Structure and organization of the mitochondrial genome of the canine heartworm, *Dirofilaria immitis*

M. HU, R. B. GASSER*, Y. G. ABS EL-OSTA *and* N. B. CHILTON

Department of Veterinary Science, The University of Melbourne, 250 Princes Highway, Werribee, Victoria 3030, Australia

(*Received 28 November 2002; revised 22 January 2003; accepted 22 January 2003*)

SUMMARY

This study determined the complete mitochondrial (mt) genome sequence of the canine heartworm, *Dirofilaria immitis*†, and compared its structure, organization and other characteristics with *Onchocerca volvulus* and other secernentean nematodes. The *D. immitis* mt genome is 13814 bp in size and contains 36 of the 37 genes typical of metazoan organisms, and lacks the ATP synthetase subunit 8 gene. All of the genes are transcribed in the same direction. For the entire genome, the nucleotide contents are ~55% (T), ~19% (each for A and G) and ~7% (C), which is very similar to those of the protein-coding genes. In the latter genes, most (~69%) third codon positions have a T, but rarely (~1–9%) have an A or a C. The C content (8–12%) is higher at the first and second codon positions compared with the third position (~1%). These nucleotide biases have a significant effect on the codon usage patterns and, thus, on the amino acid composition of the proteins. The mt genome organization of *D. immitis* is essentially the same as that of *O. volvulus*, but is distinctly different from other secernentean nematodes sequenced thus far. Irrespective of transpositions of transfer RNA (*trn*) genes and the non-coding, AT-rich region, there are 4 gene- or gene block-translocations between the mt genome of *D. immitis* and those of *Caenorhabditis elegans*, *Ascaris suum* and the 2 human hookworms, *Ancylostoma duodenale* and *Necator americanus*. For *D. immitis*, the 22 *trn* genes have secondary structures typical of other secernentean nematodes, and possess a TV-replacement loop instead of a TΨC arm and loop. Like *O. volvulus*, the mt *trn*K and *trn*P of *D. immitis* use the anticodons CUU and AGG, whereas in other nematodes, UUU and UGG are employed, respectively. Also, the secondary structures of the 2 ribosomal RNA (*rrn*) genes are similar to the models for other nematodes. Overall, the availability of the complete *D. immitis* mt genome sequence provides a resource for future studies of the comparative mt genomics and of the population genetics and/or phylogeny of parasitic nematodes.

Key words: *Dirofilaria immitis*, mitochondrial genome, mitochondrial genes, ribosomal RNA genes, transfer RNA genes.

INTRODUCTION

The majority of metazoan organisms, with few exceptions, possess circular mitochondrial (mt) genomes which usually vary in size from 14 to 19 kb (Boore, 1999; Saccone *et al*. 1999). Their mt genomes usually comprise genes for a complete set of transfer RNAs (*trn*), for 2 ribosomal RNAs (*rrn*) and for 12–13 protein subunits of the enzymes involved in oxidative phosphorylation, namely the subunits I, II and III of cytochrome oxidase (*cox*1, *cox*2 and *cox*3), the cytochrome b (*cob*), the subunits 6 and/or 8 of the ATPase complex (*atp*6 and/or *atp*8) and the subunits 1–6 and 4L of the NADH dehydrogenase (*nad*1–6 and *nad*4L) (Wolstenholme, 1992; Boore, 1999). There is also a non-coding, AT-rich (control or D-loop) region, which contains signalling elements for the regulation of replication and transcription (reviewed by Shadel & Clayton, 1997).

Over 200 mt genome sequences have been published for 11 metazoan phyla, but >150 represent chordates. Despite technological advances in mt genome sequencing, there are still major gaps in our knowledge of the mitochondrial genomics for the phylum Nematoda. Complete mt genome sequences have been determined only for 5 secernentean nematodes, including *Caenorhabditis elegans* (Rhabditida) (Okimoto *et al*. 1992), *Ascaris suum* (Ascaridida) (Okimoto *et al*. 1992), *Ancylostoma duodenale*, *Necator americanus* (Strongylida) (Hu, Chilton & Gasser, 2002) and *Onchocerca volvulus* (Spirurida) (Keddie, Higazi & Unnasch, 1998), and for the adenophorean nematode, *Trichinella spiralis* (Enoplida) (Lavrov & Brown, 2001). Current information for the class Secernentea reveals that the free-living nematode, *C. elegans*, and the parasitic nematodes *As. suum*, *An. duodenale* and *N. americanus* all possesses essentially the same mt genome arrangement (with the exception of the position of the AT-rich region) (Okimoto *et al*. 1992; Hu *et al*. 2002), whereas the filarioid, *O. volvulus*, has a distinctly different organization (Keddie *et al*. 1998).

While some workers have proposed that mt genome arrangements do not usually vary significantly within taxonomic groups (Boore, 1999), recent evidence demonstrates clearly that mt gene rearrangements can occur between species of the same family or even of the same genus (e.g. Le *et al*. 2000;

* Corresponding author. Tel: +61 3 9731 2283. Fax: +61 3 9731 2366. E-mail: robinbg@unimelb.edu.au

† The complete nucleotide sequence for the mitochondrial genome of *Dirofilaria immitis* is available from the EMBL database under the Accession Number AJ537512.

Parasitology (2003), **127**, 37–51.
DOI: 10.1017/S0031182003003275 Printed in the United Kingdom

Rawlings, Colline & Bieler, 2001). Given the paucity of mt genome information, these aspects have not yet been addressed for nematodes. In the present study, we establish the complete mt genome sequence of *D. immitis*, a parasite of veterinary importance, which belongs to the same order as *O. volvulus* but a distinct family. These 2 filarioid nematodes differ significantly in their transmission and predilection site in the host, and in the diseases they cause. While the former parasite is transmitted by mosquitoes and causes heartworm disease mainly in canids in subtropical and tropical climatic zones (e.g. Cringoli *et al*. 2001; Fan *et al*. 2001), the latter is transmitted by blackflies (of the *Simulium* species complex) and causes 'river blindness' (onchocerciasis) in humans in West Africa (Unnasch, 2002). We also compare the organization (and other characteristics) of the *D. immitis* mt genome with that of *O. volvulus* and the 5 other secernentean nematodes for which complete genome information is currently available.

MATERIALS AND METHODS

Parasite and DNA extraction

Adult worms of *D. immitis* were collected at necropsy from the pulmonary artery of a dog from Australia (Victoria), washed in physiological saline and then frozen at −70 °C until use. Total genomic DNA was extracted from (~2 cm) portions of an individual worm by sodium dodecyl-sulphate/proteinase K treatment (Gasser *et al*. 1993), column-purified (Wizard™ DNA Clean-up; Promega, Madison, WI, USA) and eluted in 40 μl of H_2O. The specific identity of the worm was verified using the sequence of the first internal transcribed spacer (ITS-1) of ribosomal DNA, which was compared with that reported recently for *D. immitis* (Accession number: AF217800) (Mar *et al*. 2002).

Long-PCR amplification, agarose gel electrophoresis and DNA sequencing

The entire mt genome of *D. immitis* was amplified by long-PCR (Expand™ 20 kb^{Plus} kit, Roche) from total genomic DNA from a single worm in 2 overlapping fragments (~5 kb and ~9 kb) using the 2 oligonucleotide primer sets COIF-MH4R and MH37F-MH28R, respectively (see Fig. 1). These primers were constructed to mt sequences which are relatively conserved among *As. suum*, *C. elegans* and *O. volvulus* (see Okimoto *et al*. 1992; Keddie *et al*. 1998). Primer MH28R was designed to the *cox*1 gene, while primers MH4R and MH37F were designed to the *rrn*S gene. Primer COIF was originally constructed from platyhelminths (see Bowles, Blair & McManus, 1992). Long-PCR amplification was performed according to the protocol of Hu *et al*. (2002): 92 °C for 2 min (initial denaturation); then 92 °C/10 s (denaturation), 50 °C/30 s (annealing), 68 °C or 60 °C/10 min (extension) for 10 cycles; followed by 92 °C/10 s, 50 °C/30 s; 68 °C or 60 °C/10 min for 20 cycles, with a cycle elongation of 10 s for each cycle, and a final extension at 68 °C or 60 °C/7 min. The extension temperatures of 68 °C and 60 °C were used specifically for primer sets MH37F-MH28R and COIF-MH4R, respectively. Amplicons were detected in 1% (w/v) agarose gels after ethidium-bromide staining and ultraviolet transillumination. Products were purified over mini-spin columns (Wizard™ PCR-Prep, Promega) and then used as templates in dye-terminator cycle-sequencing reactions according to the supplier's (Perkin Elmer) protocol, employing a primer walking strategy. The primers employed for sequencing and their relative positions in the mt genome are presented in Fig. 1. For some T-rich regions within the *D. immitis* mt genome, short regions were PCR-amplified and cloned into the plasmid vector pGEM-T™ Easy (Promega) and transformed into *Escherichia coli* JM109, according to the manufacturer's protocol. Following propagation and plasmid purification over columns (Wizard™ *Plus* SV Minipreps DNA, Promega), inserts were sequenced using the vector primers T7 and SP6 Promega.

Sequence analyses and characterization of the mt genome of D. immitis

Sequences were assembled manually and aligned with the complete mt genome sequence of *O. volvulus* (Accession number: AF015193; Keddie *et al*. 1998) using the program Clustal X (Thompson *et al*. 1997). Also, protein genes (designated according to Le, Blair & McManus, 2000), and translation initiation and termination codons were identified based on the comparison with those reported previously for *O. volvulus*. The inferred amino acid sequences and codon usage of protein genes were obtained using the computer program MacVector version 7.0 (Oxford Molecular Group). The amino acid sequences inferred for *D. immitis* were aligned with those of *C. elegans*, *As. suum* (Accession numbers: X54252 and X54253; Okimoto *et al*. 1992), *O. volvulus* (Keddie *et al*. 1998), *An. duodenale* and *N. americanus* (Accession numbers: AJ417718 and AJ417719; Hu *et al*. 2002) using Clustal X. Based on the alignment, amino acid sequence identity (%) to homologous genes was calculated. Most *trn* genes in the mt genome of *D. immitis* were identified using the tRNAscan program, available at http://www.queensu.ca/micr/faculty/kropinski/online.html (Lowe & Eddy, 1997), whereas others were identified by eye based on their potential *trn* secondary structures and/or anticodon sequences (Wolstenholme *et al*. 1987). The 2 *rrn* genes were identified based on their sequence similarity to those of *O. volvulus* (see Keddie *et al*. 1998) and their potential

https://doi.org/10.1017/S0031182003003275 Published online by Cambridge University Press

to form *rrn*-like secondary structures. The secondary structures of the *rrn* genes were inferred by analogy to previously published structures (see Hu *et al*. 2002). The stem-loop structures of non-coding regions were inferred using the Mfold program, available at http://www.queensu.ca/micr/faculty/kropinski/online.html (Santalucia, 1998).

RESULTS AND DISCUSSION

Sequencing

The ~5 kb and ~9 kb amplicons spanning the entire mt genome of *D. immitis* (Fig. 1) were subjected to sequencing. Over 90% of the genome could be sequenced directly from these amplicons, except for 2 regions containing 1–2 poly-T sequence tracts (see Fig. 1) located to the genes *nad*2 (1 with 18Ts) and *nad*6 (2 with 18–20 Ts), which had to be cloned and then sequenced to obtain readable sequence beyond these tracts. Hence, region 1 (~0·5 kb) and region 2 (~1·2 kb) containing the 3 poly-T tracts (Fig. 1) were amplified using primers Di19F-Di12R and Di49F-Di48R, respectively, and then cloned and sequenced, following conventional propagation and purification of the recombinant plasmids.

General features of the mt genome of D. immitis

The circular mt genome of *D. immitis* is 13814 bp in size (Fig. 2), which is amongst the smallest metazoan mt genomes sequenced to date, such as the platyhelminth *Taenia crassiceps* (13503 bp) (Le *et al*. 2002). It is 20–210 bp longer than the mt genomes of *C. elegans* (see Okimoto *et al*. 1992), *O. volvulus* (see Keddie *et al*. 1998), *An. duodenale* and *N. americanus* (see Hu *et al*. 2002), and 470 bp shorter than that of *As. suum* (see Okimoto *et al*. 1992). Other features of the mt genome of *D. immitis* and its products, including the positions, lengths and start/stop codons of individual genes as well as amino acid sequence lengths of predicted proteins, are shown in Table 1.

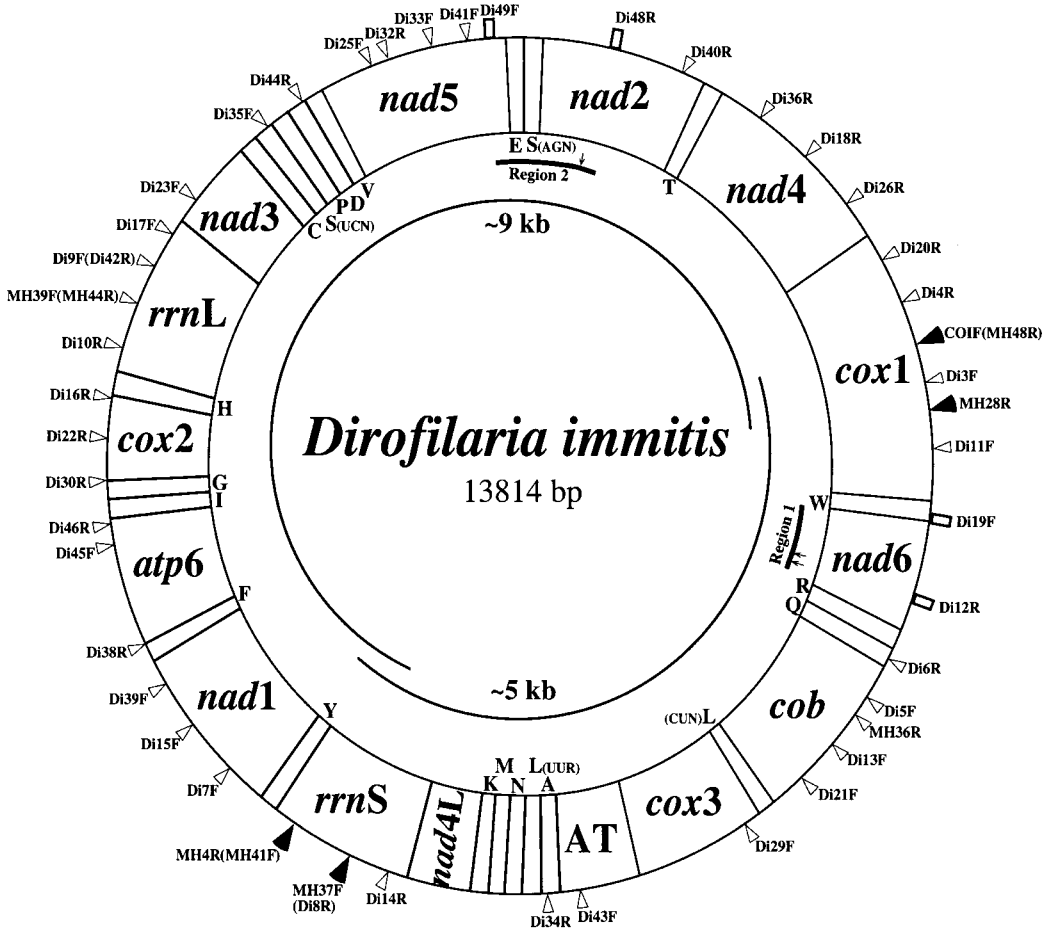
All of the 36 genes (12 protein-coding, 2 *rrn* and 22 *trn* genes) typically found in secernentean nematodes were identified in the mt genome of *D. immitis*. They are all transcribed in the same direction (Fig. 1), which is consistent with most other nematodes (Okimoto *et al*. 1992; Keddie *et al*. 1998; Hu *et al*. 2002), flatworms (Le *et al*. 2002), annelids, and some molluscs, brachiopods and cnidarians (see Boore, 1999). The *atp*8 gene is absent, consistent with other secernentean nematodes, flatworms (von Nickisch-Rosenegk, Brown & Boore, 2001; Le, Blair & McManus, 2002) and some molluscs (Hoffmann, Boore & Brown, 1992) but in contrast to *T. spiralis* (see Lavrov & Brown, 2001). Whether the *atp*8 gene absent from the mt genome of *D. immitis* is present in the nuclear genome remains to be established.

Consistent with the small size of the *D. immitis* mt genome, the genes are arranged in an 'economical fashion' in that most genes have no bases or only a few nucleotides between their coding regions, and there is evidence that 5 pairs of genes overlap by 1–21 nucleotides (Fig. 2). Specifically, (i) there is a 1-nucleotide overlap between *trn*E and *trn*S (AGN). Regarding such an overlap, study of RNA editing in vertebrates has shown that the downstream *trn* is released in its intact form, whereas the upstream *trn* is truncated and completed by editing (Yokobori & Pääbo, 1995, 1997); (ii) a 4-nucleotide overlap is present between *trn*S (AGN) and *nad*2; (iii) a 1-nucleotide overlap occurs between *trn*T and *nad*4; (iv) the last 2 nucleotides of *trn*Y overlap with the first 2 nucleotides of *nad*1; (v) there are 21 nucleotides shared between *nad*1 and *trn*F. An overlap of a similar length is also present in the mt genome of *O. volvulus* (see Keddie *et al*. 1998), but not in any other nematode species sequenced to date.

Mitochondrial genome organization

The organization of mt genome of *D. immitis* is the same as *O. volvulus* (the only other filarioid nematode for which mt sequence data have been published), but is distinctly different from other secernentean nematodes of different orders which all possess essentially the same genome organization based on the arrangements of protein-coding and *rrn* genes (and excluding the position of the AT-rich region) (Fig. 3). Discounting the numerous differences in the positions of *trn* genes and the AT-rich region, a minimum of 4 rearrangements is required to interconvert the mt gene arrangement of *D. immitis* with those of *C. elegans*, *As. suum*, and *An. duodenale* and *N. americanus* (Fig. 3). There are 2 gene- (*nad*2 and *nad*6) and 2 gene block- (*nad*4L-*rrn*S-*nad*1-*atp*6 and *nad*4-*cox*1) translocations (Fig. 3). Regarding gene rearrangements in animal mt genomes, several models have been proposed to explain the possible mechanisms, including the 'duplication/random loss model' (Moritz, Dowling & Brown, 1987; Macey *et al*. 1997) and the 'intra-mitochondrial recombination model' (Lunt & Hyman, 1997; Dowton & Campbell, 2001). Also, some other studies have focussed on the correlation between mt genome rearrangements and biological or molecular phenomena, such as the evolution of parasitism, high mt AT content, accelerated rate of mt genetic divergence as well as adaptive radiation (reviewed by Dowton, Castro & Austin, 2002). Interestingly, the *nad*2 and *nad*6 genes of *D. immitis* both contain relatively long poly-T (⩾18) tracts (see Fig. 1) (which originally caused considerable problems when subjected to direct sequencing). Biologically, such tracts may facilitate slipped-strand mispairing because of an increased chance of incidental homology between distant mt

https://doi.org/10.1017/S0031182003003275 Published online by Cambridge University Press



Primer	Sequence (5'-3')	Primer	Sequence (5'-3')
COIF	TTTTTTGGGCATCCTGAGGTTTAT	Di19F	CTGATGTTATTATTCTATGTG
MH4R	TCTACTTTACTACAACTTACTCC	Di20R	CCCAATCCTGCTCAATAACC
MH28R	CTAACTACATAATAAGTARCATG	Di21F	CTATTTTGCGTTCTGTTCCTG
MH36R	CCTCAAACTAAAACATAACC	Di22R	CATACTCATAACTCAATACCAC
MH37F	GGAGTAAAGTTGTATTTAAAC	Di23F	TTTGAGAGTAAAGTAAGATC
MH39F	TAAATGGCAGTCTTAGCGTGA	Di25F	GTTCATAGTAGTACTTTAGTTAC
MH41F	GGAGTAAGTTGTAGTAAAGTAGA	Di26R	GAACTAAATCAATTAAACATC
MH44R	TCACGCTAAGACTGCCATTTA	Di29F	ATAGTTATTATCCTTTGATG
MH48R	ATAAACCTCAGGATGCCCAAAAAA	Di30R	AAATAATTAGCCTACAAACC
Di3F	CTTTTTACTATTGGAGGTTTG	Di32R	TCCTACATAAAAAACAAAAG
Di4R	AAAAGTTCAACTACTCCCAGG	Di33F	TTTTGGTTCAGTTGCAAGTG
Di5F	AAGGGTTTAGTTTATGGTAGG	Di34R	ATATCTTATGAATGCAAATC
Di6R	AAATCTTTCGAACCAAAAAC	Di35F	ATCATTAGGAGTTTTAAGCC
Di7F	TTGACTTTGTTGGAGCGTCA	Di36R	CTATAAAAAACCAAACCAACC
Di8R	TAAACAAAACTTTACTCCCG	Di38R	CAACAACGCCTAAAAAAACC
Di9F	TAATATTTTCTTGGGAATGG	Di39F	TATCCTCGTTTTCGTTATGA
Di10R	CCATAACAAAGTAACAAACAG	Di40R	CAATCCTCAAAGATATCAAAGG
Di11F	CTTTTCAGATTATTTCTTCG	Di41F	GTTGCTTTTAATCGTTTTGAG
Di12R	TTAAAATCATAATAAACATAC	Di42R	CCATTCCCAAGAAAATATTA
Di13F	GTTGTTGTTGGTTGTTGTTC	Di43F	CATGTGTTAAGATATGTGTTTAA
Di14R	ATCTTAATACAACTAAAACC	Di44R	AACCTACTTACAGACAATAAG
Di15F	GTATATTTTATTATGTTGTCAGG	Di45F	GGTGTTGCTTTGACTTTGCG
Di16R	AAAAAATTAACACATCAACC	Di46R	AAACGCGAAAAAAATATAACTCTG
Di17F	GATATTAGTTTAGTTCGTCG	Di48R	GTAAATAACACACAACCATCCC
Di18R	CAAATAGAAGAATAAGCCGC	Di49F	TTTTGATAGTCTTGTTATGGG

Fig. 1. Schematic representation of the circular mitochondrial genome of *Dirofilaria immitis*. Each transfer RNA gene is identified by the one letter amino acid code on the inner side of the map. All genes are transcribed in the clockwise direction. Black triangles represent the 2 sets of primers employed for long-PCR amplification of 2 portions (shown by large curves) of the mitochondrial genome, whereas the open triangles represent primers used for DNA sequencing. According to the transcriptional direction, F and R in the primer codes indicate 'forward' and 'reverse', respectively. Regions 1 and 2 each containing 1–2 poly-T tracts (indicated by small arrows) were amplified using primer sets Di19F-Di12R and Di49F-Di48R (indicated by small rectangles), respectively. Sequences of all primers used for amplification or sequencing are given in the box.

https://doi.org/10.1017/S0031182003003275 Published online by Cambridge University Press

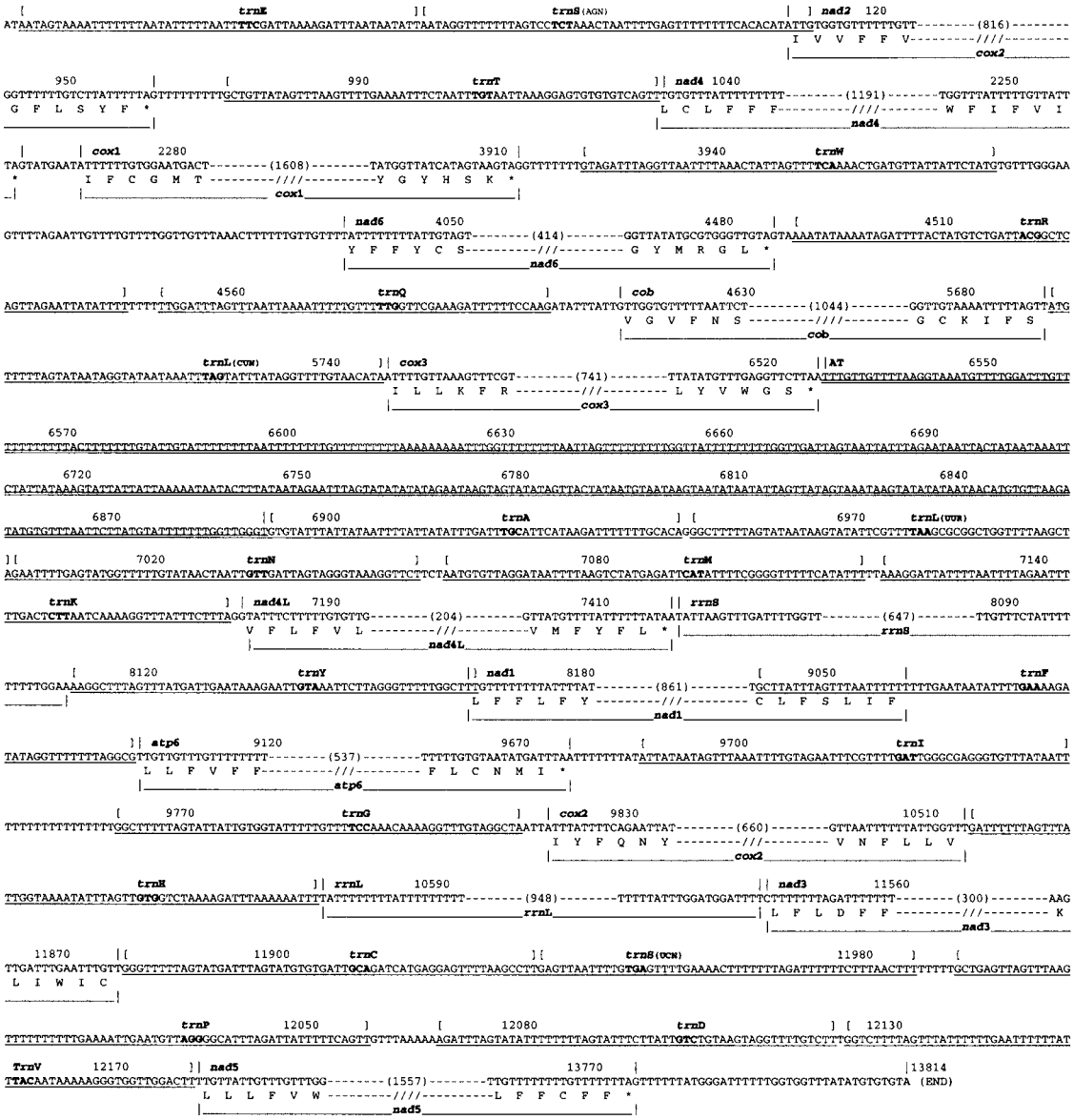


Fig. 2. Linear representation of the mitochondrial genome of *Dirofilaria immitis*, indicating relevant sequence information. The numbers of nucleotides omitted from within the genome sequence are indicated in parentheses. The sequences of the AT-rich region (double-underlined) and the transfer RNA genes (*trn*) are shown in full. For each *trn* gene (underlined), the first and last nucleotides are denoted by '[' and ']', respectively. Trinucleotides representing anticodons (bold-type) are shown directly under each *trn* gene designation. The beginning and end of each ribosomal and protein-coding gene are marked by '|'. Except for the overlap between genes *nad*1 and *trn*F, the first and last 6 codons of each protein-coding gene are shown, along with the translated amino acids. The stop codon of each protein-coding gene is denoted by '*'. All genes are transcribed from left to right.

genes (Dowton *et al*. 2002), and may be associated with the translocation of genes. Regarding the 2 gene block translocations, a likely explanation appears to relate to intra-mitochondrial recombination rather than duplication/random loss (reviewed by Dowton *et al*. 2002).

Protein genes

Twelve protein-coding genes have been identified in the mt genome of *D. immitis*. Four of them use ATT as a translation initiation codon (Table 1). Four others employ TTG, which is commonly used as an initiation codon in nematodes (see Table 3 in Hu *et al*. 2002). The codons GTT, GTA, CTT and TAT are inferred to be used for the initiation of the other 4 protein genes, but are rarely used in the nematodes studied to date, including *O. volvulus* (see Okimoto *et al*. 1992; Keddie *et al*. 1998; Lavrov & Brown, 2001; Hu *et al*. 2002).

Eight of the 12 protein genes (*nad*2, *nad*4, *cox*1, *nad*6, *cox*3, *nad*4L, *atp*6 and *nad*5) use a complete translation termination codon, TAG or TAA. The other 4 (*cob*, *nad*1, *cox*2 and *nad*3) use truncated codons, such as T alone (Table 1). Each of the latter genes is followed by a *trn*, but in no case do they

https://doi.org/10.1017/S0031182003003275 Published online by Cambridge University Press

Table 1. Positions and nucleotide (nt) sequence lengths of individual mitochondrial genes of *Dirofilaria immitis*, and start and stop codons for protein-coding genes as well as the lengths of their predicted amino acid (aa) sequences

		Sequence lengths		Codons	
Gene	nt positions	No. of nt	No. of aa	Start	Stop
*trn*E	3–58	56			
*trn*S (AGN)	58–111	53			
*nad*2	108–962	855	285	ATT	TAG
*trn*T	972–1031	60			
*nad*4	1031–2260	1230	410	TTG	TAG
*cox*1	2268–3914	1647	549	ATT	TAG
*trn*W	3923–3979	57			
*nad*6	4036–4488	453	151	TAT	TAG
*trn*R	4491–4545	55			
*trn*Q	4550–4603	54			
cob	4613–5693	1081	360	GTT	T
*trn*L (CUN)	5694–5748	55			
*cox*3	5749–6528	780	260	ATT	TAA
A-T	6529–6890	362			
*trn*A	6891–6947	57			
*trn*L (UUR)	6949–7001	53			
*trn*N	7002–7057	56			
*trn*M	7061–7118	58			
*trn*K	7121–7177	57			
*nad*4L	7179–7421	243	81	GTA	TAA
*rrn*S	7422–8108	687			
*trn*Y	8111–8166	56			
*nad*1	8165–9062	898	299	TTG	T
*trn*F	9042–9102	61			
*atp*6	9103–9678	576	192	TTG	TAA
*trn*I	9688–9746	59			
*trn*G	9762–9817	56			
*cox*2	9821–10517	697	232	ATT	T
*trn*H	10518–10574	57			
*rrn*L	10575–11542	968			
*nad*3	11543–11879	337	112	CTT	T
*trn*C	11880–11936	57			
*trn*S (UCN)	11937–11988	52			
*trn*P	11994–12059	66			
*trn*D	12069–12123	55			
*trn*V	12125–12181	57			
*nad*5	12182–13777	1596	532	TTG	TAG

overlap by 1 or 2 nucleotides with their downstream gene to complete the termination codon. Therefore, in these cases, after transcription and processing, the incomplete stop codon T could be converted to TAA by polyadenylation (Ojala, Montaya & Attardi, 1981).

After the recognition of translation initiation and termination codons, the lengths of all 12 mt protein genes of *D. immitis* were determined. Table 2 shows the lengths of inferred amino acid sequences and their identities with homologues in other secernentean nematodes. The amino acid sequences inferred for *D. immitis* are very similar in length ($\leqslant$2 amino acid differences) to those of *O. volvulus* and are most divergent from those of *As. suum* ($\leqslant$24 amino acid differences). Proteins COX1, COX3, NAD1, NAD2, NAD4L, NAD5 and NAD6 are 3-24 amino acids longer, and ATP6 and COB are 5-7 amino acids shorter compared with other secernentean nematodes, except *O. volvulus*. Comparisons also revealed that COX1 is amongst the most conserved proteins, whereas NAD6 and ATP6 are the least conserved (see Table 2).

The nucleotide compositions of the entire mt genome sequence of *D. immitis*, and its coding and non-coding regions are compared in Table 3. The whole genome has a T content of 54·9%, the A or G content is 19·3%, and the C content is 6·6%. These percentages are similar to those for *O. volvulus*, but distinct from those of other secernentean nematodes in that the percentage of A (22–28%) is higher and the percentages for T (44–49%) and G (15–20%) lower (see Table 1 in Hu *et al.* 2002). In protein-coding genes, the third codon position has a higher AT content (78%) than the first (70·1%) and second (69·9%) codon positions (Table 3). The increased

https://doi.org/10.1017/S0031182003003275 Published online by Cambridge University Press

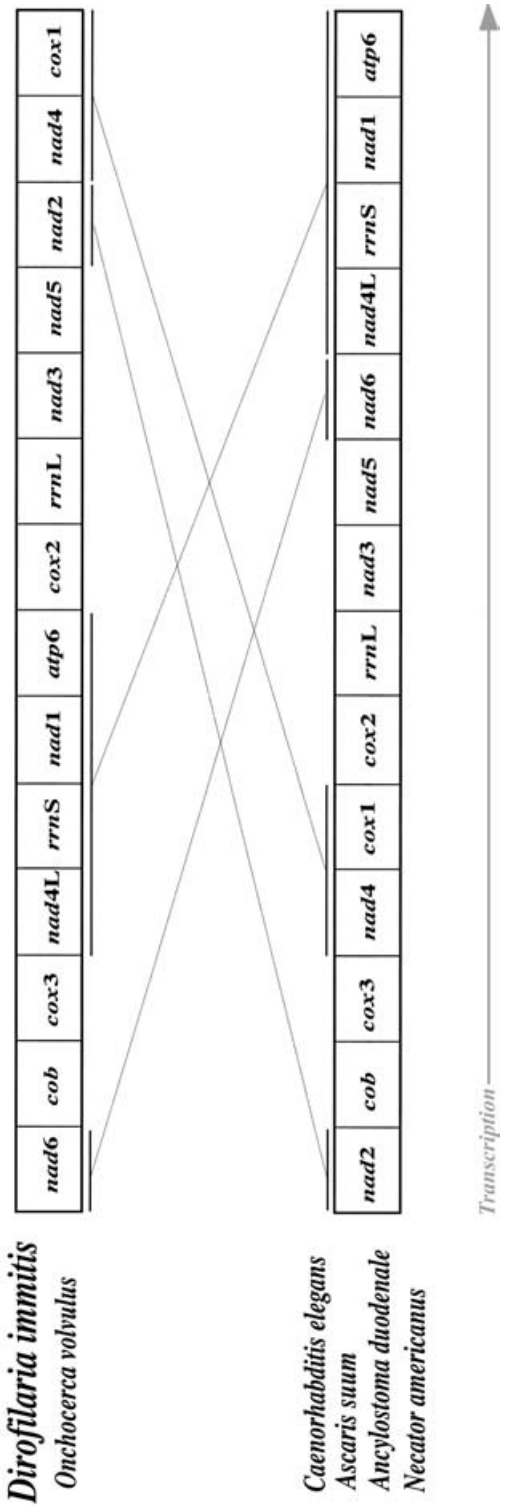


Fig. 3. Schematic, linear representation of the mitochondrial genome arrangement of *Dirofilaria immitis* (showing only protein and ribosomal RNA genes) compared with those published for other secernentean nematodes (see Okimoto *et al.* 1992; Keddie *et al.* 1998; Hu *et al.* 2002).

AT content at the third position relates mainly to the high percentage (68·9%) of T, whereas the percentages of A and C at this position are low (9·1% and 1·1%, respectively). The greatly reduced frequencies of A and C at the third codon position appear to reflect the mutational pattern in the mt genome, as the nucleotides at this position are under the least 'selective pressure' (Sharp & Matassi, 1994). Interestingly, the frequency of C (7·7–12%) at the first and second codon positions is higher than at the third (1·1%) (Table 3). These findings suggest that, although the mt genome of *D. immitis* favours T against C, there is some selection for a higher C content at the former 2 positions which is critical for the maintenance of certain, key amino acids. These amino acids (relating to codons containing a C at the second position) are alanine (2·1%), proline (2·1%), serine (5·2%) and threonine (2·7%) (Table 4). For instance, alanine and proline are non-polar, hydrophobic amino acids which are membrane-bound and, thus, are considered integral in the maintenance of mitochondrial structure and function (Asakawa *et al.* 1991). Of all 20 amino acids, leucine, phenylalanine and serine are the most commonly used by *D. immitis*, whereas arginine, glutamine and histidine are least employed (Tables 4 and 5).

The bias in the nucleotide composition of the mt genome of *D. immitis* is reflected in its codon usage (Table 5). Of 64 possible codons, 62 are used. Codons ACC and CGC are not employed, whereas T-rich codons (with T at 2 of the 3 codon positions), such as TTT (18·9%), TTA (3·4%), TTG (8·6%), ATT (5·7%), GTT (7·2%), TAT (6·6%) and TGT (3·0%), are used more frequently than other codons, except CTT and TTC (<0·1%). These findings reveal that this genome is strongly biased against C. When each codon family is examined, the usage of synonymous codons in the proteins representing the *D. immitis* mt genome follows the same pattern as the nucleotide frequency (i.e. T > G > A > C). This bias is evident for both the 4-fold and 2-fold degenerate codon families, suggesting that the third codon position usually reflects the mutational bias.

Transfer RNA genes

The 22 *trn*s encoded in the mt genome of *D. immitis* vary in size from 52 to 66 nucleotides. While their locations in this genome are the same as for *O. volvulus*, they are distinctly different from those in other nematodes. Secondary structures predicted for the 22 *trn*s of *D. immitis* (Fig. 4) are similar to those of all other secernentean nematodes examined to date (Okimoto *et al.* 1992; Keddie *et al.* 1998; Hu *et al.* 2002). Twenty of them possess a TV-replacement loop, typical for nematodes (Wolstenholme *et al.* 1987). The DHU arms of the 2 serine *trn*s are replaced by a DHU-replacement loop

https://doi.org/10.1017/S0031182003003275 Published online by Cambridge University Press

Table 2. Predicted amino acid sequence lengths and identity of individual mitochondrial proteins of *Dirofilaria immitis* (*Di*) with those of the other secernentean nematodes, *Onchocerca volvulus* (*Ov*), *Ancylostoma duodenale* (*Ad*), *Necator americanus* (*Na*), *Ascaris suum* (*As*) and the free-living nematode *Caenorhabditis elegans* (*Ce*)

	Predicted amino acid sequence lengths						Amino acid identity (%)				
Protein	*Di*	*Ov*	*Ad*	*Na*	*As*	*Ce*	*Di/Ov*	*Di/Ad*	*Di/Na*	*Di/As*	*Di/Ce*
ATP6	192	194	199	198	199	199	76·3	23·1	23·2	21·1	24·1
COB	360	361	370	371	365	370	84·4	48·4	48·4	40·0	45·7
COX1	549	549	525	525	525	525	94·5	50·5	44·2	50·5	50·5
COX2	232	233	232	232	232	231	89·2	42·1	39·1	45·3	40·1
COX3	260	260	255	255	255	255	79·6	34·9	32·8	34·7	33·6
NAD1	299	299	291	291	290	291	79·6	47·8	47·2	51·3	47·2
NAD2	285	284	282	281	281	282	75·4	26·8	30·5	34·4	41·9
NAD3	112	112	112	112	111	111	73·2	34·8	37·5	39·3	34·8
NAD4	410	411	410	410	409	409	84·6	44·3	42·8	44·3	42·8
NAD4L	81	81	78	78	77	77	75·0	33·8	33·8	37·5	38·9
NAD5	531	532	527	527	528	528	80·1	36·1	38·7	37·9	35·9
NAD6	151	150	145	145	144	145	67·3	26·7	26·7	29·3	26·0

Table 3. Nucleotide compositions (%) for the entire or regions of the mitochondrial genome of *Dirofilaria immitis*

Nucleotide	Length (bp)	A	C	T	G	A+T
Entire sequence	13824	19·3	6·6	54·9	19·3	74·2
Protein genes	10389	16·7	6·9	56·0	20·4	72·7
Codon position:						
1st	3463	21·8	7·7	48·3	22·2	70·1
2nd	3463	19·2	12·0	50·7	18·0	69·9
3rd	3463	9·1	1·1	68·9	20·9	78·0
RNA genes	2901	27·2	6·1	50·3	16·4	77·5
AT-rich region	362	31·8	1·9	54·1	12·2	85·9

which is not necessarily exclusive to nematodes. Whilst all of these *trn*s lack a T or D arm (considered crucial for the delivery of the amino-acyl *trn* to the ribosome; Wolstenholme, Okimoto & Macfarlane, 1994), recent studies of *As. suum* have identified 2 translation elongation factors Tu (EF-Tu1 and EF-Tu2) which recognize and subsequently deliver the amino-acyl *trn*s to the mitochondrial ribosomes (Ohtsuki *et al*. 2001, 2002). Clearly, this finding could have important implications for studying translational processes in the mitochondria of other parasitic nematodes.

Each *trn* in the *D. immitis* mt genome is predicted to have an amino-acyl acceptor stem of 7 nucleotides, an anticodon stem of 5 nucleotides and an anticodon loop of 7 nucleotides (Fig. 4). However, 1–3 mismatches appear in the amino-acyl stem of 16 *trn*s. Only 6 of the 22 *trn*s have perfect base-pairing in this stem, allowing for G-U pairing. Like *O. volvulus*, the most notable mismatches are present in *trn*Y (see Keddie *et al*. 1998). While mismatches in the amino-acyl stem are common for metazoan mt *trn*, it is suggested that they are corrected effectively by RNA-editing (Yokobori & Pääbo, 1995, 1997; Lavrov, Brown & Boore, 2000).

Twenty of the *D. immitis* mt *trn*s have a 4 nucleotide DHU stem, and 5–15 nucleotides in the DHU loop (Fig. 4). Their TV-replacement loop varies in size from 4–8 nucleotides. The first of the 2 nucleotides separating the amino-acyl stem from the DHU arm is a U in 16 of the *trn*s; the second is an A in 19 of them. The nucleotide separating the DHU arm from the anticodon stem is an A for *trn*A, *trn*N, *trn*H, *trn*L (CUN), *trn*L (UUR), *trn*M, *trn*F, *trn*W and *trn*Y, a G for *trn*R and *trn*C, and a U for *trn*D, *trn*Q, *trn*E, *trn*G, *trn*I, *trn*K, *trn*P, *trn*T and *trn*V. The two nucleotides immediately preceding the anticodon are always U, except for *trn*K and *trn*S (AGN). The nucleotide after the anticodon is always a purine, except for *trn*I and *trn*L where it is a U.

Twenty of the 22 anticodons are the same as those of the other secernentean nematodes studied to date. The 2 other anticodons are identical to those of *O. volvulus* (see Keddie *et al*. 1998). These are CUU in *trn*K rather than UUU, and AGG in *trn*P rather

https://doi.org/10.1017/S0031182003003275 Published online by Cambridge University Press

Table 4. Amino acid composition of proteins inferred from the mitochondrial genome sequence of *Dirofilaria immitis*

Amino acid	Codon	Number of codons	Codons usage (%)
Non-polar			
Alanine	(GCN)	72	2·1
Isoleucine	(ATY)	199	5·8
Leucine	(CTN)	33	1·0
Leucine	(TTR)	419	12·1
Methionine	(ATR)	142	4·1
Phenylalanine	(TTY)	658	19·1
Proline	(CCN)	74	2·1
Tryptophan	(TGR)	73	2·1
Valine	(GTN)	309	9·0
Total		1976	57·2
Polar			
Asparagine	(AAY)	93	2·7
Cysteine	(TGY)	107	3·1
Glutamine	(CAR)	51	1·5
Glycine	(GGN)	237	6·9
Serine	(AGN)	152	4·4
Serine	(TCN)	179	5·2
Threonine	(ACN)	92	2·7
Tyrosine	(TAY)	231	6·7
Total		1142	33·1
Acidic			
Aspartate	(GAY)	92	2·7
Glutamate	(GAR)	60	1·7
Total		152	4·4
Basic			
Arginine	(CGN)	55	1·6
Histidine	(CAY)	53	1·5
Lysine	(AAR)	76	2·2
Total		184	5·3

Table 5. Nucleotide codon usage for 12 protein-coding genes of the mitochondrial genome of *Dirofilaria immitis* (nt: nucleotide; aa: amino acid; nc: numbers of codons; *: stop codon; total no. of codons is 3463)

nt	aa	nc	%	nt	aa	nc	%	nt	aa	nc	%	nt	aa	nc	%
TTT	Phe	656	18·9	TCT	Ser	162	4·7	TAT	Tyr	230	6·6	TGT	Cys	103	3·0
TTC	Phe	2	<0·1	TCC	Ser	1	<0·1	TAC	Tyr	1	<0·1	TGC	Cys	4	0·1
TTA	Leu	119	3·4	TCA	Ser	8	0·2	TAA	*	4	0·1	TGA	Trp	32	0·9
TTG	Leu	297	8·6	TCG	Ser	8	0·2	TAG	*	5	0·1	TGG	Trp	41	1·2
CTT	Leu	25	0·7	CCT	Pro	68	2·0	CAT	His	52	1·5	CGT	Arg	43	1·2
CTC	Leu	1	<0·1	CCC	Pro	1	<0·1	CAC	His	1	<0·1	CGC	Arg	0	0·0
CTA	Leu	2	<0·1	CCA	Pro	1	<0·1	CAA	Gln	9	0·3	CGA	Arg	5	0·1
CTG	Leu	5	0·1	CCG	Pro	4	0·1	CAG	Gln	42	1·2	CGG	Arg	7	0·2
ATT	Ile	198	5·7	ACT	Thr	88	2·5	AAT	Asn	90	2·6	AGT	Ser	117	3·4
ATC	Ile	1	<0·1	ACC	Thr	0	0·0	AAC	Asn	3	0·1	AGC	Ser	3	0·1
ATA	Met	39	1·1	ACA	Thr	2	<0·1	AAA	Lyn	19	0·5	AGA	Ser	15	0·4
ATG	Met	103	3·0	ACG	Thr	2	<0·1	AAG	Lyn	57	1·6	AGG	Ser	17	0·5
GTT	Val	248	7·2	GCT	Ala	61	1·8	GAT	Asp	90	2·6	GGT	Gly	155	4·5
GTC	Val	2	<0·1	GCC	Ala	4	0·1	GAC	Asp	2	<0·1	GGC	Gly	12	0·3
GTA	Val	19	0·5	GCA	Ala	1	<0·1	GAA	Glu	14	0·4	GGA	Gly	25	0·7
GTG	Val	40	1·2	GCG	Ala	6	0·2	GAG	Glu	46	1·3	GGG	Gly	45	1·3

than UGG (Fig. 4). Anticodon CUU is also present in the *trn*K of platyhelminths (Le, Blair & McManus, 2001), hemichordates (Castresana, Feldmaier-Fuchs & Pääbo, 1998), echinoderms (Giorgi, 1996) and some arthropods (Black & Roehrdanz, 1998; Campbell & Barker, 1999).

Ribosomal RNA genes

The *rrn*S and *rrn*L genes are separated by *nad*1, *atp*6, *cox*2 and 5 *trn*s (Fig. 1). The lengths of these *rrn* genes are 687 bp and 968 bp, respectively, which is similar to those of other nematodes (Okimoto *et al.*

https://doi.org/10.1017/S0031182003003275 Published online by Cambridge University Press

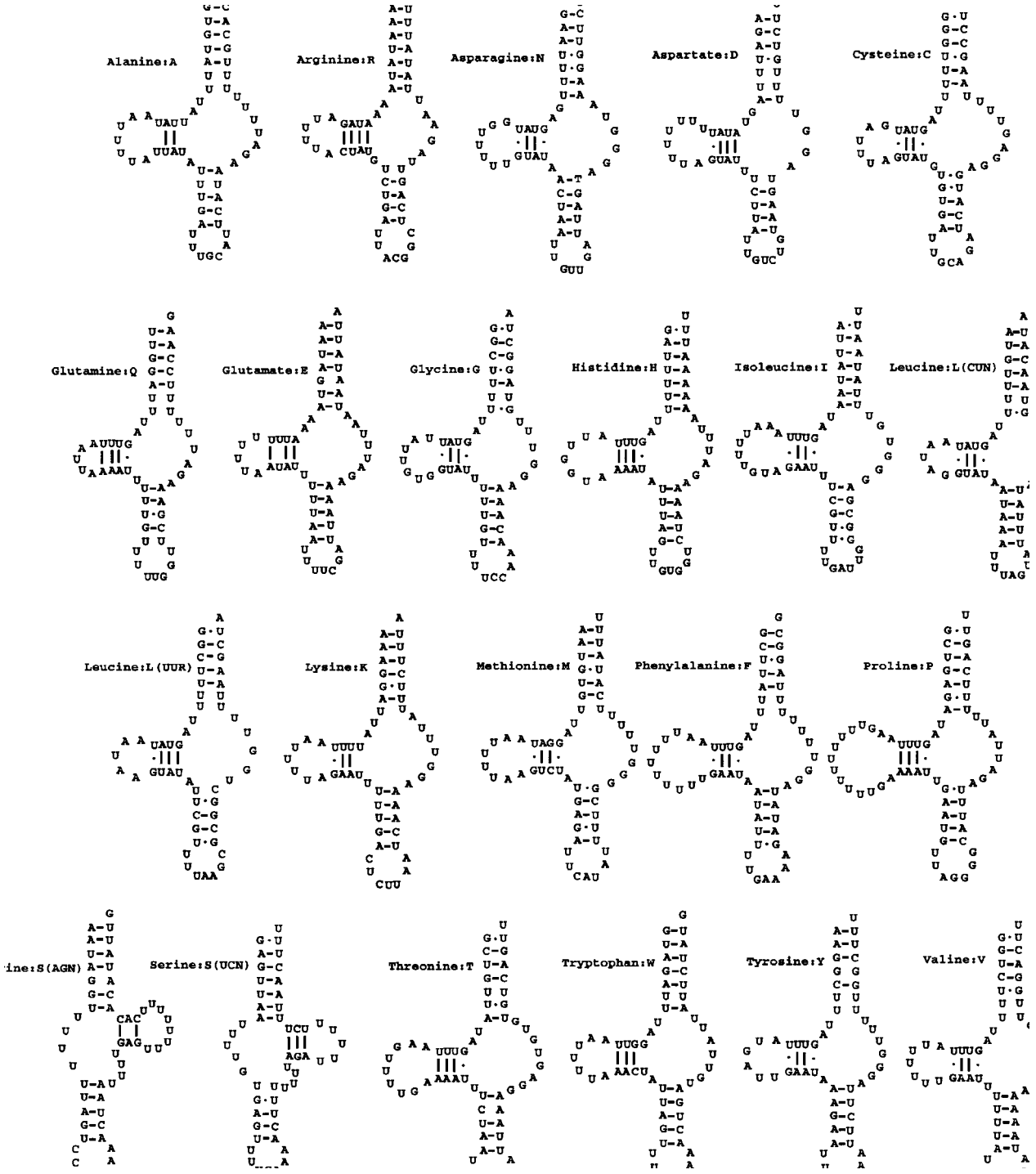


Fig. 4. Predicted secondary structures of the 22 transfer RNA genes in the mitochondrial genome of *Dirofilaria immitis*.

1992; Keddie *et al*. 1998; Lavrov & Brown, 2001; Hu *et al*. 2002) but shorter than for most other Metazoa (*cf* Wolstenholme, 1992, Table VI). The sequence identities in the *rrn*S and *rrn*L genes between *D. immitis* and *O. volvulus* are 89·4% and 83·1%, respectively. Interestingly, the secondary structure predicted for the *rrn*S gene is similar to that of the hookworms *An. duodenale* and *N. americanus*, in that most stems are conserved, except for stems 22 and 37 (Fig. 5) which are absent. Also, the *rrn*L gene structure is similar to that of both hookworms, and the nucleotides predicted to be associated with amino-acyl binding and peptidyl-transfer appear to be conserved (Fig. 6).

Non-coding regions

The longest non-coding region (362 bp) in the mt genome of *D. immitis*, located between *cox*3 and *trn*A (Fig. 1), is 85·9% AT-rich. It is proposed to represent the control (AT-rich) region, since there is no other region of a similar length with such a high AT content. The presence of pairs of inverted repeat sequences in the structure of the AT-rich region

https://doi.org/10.1017/S0031182003003275 Published online by Cambridge University Press

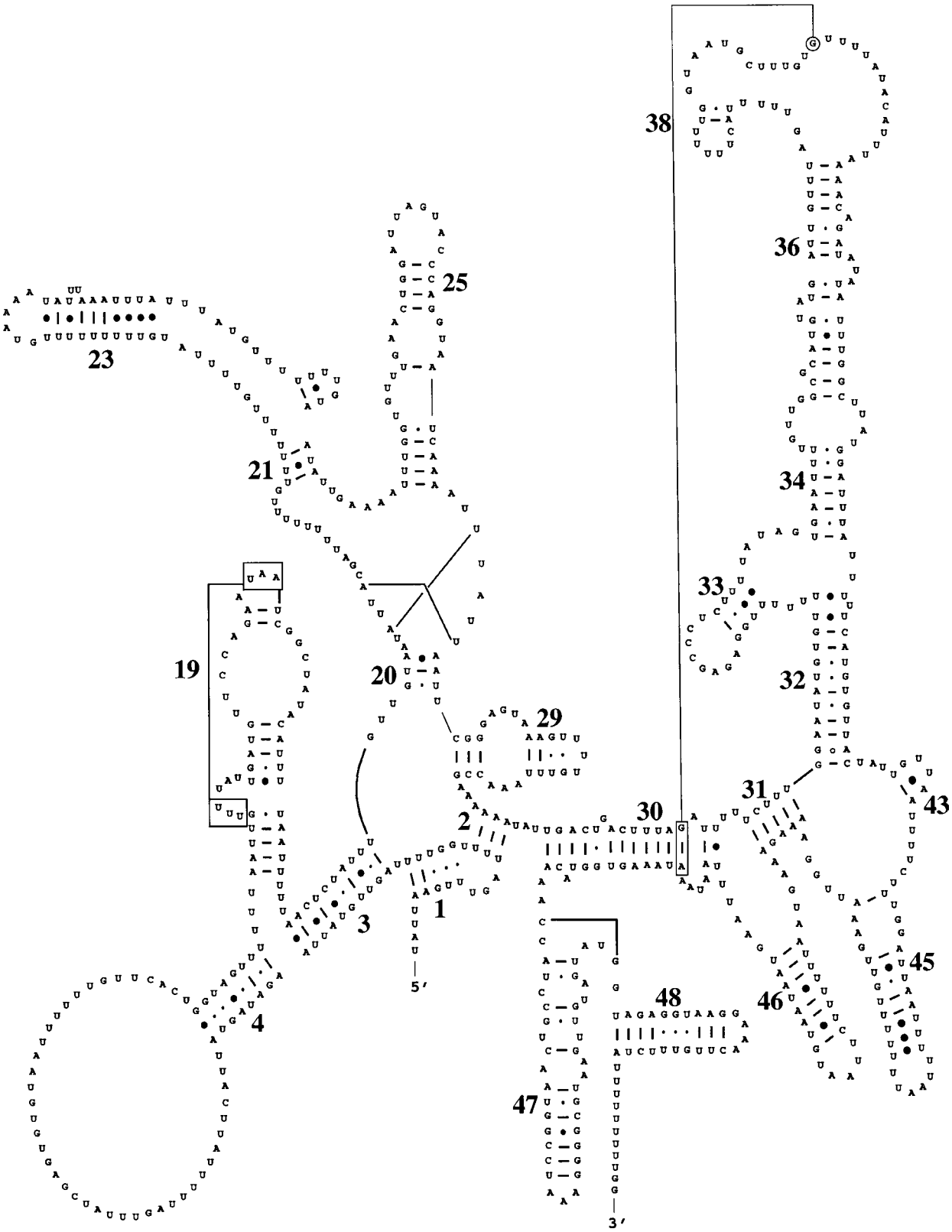


Fig. 5. Predicted secondary structure for the small ribosomal RNA gene subunit (*rrn*S) in the *Dirofilaria immitis* mitochondrial genome. Canonical base pairs C : G and U : A are indicated by lines. G : U pairs are denoted by small dots, and other non-canonical pairings by large dots. Proposed tertiary interactions are represented by long, straight lines. Bold numbers (1–48) identify secondary structure elements considered to be conserved.

https://doi.org/10.1017/S0031182003003275 Published online by Cambridge University Press

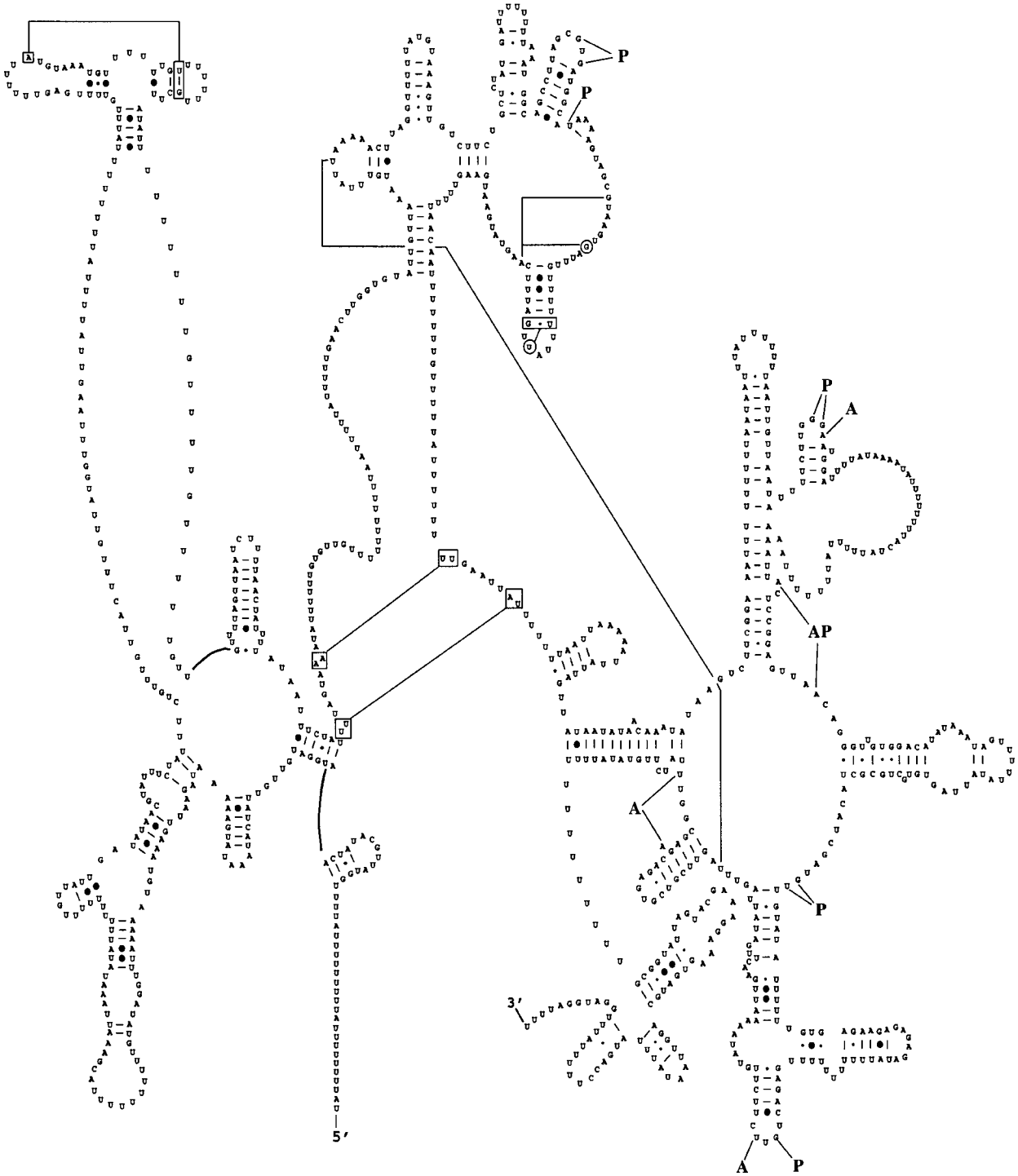


Fig. 6. Predicted secondary structure of the large ribosomal RNA gene subunit (*rrn*L) in the *Dirofilaria immitis* mitochondrial genome. Canonical base pairs C : G and U : A are indicated by lines, G : U pairs by dots, and other non-canonical pairings by large dots. Proposed tertiary interactions are represented by long, straight lines. Sites of amino-acyl binding (A), peptidyl-transfer (P) or both (AP) are indicated in bold-type.

(Fig. 7), common also to other nematodes (Hu *et al*. 2002), suggests functional significance. While the AT-rich region in *D. immitis* does not contain any tandem-repeat motifs (such as the CR1-CR6 in *C. elegans* or the AT dinucleotide repeats in *As. suum*; see Okimoto *et al*. 1992), it does contain a poly-A region typical of other secernentean nematodes (Hu *et al*. 2002).

Overall, the mt genome of *D. immitis* is compact, with only a small number of short non-coding regions between coding genes (Fig. 2). The longest such region (between the genes *trn*W and *nad*6) is 56 bp, being similar in size to a *trn* gene. However, a secondary structure could not be inferred for this region, which indicates that it is neither a *trn* gene nor a duplicated *trn* pseudogene (Beagley,

https://doi.org/10.1017/S0031182003003275 Published online by Cambridge University Press

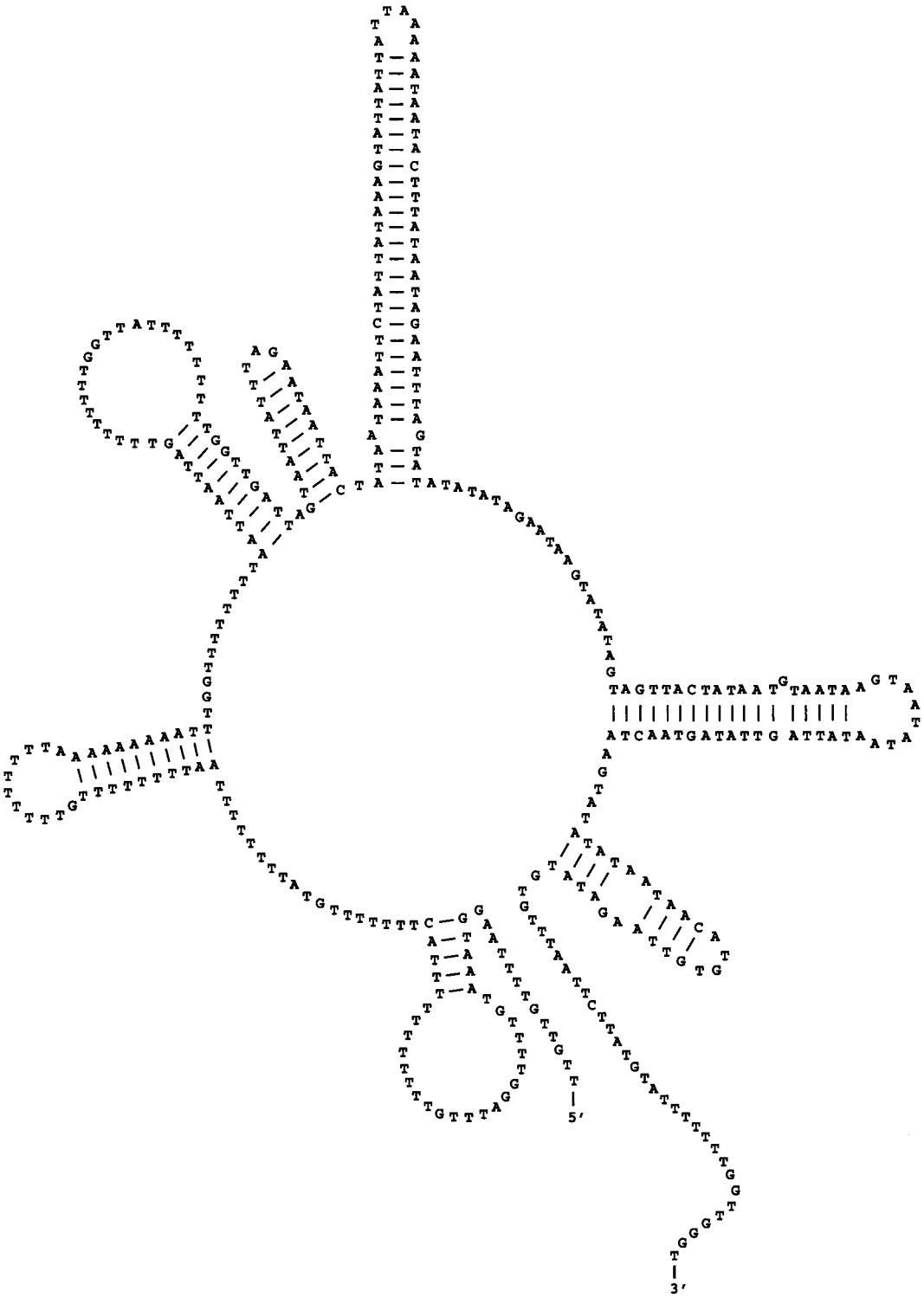


Fig. 7. Predicted secondary structure for the non-coding, AT-rich region in the mitochondrial genome of *Dirofilaria immitis*.

Okimoto & Wolstenholme, 1999). Also, it does not have a stem-and-loop structure, which suggests that it is not involved in the synthesis of the second strand of the mt genome (Boore & Brown, 1994).

Conclusion

The mt genome of the canine heartworm, *D. immitis*, exhibits several features, such as size, gene content and secondary structures of the *trn* and *rrn* genes, which are typical of other secernentean nematodes. However, this genome possesses specific characteristics shared solely by that of *O. volvulus*, the only other filarioid nematode whose complete mt genome sequence has been determined. While the organization of the *D. immitis* mt genome is very similar to that of *O. volvulus*, it is distinctly different from that of all other secernentean nematodes studied to date.

https://doi.org/10.1017/S0031182003003275 Published online by Cambridge University Press

Also, this genome is particularly T-rich, and some regions contain poly-T sequence tracts, and 2 *trn*s use distinct anticodons compared with non-filarioid nematodes. In conclusion, the present study describes the first mt genome sequence and structure for any filarioid nematode of veterinary importance. This information adds significantly to the knowledge of mt genomics of parasitic nematodes, provides a resource for the design of primers for mt genome sequencing projects (via comparison among all mt genome sequences for nematodes) and for future comparative mt genome analyses and phylogenetic studies of nematodes, and may yield genetic markers for molecular epidemiological and population genetics investigations into filarioid parasites.

The authors would like to thank Ian Beveridge for collecting the *D. immitis* used in this study. Min Hu is the recipient of a postgraduate scholarship from The University of Melbourne.

REFERENCES

ASAKAWA, S., KUMAZAWA, Y., ARAKI, T., HIMERO, H., MIURA, K. & WATANABE, K. (1991). Strand specific nucleotide composition bias in echinoderm and vertebrate mitochondrial genomes. *Journal of Molecular Evolution* **32**, 511–520.

BEAGLEY, C. T., OKIMOTO, R. & WOLSTENHOLME, D. R. (1999). *Mytilus* mitochondrial DNA contains a functional gene for a $tRNA^{ser(UCN)}$ with a dihydrouridine arm-replacement loop and a pseudo- $tRNA^{ser(UCN)}$ gene. *Genetics* **152**, 641–652.

BLACK, W. C. 4TH & ROEHRDANZ, R. L. (1998). Mitochondrial gene order is not conserved in arthropods: prostriate and metastriate tick mitochondrial genomes. *Molecular Biology and Evolution* **15**, 1772–1785.

BOORE, J. L. (1999). Animal mitochondrial genomes. *Nucleic Acids Research* **27**, 1767–1780.

BOORE, J. L. & BROWN, W. M. (1994). Complete DNA sequence of the mitochondrial genome of the black chiton, *Katharina tunicata*. *Genetics* **138**, 423–443.

BOORE, J. L. & BROWN, W. M. (1995). Complete sequence of the mitochondrial DNA of the annelid worm *Lumbricus terrestris*. *Genetics* **141**, 305–319.

BOWLES, J., BLAIR, D. & MCMANUS, D. P. (1992). Genetic variants within the genus *Echinococcus* identified by mitochondrial DNA sequencing. *Molecular and Biochemical Parasitology* **54**, 165–173.

CAMPBELL, N. J. H. & BARKER, S. C. (1999). The novel mitochondrial gene arrangement of the cattle tick, *Boophilus microplus*: fivefold tandem repetition of a coding region. *Molecular Biology and Evolution* **16**, 732–740.

CASTRESANA, J., FELDMAIER-FUCHS, G. & PÄÄBO, S. (1998). Codon reassignment and amino acid composition in hemichordate mitochondria. *Proceedings of the National Academy of Sciences, USA* **95**, 3703–3707.

CRINGOLI, G., RINALDI, L., VENEZIANO, V. & CAPELLI, G. (2001). A prevalence survey and risk analysis of filariosis in dogs from the Mt. Vesuvius of southern Italy. *Veterinary Parasitology* **102**, 243–252.

DOWTON, M. & CAMPBELL, N. J. H. (2001). Intramitochondrial recombination – is it why some mitochondrial genes sleep around? *Trends in Ecology and Evolution* **16**, 269–271.

DOWTON, M., CASTRO, L. R. & AUSTIN, A. D. (2002). Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: the examination of genome 'morphology'. *Invertebrate Systematics* **16**, 345–356.

FAN, C.-K., SU, K.-E., LIN, Y.-H., LIAO, C.-W., DU, W.-Y. & CHIOU, H.-Y. (2001). Seroepidemiologic survey of *Dirofilaria immitis* infection among domestic dogs in Taipei city and mountain aboriginal districts in Taiwan (1998–1999). *Veterinary Parasitology* **102**, 113–120.

GASSER, R. B., CHILTON, N. B., HOSTE, H. & BEVERIDGE, I. (1993). Rapid sequencing of rDNA from single worms and eggs of parasitic helminths. *Nucleic Acids Research* **21**, 2525–2526.

GIORGI, C. E., MARTIRADONNA, A., LANAVE, C. & SACCONE, C. (1996). Complete sequence of the mitochondrial DNA in the sea urchin *Arbacia lixula*: conserved features of the echinoid mitochondrial genome. *Molecular Phylogenetics and Evolution* **5**, 323–332.

HOFFMANN, R. J., BOORE, J. L. & BROWN, W. M. (1992). A novel mitochondrial genome organization for the blue mussel, *Mytilus edulis*. *Genetics* **131**, 397–412.

HU, M., CHILTON, N. B. & GASSER, R. B. (2002). The mitochondrial genomes of the human hookworms, *Ancylostoma duodenale* and *Necator americanus* (Nematoda: Secernentea). *International Journal for Parasitology* **32**, 145–158.

KEDDIE, E. M., HIGAZI, T. & UNNASCH, T. R. (1998). The mitochondrial genome of *Onchocerca volvulus*: sequence, structure and phylogenetic analysis. *Molecular and Biochemical Parasitology* **95**, 111–127.

LAVROV, D. V. & BROWN, W. M. (2001). *Trichinella spiralis* mtDNA: a nematode mitochondrial genome that encodes a putative ATP8 and normally structured tRNAs and has a gene arrangement relatable to those of coelomate metazoans. *Genetics* **157**, 621–637.

LAVROV, D. V., BROWN, W. M. & BOORE, J. L. (2000). A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*. *Proceedings of the National Academy of Sciences, USA* **97**, 13738–13742.

LE, T. H., BLAIR, D., AGATSUMA, T., HUMAIR, P.-F., CAMPBELL, N. J. H., IWAGAMI, M., LITTLEWOOD, D. T. J., PEACOCK, B., JOHNSTON, D. A., BARTLEY, J., ROLLINSON, D., HERNIOU, E. A., ZARLENGA, D. S. & MCMANUS, D. P. (2000). Phylogenies inferred from mitochondrial gene orders – a cautionary tale from the parasitic flatworms. *Molecular Biology and Evolution* **17**, 1123–1125.

LE, T. H., BLAIR, D. & MCMANUS, D. P. (2000). Mitochondrial genomes of human helminths and their use as markers in population genetics and phylogeny. *Acta Tropica* **77**, 243–256.

LE, T. H., BLAIR, D. & MCMANUS, D. P. (2001). Complete DNA sequence and gene organization of the mitochondrial genome of the liverfluke, *Fasciola hepatica* L. (Platyhelminthes, Trematoda). *Parasitology* **123**, 609–621.

LE, T. H., BLAIR, D. & MCMANUS, D. P. (2002). Mitochondrial genomes of parasitic flatworms. *Trends in Parasitology* **18**, 206–213.

LOWE, T. M. & EDDY, S. R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* **25**, 955–964.

LUNT, D. H. & HYMAN, B. C. (1997). Animal mitochondrial DNA recombination. *Nature, London* **387**, 247.

MACEY, J. R., LARSON, A., ANANJEVA, N. B., FANG, Z. & PAPENFUSS, T. J. (1997). Two novel gene orders and the role of light-strand replication in arrangement of the vertebrate mitochondrial genome. *Molecular Biology and Evolution* **14**, 91–104.

MAR, P. H., YANG, I.-C., CHANG, G.-N. & FEI, A. C.-Y. (2002). Specific polymerase chain reaction for differential diagnosis of *Dirofilaria immitis* and *Dipetalonema reconditum* using primers derived from internal transcribed spacer region 2 (ITS2). *Veterinary Parasitology* **106**, 243–252.

MORITZ, C., DOWLING, T. E. & BROWN, W. M. (1987). Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annual Review of Ecology and Systematics* **18**, 269–292.

VON NICKISCH-ROSENEGK, M., BROWN, W. N. & BOORE, J. L. (2001). Complete sequence of the mitochondrial genome of the tapeworm *Hymenolepis diminuta*: gene arrangements indicate the platyhelminths are eutrochozoans. *Molecular Biology and Evolution* **18**, 721–730.

OHTSUKI, T., SATO, A., WATANABE, Y.-I. & WATANABE, K. (2002). A unique serine-specific elongation factor Tu found in nematode mitochondria. *Nature, Structural Biology* **9**, 669–673.

OHTSUKI, T., WATANABE, Y.-I., TAKEMOTO, C., KAWAI, G., UEDA, T., KITA, K., KOJIMA, S., KAZIRO, Y., NYBORG, J. & WATANABE, K. (2001). An 'elongated' translation factor Tu for truncated tRNAs in nematode mitochondria. *Journal of Biological Biochemistry* **276**, 21571–21577.

OJALA, D., MONTAYA, J. & ATTARDI, G. (1981). tRNA punctuation model of RNA processing in human mitochondria. *Nature, London* **290**, 470–474.

OKIMOTO, R., MACFARLANE, J. L., CLARY, D. O. & WOLSTENHOLME, D. R. (1992). The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*. *Genetics* **130**, 471–498.

RAWLINGS, T. A., COLLINE, T. M. & BIELER, R. (2001). A major mitochondrial gene rearrangement among closely related species. *Molecular Biology and Evolution* **18**, 1604–1609.

SACCONE, C., DE GIORGI, C., GISSI, C., PESOLE, G. & REYES, A. (1999). Evolutionary genomics in metazoa: the mitochondrial DNA as a model system. *Gene* **238**, 195–209.

SANTALUCIA, J. (1998). A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences, USA* **95**, 1460–1465.

SHADEL, G. S. & CLAYTON, D. A. (1997). Mitochondrial DNA maintenance in vertebrates. *Annual Review of Biochemistry* **66**, 409–435.

SHARP, P. M. & MATASSI, G. (1994). Codon usage and genome evolution. *Current Opinions of Genetic Development* **4**, 851–860.

THOMPSON, J. D., GIBSON, T. J., PLEWNIAK, F., JEANMOUGIN, F. & HIGGINA, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25**, 4876–4882.

UNNASCH, T. R. (2002). River blindness. *Lancet* **360**, 182–183.

WOLSTENHOLME, D. R., MACFARLANE, J. L., OKIMOTO, R., CLARY, D. O. & WAHLEITHNER, J. A. (1987). Bizarre tRNAs inferred from DNA sequences of mitochondrial genomes of nematode worms. *Proceedings of the National Academy of Sciences, USA* **84**, 1324–1328.

WOLSTENHOLME, D. R. (1992). Animal mitochondrial DNA: structure and evolution. *International Review of Cytology* **141**, 173–216.

WOLSTENHOLME, D. R., OKIMOTO, R. & MACFARLANE, J. L. (1994). Nucleotide correlations that suggest tertiary interactions in the TV-replacement loop-containing mitochondrial tRNAs of the nematodes, *Caenorhabditis elegans* and *Ascaris suum*. *Nucleic Acids Research* **22**, 4300–4306.

YOKOBORI, S.-I. & PÄÄBO, S. (1995). tRNA editing in metazoans. *Nature, London* **377**, 490.

YOKOBORI, S.-I. & PÄÄBO, S. (1997). Polyadenylation creates the discriminator nucleotide of chicken mitochondrial tRNA (Tyr). *Journal of Molecular Biology* **265**, 95–99.