contains gaps and has only 11,749 annotated gene models (E. Ghedin, personal communications), while about 18,500 genes have been proposed previously according to a genome sequence survey (Whitton et al., 2004). Furthermore, the majority of the similarities across the two filarial species are much stronger than those detected towards genes from nematodes of other clades (e.g. Table 1; Supplementary Tables 1), confirming the close evolutionary relationship between the two filarial parasites. In addition, 6% of all clusters (105/1793) lacked similarities to the current annotated *B. malayi* genes, yet showed significant similarities to regions of the *B. malayi* genome when translated. Such matches likely reveal potential protein-coding regions within the genome that have not been predicted, consistent with the expectation that ESTs from other nematodes would facilitate annotation of the *B. malayi* genome. Similarly, ESTs generated by our Parasitic Nematode Sequencing Project (<http://www.nematode.net>) are being continuously mapped to the *C. elegans* genome to refine its annotation.

To obtain an overall interpretation of the genes expressed by adult heartworm, we queried our cluster sequences against three databases containing genes from different species: (i) filarial Nematoda, which contains nucleotide sequences of the *B. malayi* draft assembly and all non-*D. immitis* filarial nematode genes in GenBank, (ii) non-filarial Nematoda, containing all non-filarial nematode nucleic acid sequences from GenBank, (iii) non-Nematoda, amino acid sequences from the non-redundant protein database NR excluding those from nematodes. In total, using the *E*-value of 1.0×10^{-5} as the cut-off, 1403 clusters (78%) were shown to contain similarities to genes from other species in at least one of the three databases and distribution of the matches likely reflects the phylogenetic relationships among metazoans (Fig. 2). The 22% of transcripts lacking similarities may represent: (i) novel genes, (ii) untranslated regions (UTRs) that have little similarity in other species, or (iii) short ORFs that are difficult to detect by

Table 1
The 25 most abundant transcripts in adult *Dirofilaria immitis*

| Cluster | #EST | Best NR <i>B. malayi</i> | Hit | Bit Score | %ID ^b | Description | Top Hit in <i>C. elegans</i> | <i>C. elegans</i> RNAi ^c | <i>C. elegans</i> Gene Ontology |
|---------|------------|--------------------------------|-----|----------------|------------------|--|---------------------------------|-------------------------------------|--|
| 1 | D100105.cl | CAA95851 14379.m00149 | 144 | 97.6 171.9 | 55.9 94.3 | <i>Caenorhabditis elegans</i> putative NADH-ubiquinone oxidoreductase subunit B15 <i>Brugia malayi</i> conserved hypothetical protein | W01A8.4 | Emb Lva Bmd | electron transfer |
| 2 | D100268.cl | CAE62144 14610.m00102 | 68 | 284.9 401.8 | 62.3 90.8 | <i>Caenorhabditis briggsae</i> CBG06190, Fat-6, delta-9 fatty acid desaturase <i>B. malayi</i> putative acyl-CoA desaturase | VZK822L.1 | Emb Gro Stp Unc, etc | fatty acid biosynthesis |
| 3 | D100534.cl | BAA97098 14972.m07436 | 67 | 97.6 102.2 | 45.0 90.3 | <i>Arabidopsis thaliana</i> putative histone subunit <i>B. malayi</i> putative histone H2A variant | R08C7.3 | Gro Unc Lva | nucleosome assembly |
| 4 | D100492.cl | CAE73320 14975.m04447 | 67 | 60.0 40.6 | 62.5 59.4 | <i>C. briggsae</i> CBG20748, Hot-3, GPI-linked <i>B. malayi</i> hypothetical protein | W02B8.5 | WT | - |
| 5 | D100542.cl | - | 66 | - | - | - | - | - | - |
| 6 | D100041.cl | AAH70989 14950.m01860 | 62 | 225.1 298.6 | 63.9 93.4 | <i>Xenopus laevis</i> MGC78865, putative subunit of actin nucleator ARP2/3 <i>B. malayi</i> putative ARP2/3 complex, 34 kDa subunit | Y6D11A.2 | Emb | regulation of actin polymerization |
| 7 | D100002.cl | CAA48632 14083.m00056 | 51 | 299.7 294.4 | 91.2 87.1 | <i>Onchocerca volvulus</i> OV25-1, small heat shock protein <i>B. malayi</i> putative small heat shock protein OV25-1 | F52E1.7b | WT | - |
| 8 | D100102.cl | CAA94883 14972.m07036 | 43 | 52.4 141.6 | 36.0 71.9 | <i>C. elegans</i> putative tRNA pseudouridine synthase D <i>B. malayi</i> tRNA pseudouridine synthase D | B0024.11 | WT | tRNA processing |
| 9 | D100008.cl | AAK29794 14400.m00023 | 38 | 209.9 270.8 | 58.9 90.6 | <i>C. elegans</i> testican, Ca2+-binding proteoglycan <i>B. malayi</i> thyroglobulin type-1 repeat family protein | C54E4.2 | WT | calcium ion binding |
| 10 | D100032.cl | CAD61199 14614.m00134 | 33 | 145.2 42.0 | 96.2 30.7 | <i>D. immitis</i> NADH-ubiquinone oxidoreductase subunit 3 <i>B. malayi</i> predicted protein | MTCE.35 | - | electron transfer |
| 11 | D100219.cl | - | 31 | - | - | - | - | - | - |
| 12 | D100010.cl | CAE58565 13281.m00013 | 29 | 83.6 113.1 | 54.7 84.7 | <i>C. briggsae</i> CBG01728, methyltransferase activity <i>B. malayi</i> hypothetical 28.7 kDa protein | F29B9.1 | WT | methyltransferase activity |
| 13 | D100017.cl | CAA94909 14963.m01811 | 27 | 302.2 138.5 | 85.4 60.4 | <i>C. elegans</i> hypothetical protein, Rhodopsin-like GPCR activity <i>B. malayi</i> AT19640p-related protein | C35A5.7 | WT | G-protein coupled receptor signaling |
| 14 | D100336.cl | CAE73126 14990.m08028 | 26 | 332.8 259.9 | 91.7 69.3 | <i>C. briggsae</i> CBG20509, protein kinase activity <i>B. malayi</i> nrhp-prov protein-related | H37N21.1 | Gro Elg Unclassified | protein amino acid phosphorylation |
| 15 | D100049.cl | CAE69187 14982.m02231 | 25 | 348.3 227.2 | 92.1 84.1 | <i>C. briggsae</i> CBG15221, putative DEAH-box RNA helicase <i>B. malayi</i> putative ATP-dependent helicase DHX35 | Y67D2.6 | WT | ATP-dependent helicase activity |
| 16 | D100161.cl | - | 25 | - | - | - | - | - | - |
| 17 | D100498.cl | CAA21542 15187.m00012 | 24 | 61.4 107.5 | 32.3 78.2 | <i>C. elegans</i> hypothetical protein <i>B. malayi</i> RNA-metabolising metallo-beta-lactamase family protein | Y37D8A.19 | Lva | growth |
| 18 | D100040.cl | CAE63201 14972.m07733 | 24 | 230.4 352.9 | 60.8 96.4 | <i>C. briggsae</i> CBG07556, Sym-1, containing Leucine-rich Repeat <i>B. malayi</i> Leucine-rich Repeat family protein | C44H4.3 | WT | - |
| 19 | D100495.cl | - | 23 | - | - | - | - | - | - |
| 20 | D100494.cl | NP_500041 13411.m00124 | 22 | 47.0 259.9 | 32.4 71.4 | <i>C. elegans</i> hypothetical protein <i>B. malayi</i> hypothetical protein | C44B12.5 | Emb Ste Sip | reproduction |
| 21 | D100451.cl | CAE72557 15233.m00030 | 20 | 221.5 297.6 | 73.6 92.4 | <i>C. briggsae</i> CBG19741, Pbs-5, 20S proteasome regulatory subunit <i>B. malayi</i> proteasome A-type and B-type family protein | K05C4.1 | Emb Ocs Lva Ste Unc | ubiquitin-dependent protein catabolism |
| 22 | D100488.cl | CAE63803 14250.m00297 | 19 | 130.7 195.5 | 59.1 85.0 | <i>C. briggsae</i> CBG08349, putative NADH-ubiquinone oxidoreductase <i>B. malayi</i> phosphopantetheine attachment site family protein | F37C12.3 | Emb Lva Sma Sip Gro | electron transfer |
| 23 | D100096.cl | AAB68798 14972.m07237 | 19 | 368.3 191.3 | 97.9 64.4 | <i>D. immitis</i> peroxidase-1, thioredoxin peroxidase <i>B. malayi</i> putative thioredoxin peroxidase-1 | F09E5.15 | Seck Sma | physiological process |
| 24 | D100088.cl | NP_502816 14920.m00420 | 18 | 91.0 215.6 | 37.7 62.3 | <i>C. elegans</i> putative elongation factor 1 beta/delta chain <i>B. malayi</i> EF-1, containing guanine nucleotide exchange domain | Y41E3.10 | Unclassified | translation elongation |
| 25 | D100485.cl | CAE60102 14036.m00111 | 18 | 177.9 247.6 | 50.5 85.9 | <i>C. briggsae</i> CBG03635 <i>B. malayi</i> hypothetical protein | C36E8.2 | WT | heme binding |

^aNumber of ESTs contained in each cluster

^bPercent identity (%ID) at amino acid level

^cRNAi phenotype (RNAi) described using abbreviations listed at www.wormbase.org.

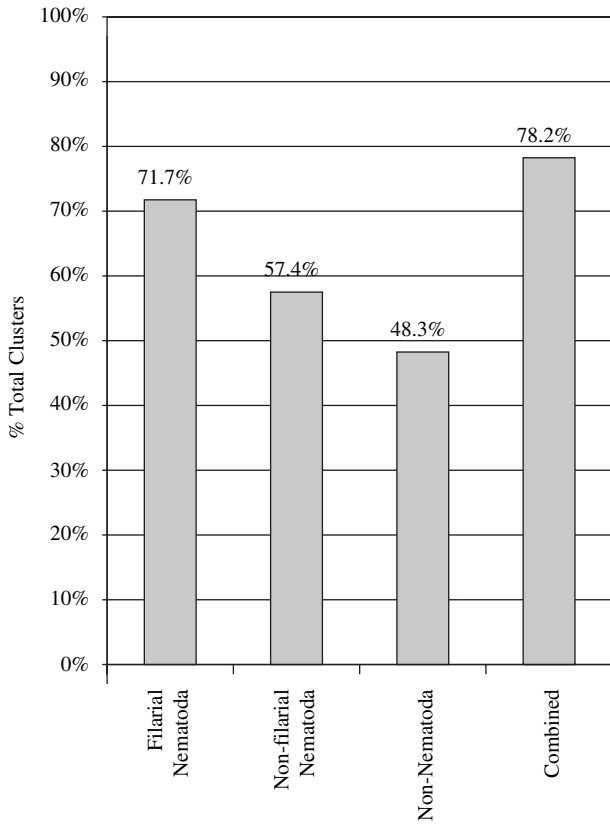


Fig. 2. Identification of amino acid similarities between *Dirofilaria immitis* clusters and sequences from other species. *E*-value of 1.0×10^{-5} was used as the cut-off for WU-BLASTX or WU-TBLASTX searches, depending on which database was searched against. Three databases used are: (i) filarial nematoda (*Brugia malayi* draft genome and all other GenBank nucleotide sequences with filarial origins excluding *D. immitis*), (ii) non-filarial nematoda (all GenBank nucleotide sequences from non-filarial nematodes), and (iii) non-nematoda (all NR protein sequences excluding those from nematodes). Overall, 78.2% *D. immitis* clusters contain similarities to genes from other species.

sequence comparison with our parameters per se. Indeed, contig sequences without matches tend to have shorter predicted ORFs (highest frequency is located between 20 and 40 amino acids), while contigs with matches have a bimodal distribution of ORF lengths, with frequency peaking at 40–60 and 160–180 amino acids, respectively (Fig. 3).

Of the 1403 (78%) clusters in which similarities were detected, 731 contained matches to all three databases (Fig. 4), implying that these are genes involved in the molecular and cellular processes conserved in metazoans. In contrast, 236 contained similarities to coding sequences from both filarial and non-filarial nematodes but not to non-nematode proteins. Consequently, they are potential candidates for nematode-specific genes. Furthermore, 54 showed similarities to only non-nematode proteins, suggesting gene-loss events or accelerated changes in other nematode lineages. Incomplete genomes and thus lack of representations for many nematode species could also contribute to this observation. Finally, 282 clusters with matches, or 16% (282/1793) of total heartworm clusters, contained similarities only to filarial nematodes (Fig. 4). Their lack of similarities

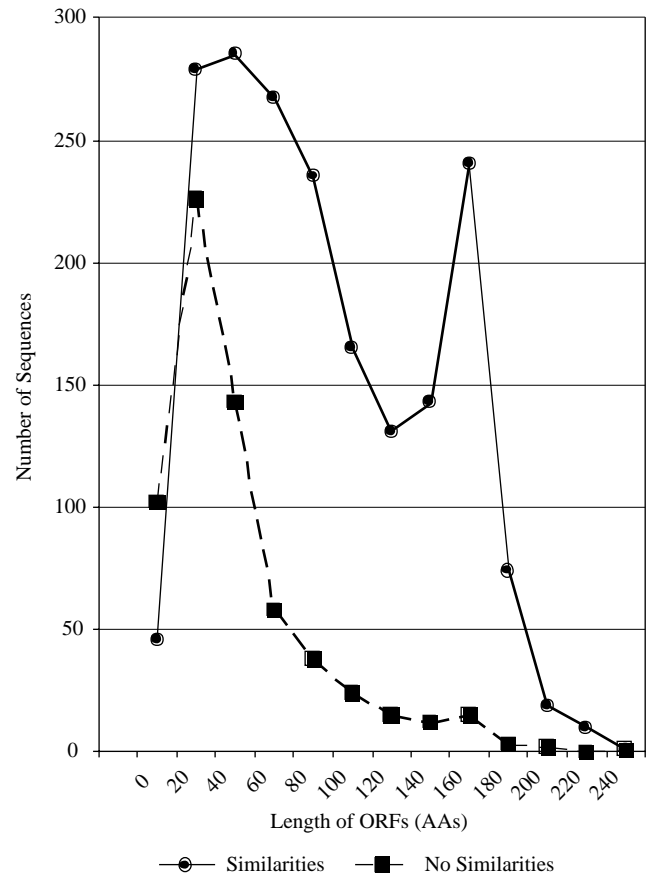


Fig. 3. Length distribution of putative open reading frames (ORFs) for *Dirofilaria immitis* contigs with (solid line) or without (dotted line) similarities to previously characterized proteins shown in amino acids (AAs). Those without similarities are substantially shorter.

to proteins from the well-studied *C. elegans* genome (The *C. elegans* Sequencing Consortium, 1998) and the draft genome of *Caenorhabditis briggsae* (Stein et al., 2003) suggests that such heartworm clusters represent putative parasitism-related genes adapted in filarial nematodes or reflect the phylogenetic distance between clade V and clade III nematodes (Blaxter et al., 1998). Nevertheless, we believe

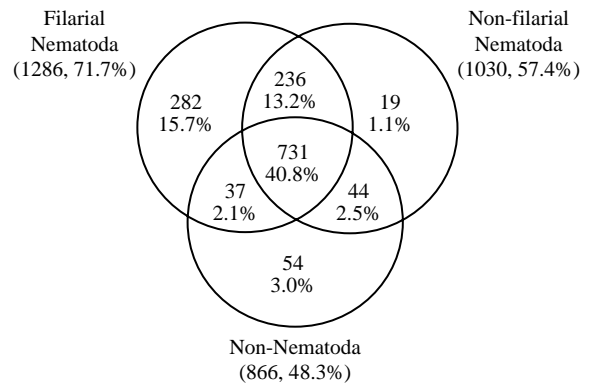


Fig. 4. Distribution of the amino acid similarities identified between *Dirofilaria immitis* clusters and sequences from other species. 1403 heartworm clusters (78% of the total) containing similarities identified by BLAST searches were shown in the Venn diagram. 282 clusters contained similarities shared only by filarial nematodes, representing a putative group of filarial specific genes.

Table 2a
Electron transfer-related transcripts in adult *Dirofilaria immitis* containing 248 expressed sequence tags (ESTs)

| Complex | Cluster | #EST ^a | Best NR hit | Bit score | %ID ^b | Description |
|------------|------------|-------------------|-------------|-----------|--|---|
| I | DI00105.cl | 144 | CAA95851 | 97.6 | 55.9 | <i>Caenorhabditis elegans</i> putative NADH-ubiquinone oxidoreductase subunit B15 |
| | DI00032.cl | 33 | CAD61199 | 145.2 | 96.2 | <i>D. immitis</i> NADH-ubiquinone oxidoreductase subunit 3 |
| | DI00488.cl | 19 | CAE63803 | 130.7 | 59.1 | <i>Caenorhabditis briggsae</i> NADH-ubiquinone oxidoreductase subunit |
| | DI00215.cl | 9 | CAD61199 | 109.3 | 94.8 | <i>D. immitis</i> NADH-ubiquinone oxidoreductase subunit 3 |
| | DI00432.cl | 6 | CAD61200 | 352.5 | 95.4 | <i>D. immitis</i> NADH-ubiquinone oxidoreductase subunit 5 |
| | DI00770.cl | 2 | CAD61196 | 144.1 | 100.0 | <i>D. immitis</i> NADH-ubiquinone oxidoreductase subunit 1 |
| | DI01411.cl | 1 | CAD61196 | 249.7 | 86.0 | <i>D. immitis</i> NADH-ubiquinone oxidoreductase subunit 1 |
| | DI01441.cl | 1 | CAD61192 | 80.0 | 92.6 | <i>D. immitis</i> NADH-ubiquinone oxidoreductase subunit 6 |
| DI02390.cl | 1 | CAA21667 | 44.1 | 61.2 | <i>C. elegans</i> NADH-ubiquinone oxidoreductase subunit 1 beta subcomplex 8 | |
| II | DI00310.cl | 2 | BAA11232 | 238.1 | 70.8 | <i>Ascaris suum</i> succinate dehydrogenase cytochrome b large subunit |
| | DI02456.cl | 1 | BAA23716 | 146.6 | 68.4 | <i>A. suum</i> succinate dehydrogenase iron-sulfur subunit |
| III | DI00209.cl | 3 | CAD61193 | 182.5 | 68.8 | <i>D. immitis</i> cytochrome b |
| | DI01098.cl | 1 | BAD11763 | 291.6 | 82.1 | <i>Brugia malayi</i> ubiquinol-cytochrome c oxidoreductase core subunit 1 |
| IV | DI00132.cl | 11 | CAD61191 | 361.3 | 94.0 | <i>D. immitis</i> cytochrome c oxidase subunit I |
| | DI00329.cl | 4 | CAD61198 | 336.0 | 93.6 | <i>D. immitis</i> cytochrome c oxidase subunit II |
| | DI01488.cl | 3 | CAD61191 | 188.5 | 95.1 | <i>D. immitis</i> cytochrome c oxidase subunit I |
| | DI01662.cl | 1 | CAD61194 | 181.4 | 90.0 | <i>D. immitis</i> cytochrome c oxidase subunit III |
| | DI01252.cl | 1 | AAX22168 | 42.7 | 95.0 | <i>Melita nitida</i> cytochrome c oxidase subunit I |
| | DI02317.cl | 1 | AAT37154 | 70.2 | 65.3 | <i>Canis familiaris</i> cytochrome c oxidase assembly protein Cox17 |
| V | DI02377.cl | 2 | CAB03179 | 52.6 | 64.1 | <i>C. elegans</i> putative Mitochondrial ATP synthase subunit G |

^a Number of ESTs contained in each cluster.

^b Percent identity (%ID) at amino acid level.

these 282 *D. immitis* genes represent the identification, for the first time, of a specific gene group that may play important roles in filarial physiology.

3.3. Transcripts expressed abundantly in adult *D. immitis*

In general, high transcript abundance in a cDNA library correlates with high transcriptional levels in the sampled material, although there are exceptions (Audic and Claverie, 1997). The top 25 most abundant clusters in the adult heartworm (Table 1) contain 989 ESTs, representing 25% of all ESTs. Of these, three contained no detectable similarities to any NR or *B. malayi* proteins. Of the 22 clusters with matches, 21 had better similarity scores to *B. malayi* or other filarial proteins than against the NR database, with the only exception (DI00492.cl) containing weak similarities to both NR and *B. malayi* proteins. In addition, even if the *B. malayi* genes were excluded from the analysis, the top matches for 19 of the 22 most abundant clusters were to other nematode proteins, suggesting that the most abundant *D. immitis* transcripts are well conserved across many nematode species. Besides those clusters discussed in detail below (Sections 3.3.1 and 3.3.2), the abundant transcripts also represent genes involved in: (i) cellular metabolism (DI00268.cl, DI00102.cl, DI00010.cl, DI00498.cl, and DI00451.cl), (ii) cellular organization (DI00534.cl, DI00041.cl, DI00008.cl), (iii) signal transduction (DI00492.cl, DI00017.cl and DI00336.cl), and (iv) other biological processes (Table 1).

3.3.1. Abundant *D. immitis* adult transcripts include genes encoding subunits of the mitochondrial electron transfer complex I

The most abundant transcript, DI00105.cl (144 ESTs), represents 3.6% of total ESTs. The 11 contigs from this cluster showed 50–62% identities, at the amino acid level, to the *C. elegans* hypothetical protein W01A8.4, the putative B15 subunit of NADH-ubiquinone oxidoreductase (complex I). Complex I is a large enzyme complex in mitochondria consisting of at least 34 polypeptide chains in mammals, serving as the entry point of the aerobic electron transfer chain linked to oxidative phosphorylation, which generates most of the adenosine 5'-triphosphate required by mammalian cells (Grigorieff, 1998; Friedrich and Bottcher, 2004). Interestingly, two other abundant clusters apparently encode subunits of the same complex (Table 1): DI00032.cl (33 ESTs), the previously identified complex I subunit 3 encoded in the *D. immitis* mitochondrial genome (Hu et al., 2003), and DI00488.cl (19 ESTs). Furthermore, we were able to identify at least six other clusters (20 ESTs; Table 2a) encoding mitochondria-encoded complex I subunits 1, 5, 6 and one more nucleus-encoded subunit. Overall, 216 ESTs (5.4% of all ESTs) were identified to represent at least seven different subunits of the same protein complex, encoded in either mitochondrial or nuclear genomes, offering converging supports to their abundant expression and implying a critical function for this electron transfer component in adult heartworm.

Adult *D. immitis*, like many other adult filarial parasitic nematodes, was thought to be a homolactic fermenter (Kohler,

1991; Tielens, 1994), generating the majority of its energy through an anaerobic glycolytic pathway in the cytosol by fermenting the host blood glucose mainly to lactate (Hutchison and McNeill, 1970; Jaffe and Doremus, 1970). It was also shown that lactate accounted for only 55% of the carbon utilized from 1-¹⁴C-glucose in an in vitro anaerobic adult incubation (Hutchison and Turner, 1979), suggesting the adult parasite may also rely on other, more efficient, energy generation. To further explore the energy mechanism in adult heartworm, we searched for ESTs containing similarities to other electron transfer complexes (Table 2a). A smaller number of ESTs (30), encoding nine polypeptides in total were identified (two for complex II, two for complex III, four for complex IV, one for complex V), in contrast to the higher abundance of complex I subunits (216 ESTs). Indeed, the mammalian-type aerobic electron transfer-based energy generation is unlikely to function in the adult heartworm. Firstly, worms have no circulatory system and respiratory organs and adult heartworms are probably too large to maintain an aerobic metabolism through simple diffusion. Second, many of the tricarboxylic acid cycle enzymes have low activity levels inside adult heartworm mitochondria (McNeill and Hutchison, 1971). Together with this evidence, our observations argue for an alternative energy generation mechanism, relying heavily on the electron transfer complex I, in adult *D. immitis*.

Interestingly, studies on the adult stages of the clade III pig intestine parasitic nematode *Ascaris suum* have revealed such an energy generation mechanism, the malate dismutation pathway or the PEPCK-succinate pathway, that requires the complexes I and II but not the others (Tielens and Van Hellemond, 1998; Kita and Takamiya, 2002; Tielens et al., 2002; van Hellemond et al., 2003). A developmental switch of energy generation occurs in *A. suum* at around stage L3, wherein an anaerobic electron transfer pathway in adults replaces the mammalian-type aerobic electron transfer in embryos and larvae. Among the changes during this switch, a drastic decrease of the enzymatic activities and the expression levels of complexes III and IV in adult *A. suum* (Kita and

Takamiya, 2002) resembles closely the sampling pattern we observed in adult *D. immitis*, even though our sample may contain a number of developing embryos and larvae. In addition, two biomedical benchmarks for the anaerobic pathway were shown to be similar in the adults of these two parasites (Kuramochi et al., 1995). Furthermore, phylogenetic analysis suggested that parasitism in nematodes arose independently at least four times during evolution from the related free-living nematodes, with ascarids and filarial parasites evolving from the same branch (Blaxter et al., 1998). Therefore, the anaerobic electron transfer may have been acquired by a common ancestor during evolution and retained in both *D. immitis* and *A. suum* adults. Further gene expression profiling and biochemical studies will be needed to test our hypothesis in adult heartworm. A unique energy metabolism pathway in the adult parasite distinct from that of the host would offer a promising target for next generation chemotherapy.

3.3.2. Abundant *D. immitis* adult transcripts and antioxidant defenses

Blood-feeding parasites such as adult *D. immitis* may require a high level of antioxidant defenses (Oliveira and Oliveira, 2002). Blood in the host contains large quantities of free inorganic iron and the common biological iron chelator haem, both of which are known to be potent generators of reactive oxygen species (Tappel, 1955; Halliwell and Gutteridge, 1999). In addition, cytotoxic responses by the host immune system also contribute to the formation of damaging free radicals (Selkirk et al., 1998; Maizels et al., 2001), making oxidative stress an obvious threat to parasites such as adult *D. immitis*. Currently, at least two interrelated strategies have been discovered in diminishing oxidative stress in general: antioxidant enzymes to detoxify damaging free radicals directly and molecular chaperones, such as small heat shock proteins and Hsp70, to protect target proteins by covering sensitive sites or helping with refolding (Papp et al., 2003). Interestingly, we found abundant transcripts encoding

Table 2b
Antioxidation-related transcripts in adult *Dirofilaria immitis* containing 110 expressed sequence tags (ESTs)

| Protein | Cluster | #EST ^a | Best NR Hit | Bit Score | %ID ^b | Description |
|------------|------------|-------------------|-------------|-----------|---|--|
| Enzyme | DI00096.cl | 19 | AAB68798 | 368.3 | 97.9 | <i>D. immitis</i> peroxidoxin-1, thioredoxin peroxidase |
| | DI00117.cl | 9 | CAD38524 | 154.0 | 55.7 | <i>Globodera rostochiensis</i> putative glutathione peroxidase |
| | DI02379.cl | 1 | AAB61472 | 231.1 | 88.4 | <i>D. immitis</i> cytosolic Cu–Zn superoxide dismutase |
| Chaperone | DI00002.cl | 51 | CAA48632 | 299.7 | 91.2 | <i>Onchocerca volvulus</i> OV25-1, small heat shock protein |
| | DI00023.cl | 17 | AAD13154 | 453.9 | 99.2 | <i>Setaria digitata</i> heat shock protein 70 |
| | DI00396.cl | 4 | AAB05072 | 78.6 | 88.0 | <i>Caenorhabditis elegans</i> Cct-1, tcp-1 alpha subunit, chaperonin |
| | DI01129.cl | 2 | CAA06695 | 347.2 | 98.9 | <i>Brugia pahangi</i> heat shock protein 90 |
| | DI00926.cl | 1 | AAD03405 | 377.8 | 98.9 | <i>D. immitis</i> calreticulin, Ca-binding chaperone |
| | DI01259.cl | 1 | CAA06695 | 253.9 | 95.0 | <i>B. pahangi</i> heat shock protein 90 |
| | DI01743.cl | 1 | AAC17926 | 130.4 | 70.0 | <i>Brugia malayi</i> heat shock protein 70 |
| | DI02266.cl | 1 | CAA06695 | 128.3 | 98.6 | <i>B. pahangi</i> heat shock protein 90 |
| | DI02066.cl | 1 | P29779 | 76.2 | 46.0 | <i>O. volvulus</i> OV25-2, small heat shock protein |
| | DI01149.cl | 1 | CAA62240 | 75.8 | 40.5 | <i>Geobacillus stearothermophilus</i> dnaJ, heat shock protein |
| DI02315.cl | 1 | AAM22058 | 68.4 | 76.0 | <i>C. elegans</i> Cct-6, tcp-1 zeta subunit, chaperonin | |

^a Number of ESTs contained in each cluster.

^b Percent identity (%ID) at amino acid level.

Table 3
Selected metabolic pathways in adult *Dirofilaria immitis* mapped by Kyoto Encyclopedia of Genes and Genomes (KEGG)

| KEGG pathways ^a | <i>D. immitis</i> | | <i>C. elegans</i> , <i>Enzymes</i> | Coverage (<i>Di/Ce</i>) ^b (%) |
|--|-------------------|---------|---------------------------------------|---|
| | Clusters | Enzymes | | |
| Oxidative phosphorylation | 22 | 7 | 11 | 63.6 |
| Glycolysis/gluconeogenesis | 12 | 11 | 24 | 45.8 |
| Bile acid biosynthesis | 8 | 6 | 14 | 42.9 |
| Galactose metabolism | 7 | 5 | 13 | 38.5 |
| Limonene and pinene degradation | 6 | 5 | 13 | 38.5 |
| Biosynthesis of steroids | 4 | 5 | 15 | 33.3 |
| Phenylalanine metabolism | 6 | 4 | 12 | 33.3 |
| Histidine metabolism | 10 | 5 | 16 | 31.3 |
| Starch and sucrose metabolism | 24 | 9 | 30 | 30.0 |
| Reductive carboxylate cycle (CO ₂ fixation) | 3 | 3 | 10 | 30.0 |
| Fructose and mannose metabolism | 7 | 7 | 24 | 29.2 |
| Beta-Alanine metabolism | 2 | 1 | 13 | 7.7 |
| Carbon fixation | 1 | 1 | 14 | 7.1 |
| Arginine and proline metabolism | 4 | 2 | 28 | 7.1 |
| Aminosugars metabolism | 1 | 1 | 15 | 6.7 |
| Nitrogen metabolism | 1 | 1 | 19 | 5.3 |
| Glutamate metabolism | 2 | 1 | 26 | 3.8 |
| Fatty acid biosynthesis (path 1) | 0 | 0 | 10 | 0.0 |
| Urea cycle and metabolism of amino groups | 0 | 0 | 15 | 0.0 |

^a Pathways with 10% above (top) and 10% under (bottom) the mean coverage ratio.

^b Coverage for each pathway calculated as the ratio between numbers of mapped enzymes in *D. immitis* and *Caenorhabditis elegans*.

proteins of both groups in adult heartworm (DI00002.cl, small heat shock protein; DI00096.cl, thioredoxin peroxidase-1; Table 1). We were also able to identify transcripts coding for two other antioxidant enzymes (glutathione peroxidase and cytosolic Cu–Zn superoxide dismutase) as well as many other chaperones/heat shock proteins (Table 2b). In total, 14 clusters (110 ESTs), representing 2.8% of total transcripts, are likely involved in handling oxidative stress by adult heartworm, confirming the importance of its antioxidant mechanism.

3.4. *Wolbachia* transcripts in adult *D. immitis*

The endosymbiont *Wolbachia* has been found in the vast majority of filarial nematodes including *D. immitis*, while it appears to be absent in some filarial species such as *Loa loa* (Buttner et al., 2003) and non-filarial worms (Bandi et al., 1998; Fenn and Blaxter, 2004). To see if *Wolbachia* sequences are also present in our survey, *D. immitis* clusters were queried against the complete gene-set of the *Wolbachia* strain TRS of *B. malayi* (Foster et al., 2005). Eight clusters were shown to contain similarities with the *Wolbachia* genes (40–90% protein sequence identities; data not shown). The cluster with the best similarity score (DI00139.cl) shares 92% identity at the amino acid level with *Wolbachia* protein AAW71254, the E1 subunit of its pyruvate/2-2-oxoglutarate dehydrogenase. This similarity is also the top match when querying DI00139.cl against the NR database (data not shown), suggesting it represents the ortholog of AAW71254 from *Wolbachia* strain in *D. immitis*. Further studies are needed to confirm the endogenous *Wolbachia* origin.

3.5. Functional categorization and other sequence analysis

Both *D. immitis* clusters and, for comparison, *C. elegans* complete gene-set (WORMPEP v140) were mapped to the Gene Ontology (GO) terms (The Gene Ontology Consortium, 2000; Mulder et al., 2005). In total, GO terms were assigned to 341 (19%) *D. immitis* clusters and 12,681 (64%) *C. elegans* genes, respectively. As expected in surveying partial genomes, the unique mappings in *D. immitis* are drastically fewer than those for the complete *C. elegans* genome. Nevertheless, the two species have similar GO mapping distributions at the top two levels of the GO hierarchy (Supplementary Fig. 1), when the percentages of unique mappings in each species are compared. Complete listing of all *D. immitis* assignments can be viewed through the AmiGo browser at <http://www.nematode.net> (Wylie et al., 2004).

As an alternative way to classify biochemical functions, *D. immitis* clusters, as well as *C. elegans* genes, were assigned to metabolic pathways using the Kyoto Encyclopedia of Genes and Genomes (Bono et al., 1998; Kanehisa and Goto, 2000; Kanehisa et al., 2004). For the 87 metabolic pathways present in *C. elegans*, 215 (12%) total heartworm clusters were mapped to 200 metabolic enzymes with unique enzyme commission (EC) numbers. Comparing with the total 1042 metabolic enzymes mapped in *C. elegans*, an overall relative coverage of 19% (200/1042) was obtained (Supplementary Tables 2). By calculating such *D. immitis*/*C. elegans* coverage ratio for each of the 49 pathways in which 10 or more enzymes were mapped in *C. elegans*, we found eight pathways with lower coverage in *D. immitis*, and 11 pathways with higher

representations in heartworm (Table 3), when an arbitrary range of 10% under and over the mean ratio was used as an empirical cut-off. Interestingly, the top two highest coverage ratios belong to the oxidative phosphorylation pathway and the glycolysis/gluconeogenesis pathway. Since both pathways contain many enzymes involved in the anaerobic electron transfer discussed above, this observation is consistent with our hypothesis on the adult energy generation mechanism. The difference between the metabolic pathway representations in the two species may reflect the evolutionary adaptations of metabolism for their different life-styles (parasitic vs. free-living) or different environments. Complete listing of all KEGG mappings including graphical representation is available at <http://www.nematode.net> (Wylie et al., 2004).

RNA interference (RNAi) technology was first developed and has been successfully applied to genome-wide gene silencing to reveal gene functions in *C. elegans* (Fire et al., 1998; Kamath et al., 2003; Rual et al., 2004; Sonnichsen et al., 2005). *Caenorhabditis elegans* RNAi phenotypes can be further extrapolated in understanding functions of orthologous genes in other nematodes, especially the parasitic nematodes in which complexity of obligate parasitic life cycles and lack of proper laboratory culture system makes high-throughput screening impractical (Mitreva et al., 2005). In *D. immitis*, 55% of the most abundant (Table 1) and 75% of the most conserved (Supplementary Tables 1) transcripts have *C. elegans* homologs with detected phenotypes when silenced in RNAi experiments. Compared with the roughly 22% possibility of observing RNAi phenotypes in *C. elegans* at its whole genome level (reviewed by Mitreva et al., 2005), we confirmed the correlation between presence of the RNAi phenotypes and sequence conservation (and possibly levels of expression) as noticed in several other parasitic nematodes previously (McCarter et al., 2003; Mitreva et al., 2004a).

Secretion of proteins that are capable of modifying or manipulating the host environment is one of the adaptations during parasitism evolution to permit survival of parasites in the potentially hostile host (Harcus et al., 2004; Vanholme et al., 2005). Of the 1793 *D. immitis* clusters, we were able to identify putative signal peptides for secretion in 36 clusters and nine of these were further predicted as putative secreted proteins after meeting our criteria (see Section 2). Two of these (DI00012.cl and DI00588) do not contain amino acid similarities to any sequence from other species. DI00011.cl belongs to the putative filarial-specific gene group discussed above and DI00344.cl contains similarities shared by filarial and non-filarial parasites but not found in *C. elegans*, *C. briggsae* or non-nematodes. Only three clusters (DI00144.cl, DI01826.cl, and DI00202.cl) contain similarities to genes from *C. elegans* and *C. briggsae*. Interestingly, none of their best matches in the free-living worms (C24F3.2, a dual-specificity protein phosphatase; C38C6.2, alpha amylase; and CBG06129, a hypothetical protein) was predicted to be secreted (data not shown), suggesting possible adaptation for secretion by pre-existing genes in the evolution of parasitism. In addition, studies in the blood-feeding canine hookworm

(*Ancylostoma caninum*) have revealed a multi-enzyme cascade of hemoglobin proteolysis as an important source of its nutrition (Williamson et al., 2004). We were able to identify, for the first time, a putative serine protease (DI00123.cl) and a putative metalloprotease (DI00551.cl) in heartworm, in which the ingestion of host blood was reported to be active (McGonigle et al., 2001). More gene sequences from heartworm and proper biochemical studies will be required to explore whether such a proteolysis pathway is functional in *D. immitis* and whether these two putative proteases are involved in it.

4. Conclusions

The current study has provided the largest set of sequence information for *D. immitis* to date, greatly expanding our knowledge of the heartworm transcriptome. As demonstrated previously (Mitreva et al., 2005), the identification of about 1800 genes from the parasite underlines the power of EST survey when studying a large number of species in Nematoda. Our initial comparative genomic analyses, based on the numbers and strength of the matches, is consistent with the phylogenetic relationships between *D. immitis* and other nematodes such as *B. malayi* and metazoans. More importantly, we generated a group of candidate filarial-specific genes that could play specific roles during parasitism and pathology, making them a salient focus for future studies. Furthermore, identification of abundant transcripts involved in energy generation and anti-oxidation suggests critical functions for these pathways in adult physiology. Our hypothesis of an alternative energy generation mechanism in the adult heartworm based on non-mammalian type electron transfer, upon further confirmation by gene expression profiling and biochemical studies, may offer promising targets for safer next generation macrofilaricides and treatments in eliminating the adult parasites without interfering with the host metabolism. Lastly, the GO and KEGG classifications revealed putative functions for many *D. immitis* genes, offering a functional view of the adult transcriptome. Overall, the current survey provides a solid foundation for further studies in *D. immitis*, such as larger-scale EST surveys for adult and other stages by constructing specific cDNA libraries to identify more *D. immitis* genes, gene expression profiling using microarray technology to better understand the molecular basis of its developmental changes and parasitism, molecular and biochemical characterizations of gene functions in heartworm, etc. Information obtained from heartworm also offers great opportunities for further comparative genomic studies with human filarial nematodes such as *B. malayi*.

Acknowledgements

The *D. immitis* EST project at Washington University was supported by NIH-NIAID research grant AI 46593 to RKW. The project originated while JPM was a Merck Fellow of the Helen Hay Whitney Foundation. We thank Dr Prema Arasu at

North Carolina State University for providing worms, Dr John Spieth and Dr Michael Crawford for scientific discussions, Merry McLaird, Claire Murphy, Mike Dante and Sahar Abubucker for technical assistance and The Institute for Genomic Research for the use of the *B. malayi* genome assembly. The current research was not funded by Divergence Inc., of which JPM is an employee and equity holder. This work was supported by Wellcome Trust, programme grant GR066764.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ijpara.2006.03.002.

References

- Audic, S., Claverie, J.M., 1997. The significance of digital gene expression profiles. *Genome Res.* 7, 986–995.
- Bandi, C., Anderson, T.J., Genchi, C., Blaxter, M.L., 1998. Phylogeny of *Wolbachia* in filarial nematodes. *Proc. R. Soc. Lond. B. Biol. Sci.* 265, 2407–2413.
- Blaxter, M.L., De Ley, P., Garey, J.R., Liu, L.X., Scheldeman, P., Vierstraete, A., Vanfleteren, J.R., Mackey, L.Y., Dorris, M., Frisse, L.M., Vida, J.T., Thomas, W.K., 1998. A molecular evolutionary framework for the phylum Nematoda. *Nature* 392, 71–75.
- Blaxter, M., Aslett, M., Guiliano, D., Daub, J., 1999. Parasitic helminth genomics. Filarial genome project. *Parasitology* 118, S39–S51.
- Blaxter, M.L., Daub, J., Guiliano, D., Parkinson, J., Whitton, C., Project, F.G., 2002. The *Brugia malayi* genome project: expressed sequence tags and gene discovery. *Trans. R. Soc. Trop. Med. Hyg.* 96, 7–17.
- 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.
- Buttner, D.W., Wanji, S., Bazzocchi, C., Bain, O., Fischer, P., 2003. Obligatory symbiotic *Wolbachia* endobacteria are absent from *Loa loa*. *Filar. J.* 2, 10.
- Emanuelsson, O., Nielsen, H., Brunak, S., von Heijne, G., 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* 300, 1005–1016.
- Fenn, K., Blaxter, M., 2004. Quantification of *Wolbachia* bacteria in *Brugia malayi* through the nematode lifecycle. *Mol. Biochem. Parasitol.* 137, 361–364.
- Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., Mello, C.C., 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806–811.
- Foster, J., Ganatra, M., Kamal, I., Ware, J., Makarova, K., Ivanova, N., Bhattacharyya, A., Kapatral, V., Kumar, S., Posfai, J., Vincze, T., Ingram, J., Moran, L., Lapidus, A., Omelchenko, M., Kyrpides, N., Ghedin, E., Wang, S., Goltsman, E., Joukov, V., Ostrovskaya, O., Tsukerman, K., Mazur, M., Comb, D., Koonin, E., Slatko, B., 2005. The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode. *Plos. Biol.* 3, e121.
- Friedrich, T., Bottcher, B., 2004. The gross structure of the respiratory complex I: a lego system. *Biochim. Biophys. Acta* 1608, 1–9.
- Ghedini, E., Wang, S., Foster, J.M., Slatko, B.E., 2004. First sequenced genome of a parasitic nematode. *Trends Parasitol.* 20, 151–153.
- Grigorieff, N., 1998. Three-dimensional structure of bovine NADH:ubiquinone oxidoreductase (complex I) at 2.2 Å in ice. *J. Mol. Biol.* 277, 1033–1046.
- Halliwell, B., Gutteridge, J.M.C., 1999. Free radicals in biology and medicine. Oxford University Press, Oxford.
- Harcus, Y.M., Parkinson, J., Fernandez, C., Daub, J., Selkirk, M.E., Blaxter, M.L., Maizels, R.M., 2004. Signal sequence analysis of expressed sequence tags from the nematode *Nippostrongylus brasiliensis* and the evolution of secreted proteins in parasites. *Genome Biol.* 5, R39.
- Hu, M., Gasser, R.B., Abs El-Osta, Y.G., Chilton, N.B., 2003. Structure and organization of the mitochondrial genome of the canine heartworm, *Dirofilaria immitis*. *Parasitology* 127, 37–51.
- Hutchison, W.F., McNeill, K.M., 1970. Glycolysis in the adult dog heartworm, *Dirofilaria immitis*. *Comp. Biochem. Physiol.* 35, 721–727.
- Hutchison, W.F., Turner, A.C., 1979. Glycolytic end products of the adult dog heartworm, *Dirofilaria immitis*. *Comp. Biochem. Physiol.* B 62, 71–73.
- Jaffe, J.J., Doremus, H.M., 1970. Metabolic patterns of *Dirofilaria immitis* microfilariae in vitro. *J. Parasitol.* 56, 254–260.
- Kamath, R.S., Fraser, A.G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D.P., 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.
- Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., Hattori, M., 2004. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 32, D277–D280.
- Kita, K., Takamiya, S., 2002. Electron-transfer complexes in *Ascaris* mitochondria. *Adv. Parasitol.* 51, 95–131.
- Knight, D.H., 1977. Heartworm heart disease. *Adv. Vet. Sci. Comp. Med.* 21, 107–149.
- Knight, D.H., 1987. Heartworm infection. *Vet. Clin. North Am. Small Anim. Pract.* 17, 1463–1518.
- Kohler, P., 1991. The pathways of energy generation in filarial parasites. *Parasitol. Today* 7, 21–25.
- Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580.
- Kuramochi, T., Kita, K., Takamiya, S., Kojima, S., Hayasaki, M., 1995. Comparative study and cDNA cloning of the flavoprotein subunit of mitochondrial complex II (succinate-ubiquinone oxidoreductase: fumarate reductase) from the dog heartworm, *Dirofilaria immitis*. *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.* 111, 491–502.
- Maizels, R.M., Blaxter, M.L., Scott, A.L., 2001. Immunological genomics of *Brugia malayi*: filarial genes implicated in immune evasion and protective immunity. *Parasite Immunol.* 23, 327–344.
- McCall, J.W., 2005. The safety-net story about macrocyclic lactone heartworm preventives: a review, an update, and recommendations. *Vet. Parasitol.* 133, 197–206.
- McCarter, J., Dautova Mitreva, M., Martin, J., Dante, M., Wylie, T., Rao, U., Pape, D., Bowers, Y., Theising, B., Murphy, C.V., Kloek, A.P., Chiapelli, B.J., Clifton, S.W., Bird, M.D., Waterston, R., 2003. Analysis and functional classification of transcripts from the Nematode *Meloidogyne incognita*. *Genome Biol.* 4 (R26), 1–19.
- McGonigle, S., Yoho, E.R., James, E.R., 2001. Immunisation of mice with fractions derived from the intestines of *Dirofilaria immitis*. *Int. J. Parasitol.* 31, 1459–1466.
- McNeill, K.M., Hutchison, W.F., 1971. The tricarboxylic acid cycle enzymes in the adult dog heartworm, *Dirofilaria immitis*. *Comp. Biochem. Physiol.* 38B, 493–500.
- McTier, T.L., McCall, J.W., Dzimianski, M.T., Raynaud, J.P., Strickland, J.E., 1994. Use of melarsomine dihydrochloride (RM 340) for adulticidal treatment of dogs with naturally acquired infections of *Dirofilaria immitis* and for clinical prophylaxis during reexposure for 1 year. *Vet. Parasitol.* 55, 221–233.
- Mitreva, M., Elling, A.A., Dante, M., Kloek, A.P., Kalyanaraman, A., Aluru, S., Clifton, S.W., Bird, D.M., Baum, T.J., McCarter, J.P., 2004a. A survey of SL1-spliced transcripts from the root-lesion nematode *Pratylenchus penetrans*. *Mol. Genet. Genomics* 272, 138–148.
- Mitreva, M., McCarter, J.P., Martin, J., Dante, M., Wylie, T., Chiapelli, B., Pape, D., Clifton, S.W., Nutman, T.B., Waterston, R.H., 2004b. Comparative genomics of gene expression in the parasitic and free-living nematodes *Strongyloides stercoralis* and *Caenorhabditis elegans*. *Genome Res.* 14, 209–220.

- Mitreva, M., Blaxter, M.L., Bird, D.M., McCarter, J.P., 2005. Comparative genomics of nematodes. *Trends Genet.* 21, 573–581.
- Molyneux, D.H., Bradley, M., Hoerauf, A., Kyelem, D., Taylor, M.J., 2003. Mass drug treatment for lymphatic filariasis and onchocerciasis. *Trends Parasitol.* 19, 516–522.
- Mulder, N.J., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Binns, D., Bradley, P., Bork, P., Bucher, P., Cerutti, L., Copley, R., Courcelle, E., Das, U., Durbin, R., Fleischmann, W., Gough, J., Haft, D., Harte, N., Hulo, N., Kahn, D., Kanapin, A., Krestyaninova, M., Lonsdale, D., Lopez, R., Letunic, I., Madera, M., Maslen, J., McDowall, J., Mitchell, A., Nikolskaya, A.N., Orchard, S., Pagni, M., Ponting, C.P., Quevillon, E., Selengut, J., Sigrist, C.J., Silventoinen, V., Studholme, D.J., Vaughan, R., Wu, C.H., 2005. InterPro, progress and status in 2005. *Nucleic Acids Res.* 33, D201–D205 (Database Issue).
- Oliveira, P.L., Oliveira, M.F., 2002. Vampires, Pasteur and reactive oxygen species. Is the switch from aerobic to anaerobic metabolism a preventive antioxidant defence in blood-feeding parasites? *Fed. Eur. Biochem. Soc. Lett.* 525, 3–6.
- Papp, E., Nardai, G., Soti, C., Csermely, P., 2003. Molecular chaperones, stress proteins and redox homeostasis. *Biofactors* 17, 249–257.
- Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., Lopez, R., 2005. InterProScan: protein domains identifier. *Nucleic Acids Res.* 33, W116–W120.
- Rual, J.F., Ceron, J., Koreth, J., Hao, T., Nicot, A.S., Hirozane-Kishikawa, T., Vandenhaute, J., Orkin, S.H., Hill, D.E., van den Heuvel, S., Vidal, M., 2004. Toward improving *Caenorhabditis elegans* phenome mapping with an ORFeome-based RNAi library. *Genome Res.* 14, 2162–2168.
- Selkirk, M.E., Smith, V.P., Thomas, G.R., Gounaris, K., 1998. Resistance of filarial nematode parasites to oxidative stress. *Int. J. Parasitol.* 28, 1315–1332.
- Sonnichsen, B., Koski, L.B., Walsh, A., Marschall, P., Neumann, B., Brehm, M., Alleaume, A.M., Artelt, J., Bettencourt, P., Cassin, E., Hewitson, M., Holz, C., Khan, M., Lazik, S., Martin, C., Nitzsche, B., Ruer, M., Stamford, J., Winzi, M., Heinkel, R., Roder, M., Finell, J., Hantsch, H., Jones, S.J., Jones, M., Piano, F., Gunsalus, K.C., Oegema, K., Gonczyk, P., Coulson, A., Hyman, A.A., Echeverri, C.J., 2005. Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. *Nature* 434, 462–469.
- Stein, L.D., Bao, Z., Blasiar, D., Blumenthal, T., Brent, M.R., Chen, N., Chinwalla, A., Clarke, L., Clee, C., Coghlan, A., 2003. The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *Plos. Biol.* 1, E45.
- Tappel, A.L., 1955. Unsaturated lipid oxidation catalyzed by hematin compounds. *J. Biol. Chem.* 217, 721–733.
- The *C. elegans* Sequencing Consortium, 1998. Genome sequence of the Nematode *C. elegans*: a platform for investigating biology. *Science* 282, 2012–2018.
- The Gene Ontology Consortium, 2000. Gene ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29.
- Tielens, A.G., 1994. Energy generation in parasitic helminths. *Parasitol. Today* 10, 346–352.
- Tielens, A.G., Van Hellemond, J.J., 1998. The electron transport chain in anaerobically functioning eukaryotes. *Biochim. Biophys. Acta* 1365, 71–78.
- Tielens, A.G., Rotte, C., van Hellemond, J.J., Martin, W., 2002. Mitochondria as we don't know them. *Trends Biochem. Sci.* 27, 564–572.
- van Hellemond, J.J., van der Klei, A., van Weelden, S.W., Tielens, A.G., 2003. Biochemical and evolutionary aspects of anaerobically functioning mitochondria. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 358, 205–213.
- Vanholme, B., Mitreva, M., Van Criekinge, W., Logghe, M., Bird, D., McCarter, J.P., Gheysen, G., 2005. Detection of putative secreted proteins in the plant-parasitic nematode *Heterodera schachtii*. *Parasitol. Res.* 28, 1–11.
- Whitton, C., Daub, J., Quail, M., Hall, N., Foster, J., Ware, J., Ganatra, M., Slatko, B., Barrell, B., Blaxter, M., 2004. A genome sequence survey of the filarial nematode *Brugia malayi*: repeats, gene discovery, and comparative genomics. *Mol. Biochem. Parasitol.* 137, 215–227.
- Williams, S.A., Lizotte-Waniewski, M.R., Foster, J., Guiliano, D., Daub, J., Scott, A.L., Slatko, B., Blaxter, M.L., 2000. The filarial genome project: analysis of the nuclear, mitochondrial and endosymbiont genomes of *Brugia malayi*. *Int. J. Parasitol.* 30, 411–419.
- Williams, S.A., Laney, S.J., Lizotte-Waniewski, M., Bierwert, L.A., 2002. The Riverblindness genome project. *Trends Parasitol.* 18, 86–90.
- Williamson, A.L., Lecchi, P., Turk, B.E., Choe, Y., Hotez, P.J., McKerrow, J.H., Cantley, L.C., Sajid, M., Craik, C.S., Loukas, A., 2004. A multi-enzyme cascade of hemoglobin proteolysis in the intestine of blood-feeding hookworms. *J. Biol. Chem.* 279, 35950–35957.
- Wylie, T., Martin, J., Dante, M., Mitreva, M., Clifton, S.W., Chinwalla, A., Waterston, R.H., Wilson, R.K., McCarter, J.P., 2004. Nematode.net: a tool for navigating sequences from parasitic and free-living nematodes. *Nucleic Acids Res.* 32, D423–D426.