The complete mitochondrial genomes for three Toxocara species of human and animal health significance

Ming-Wei Li1,2, Rui-Qing Lin1, Hui-Qun Song1, Xiang-Yun Wu3 and Xing-Quan Zhu*1

Address: 1Laboratory of Parasitology, College of Veterinary Medicine, South China Agricultural University, Guangzhou, Guangdong Province 510642, ProC, 2Department of Veterinary Medicine, Agricultural College, Guangdong Ocean University, Zhanjiang, Guangdong Province 524088, ProC and *Key Laboratory of Marine Bio-resources Sustainable Utilization, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, Guangdong Province 510301, ProC

Email: Ming-Wei Li - lmw1960@126.com; Rui-Qing Lin - rqlin@scau.edu.cn; Hui-Qun Song - hqsong28@scau.edu.cn; Xiang-Yun Wu - wxy816@163.com; Xing-Quan Zhu* - xingquanzh@scau.edu.cn

* Corresponding author

Abstract

Background: Studying mitochondrial (mt) genomics has important implications for various fundamental areas, including mt biochemistry, physiology and molecular biology. In addition, mt genome sequences have provided useful markers for investigating population genetic structures, systematics and phylogenetics of organisms. Toxocara canis, Toxocara cati and Toxocara malaysiensis cause significant health problems in animals and humans. Although they are of importance in human and animal health, no information on the mt genomes for any of Toxocara species is available.

Results: The sizes of the entire mt genome are 14,322 bp for T. canis, 14029 bp for T. cati and 14266 bp for T. malaysiensis, respectively. These circular genomes are amongst the largest reported to date for all secerentean nematodes. Their relatively large sizes relate mainly to an increased length in the AT-rich region. The mt genomes of the three Toxocara species all encode 12 proteins, two ribosomal RNAs and 22 transfer RNA genes, but lack the ATP synthetase subunit 8 gene, which is consistent with all other species of Nematode studied to date, with the exception of Trichinella spiralis. All genes are transcribed in the same direction and have a nucleotide composition high in A and T, but low in G and C. The contents of A+T of the complete genomes are 68.57% for T. canis, 69.95% for T. cati and 68.86% for T. malaysiensis, among which the A+T for T. canis is the lowest among all nematodes studied to date. The AT bias had a significant effect on both the codon usage pattern and amino acid composition of proteins. The mt genome structures for three Toxocara species, including genes and non-coding regions, are in the same order as for Ascaris suum and Anisakis simplex, but differ from Ancylostoma duodenale, Necator americanus and Caenorhabditis elegans only in the location of the AT-rich region, whereas there are substantial differences when compared with Onchocerca volvulus, Dirofilaria immitis and Strongyloides stercoralis. Phylogenetic analyses based on concatenated amino acid sequences of 12 protein-coding genes revealed that the newly described species T. malaysiensis was more closely related to T. cati than to T. canis, consistent with results of a previous study using sequences of nuclear internal transcribed spacers as genetic markers.

Conclusion: The present study determined the complete mt genome sequences for three roundworms of human and animal health significance, which provides mtDNA evidence for the validity of T. malaysiensis and also provides a foundation for studying the systematics, population genetics and ecology of these and other nematodes of socio-economic importance.
Background

Mitochondria are sub-cellular organelles involved in oxidative phosphorylation, offering energy to organisms. They play important roles in cellular metabolism, living and apoptosis. Within these organelles, most metazoan species possess a compact, circular mitochondrial (mt) genome, which varies in size from 14 to 20 kb [1]. The metazoan mt genome usually contains a complement of genes encoding 12–13 protein subunits of the enzymes involved in oxidative phosphorylation, 22 transfer RNAs, and two ribosomal RNAs. There are no introns within genes, and no to limited spacer regions between genes [1,2]. Studying mt genomes has important implications for various fundamental areas, including mt biochemistry, physiology and molecular biology. In addition, mt genome sequences have provided useful markers for investigating population genetic structures, systematics and phylogenetics of organisms due to their maternal inheritance, higher mutation rates than nuclear genes and relatively conserved genome structures [3-5].

Although Nematoda is the second largest animal phylum, to date only 40 complete mitochondrial DNA sequences of nematode species have been deposited in GenBank™ [6-15]. In the order Ascaridada, the mt genomes of only two species have been sequenced [6,15]. The lack of knowledge of mt genomics for parasitic nematodes in this order is a major limitation for population genetic and phylogenetic studies of the pathogens in the order Ascaridada including the species in the genus Toxocara.

Toxocara canis, Toxocara cati and Toxocara malaysiensis are the common ascaridoid nematodes of dogs and cats, causing significant health problems. Infection of dogs with T. canis is quite common, with the prevalence ranging from 5.5% to 64.7% [16-18]. Infection of cats with T. cati is also found worldwide, with infection rates of up to 25.2% to 66.2% [19-21]. More importantly, T. canis and T. cati are of public health significance due to their larvae of being able to invade humans and cause diseases such as ocular larva migrans (OLM), visceral larva migrans (VLM), eosinophilic meningencephalitis (EME) and/or covert toxocariasis (CT) [22-24]. Although they are of importance in human and animal health, there is still no information on the mt genomes available for any of Toxocara species.

The objectives of the present study were to fill some of these knowledge gaps by determining the structure, organization and sequence of the complete mt genomes of T. canis, T. cati and T. malaysiensis of human and animal significance, and to provide mtDNA evidence for the recently described species T. malaysiensis [25,26]. Also, the features of mt genomes of the three ascaridoid nematodes, such as the gene arrangements, structures, compositions, as well as translation and initiation codons and codon usage patterns were compared with those of other nematodes in the same order Ascaridata, namely Ascaris suum and Anisakis simplex. The phylogenetic relationships among these ascaridoid nematodes were also investigated using the protein-coding amino acid sequences.

Results and Discussion

General features of the mt genome of three Toxocara species

The complete mt genomes of T. canis, T. cati and T. malaysiensis are 14,322 bp, 14,029 bp and 14266 bp in length, respectively. These complete mt genome sequences have been deposited in the GenBank™ under the accession numbers AM411108 (for T. canis), AM411622 (for T. cati) and AM412316 (for T. malaysiensis). All three mt genomes contain 12 protein-coding genes (cox1-3, nad1-6, nad4L, atp6 and cytb), 22 transfer RNA genes and two ribosomal RNA genes, but lack an atp8 gene (Table 1). These circular genomes are typical of other nematode mitochondrial genomes except for Trichinella spiralis in which the atp8 gene is encoded. All genes are transcribed in the same direction as found in other members of the secernentean nematodes sequenced to date, but in contrast to T. spiralis and Xiphinema americanus [8,13].

The mt genome arrangement of T. canis, T. cati and T. malaysiensis are the same as those of A. suum and A. simplex and almost identical to the genome structures of strongyloida nematodes Anclyostoma duodenale, Nectator americanus, Cooperia oncophora and rhabditid nematode Caenorhabditis elegans, with the exception of the relative position of an AT-rich region and the number of non-coding regions (NCR). In the mt genomes of the three Toxocara species, the AT-rich region is located between trnS2 and trnN, whereas it is positioned between trnA and trnP in A. duodenale, N. americanus, Co. oncophora and C. elegans. The only NCR found in Toxocara spp. was between nad4 and cox1, as has been identified in the aforementioned species. The second NCR located between nad3 and nad5 in the hookworms A. duodenale and N. americanus was not found in Toxocara spp. [9]. The genome structures of these Toxocara species differ significantly from those of Onchocerca volvulus, Strongyloides stercoralis, Dirofilaria immitis, X. americanus and T. spiralis in the location of the AT-rich region, and some tRNA and protein-coding genes. Comparison of the gene arrangements of these three Toxocara species with those of the other nine representatives of the secernentean nematodes suggests that T. canis, T. cati and T. malaysiensis are more closely related to A. suum and A. simplex than to C. elegans, A. duodenale, N. americanus, Co. oncophora, O. volvulus, D. immitis and S. stercoralis.

The nucleotide compositions of the entire mtDNA sequences for T. canis, T. cati and T. malaysiensis are biased
Table 1: Positions and nucleotide sequence lengths of mitochondrial genomes of *T. canis*, *T. cati* and *T. malaysiensis*, and start and stop codons for protein-coding genes as well as lengths of their predicted amino acid sequences and tRNA gene anticodons (starting from nad1)

| Genes | Positions and nt sequence lengths (bp) | Initiation and termination codons (Ini/Ter) | No. of amino acid anticodons |
|-------|---------------------------------------|---------------------------------------------|------------------------------|
|       | Tcan | Tcat | Tmal | Tcan | Tcat | Tmal | Tcan | Tcat | Tmal | Tcan | Tcat | Tmal | Tcan | Tcat | Tmal |
| nad1  | 1–873 (873) | 1–873 (873) | 1–873 (873) | TTG/TAG | TTG/TAA | TTG/TAG | 291 | 291 | 291 | 291 | 291 | 291 |
| atp6  | 876–1473 (598) | 875–1472 (598) | 875–1472 (598) | ATT/T | ATT/T | ATT/T | 200 | 200 | 200 | 200 | 200 | 200 |
| trn-Lys | 1474–1535 (62) | 1473–1535 (63) | 1473–1535 (61) | TTT | TTT | TTT | 369 | 369 | 369 | 369 | 369 | 369 |
| trn-LeuUUR | 1536–1590 (55) | 1536–1590 (55) | 1534–1589 (56) | ATG/T | ATG/T | ATG/T | 256 | 256 | 256 | 256 | 256 | 256 |
| trn-SerAGN | 1591–1642 (52) | 1591–1641 (51) | 1590–1642 (53) | ATT/T | ATT/T | ATT/T | 410 | 410 | 410 | 410 | 410 | 410 |
| nad2  | 1643–2486 (844) | 1643–2486 (844) | 1643–2486 (844) | ATT/T | GTT/T | ATT/T | 369 | 369 | 369 | 369 | 369 | 369 |
| trn-Ile | 2487–2547 (61) | 2486–2546 (61) | 2487–2549 (63) | TAG | TAG | TAG | 238 | 237 | 237 | 238 | 237 | 237 |
| trn-Arg | 2548–2601 (54) | 2547–2602 (56) | 2550–2604 (55) | ACG | ACG | ACG | 528 | 528 | 528 | 528 | 528 | 528 |
| trn-Glu | 2604–2658 (55) | 2603–2657 (55) | 2605–2659 (55) | TTG | TTG | TTG | 112 | 112 | 112 | 112 | 112 | 112 |
| trn-Phe | 2659–2716 (58) | 2658–2716 (59) | 2662–2718 (57) | TTT | TTT | TTT | 145 | 145 | 145 | 145 | 145 | 145 |
| cyt b  | 2717–3823 (1107) | 2717–3823 (1107) | 2719–3825 (1107) | GTG/TAG | ATG/TAA | GTG/TAG | 526 | 526 | 526 | 526 | 526 | 526 |
| cox 1  | 3824–3879 (56) | 3826–3880 (55) | 3826–3882 (57) | TAG | TAG | TAG | 369 | 369 | 369 | 369 | 369 | 369 |
| tRNA-LeuCUN | 3880–4647 (768) | 3881–4648 (768) | 3883–4650 (768) | ATG/TAG | ATG/TAA | GTG/TAG | 238 | 237 | 237 | 238 | 237 | 237 |
| cox 2  | 4648–4704 (57) | 4652–4708 (57) | 4651–4705 (55) | TAG | TAG | TAG | 526 | 526 | 526 | 526 | 526 | 526 |
| trn-Thr | 5935–6045 (111) | 5939–6054 (116) | 5936–6047 (112) | GTG/TAG | ATG/TAA | GTG/TAG | 369 | 369 | 369 | 369 | 369 | 369 |
| cox 3  | 6046–7623 (1578) | 6055–7632 (1578) | 6048–7628 (1581) | TAG | TAG | TAG | 238 | 237 | 237 | 238 | 237 | 237 |
| trn-Lys | 7926–7680 (55) | 7365–7688 (54) | 7361–7686 (56) | GCA | GCA | GCA | 528 | 528 | 528 | 528 | 528 | 528 |
| tRNA-Met | 7689–7748 (60) | 7694–7753 (60) | 7694–7753 (60) | CAT | CAT | CAT | 528 | 528 | 528 | 528 | 528 | 528 |
| trn-Asp | 7750–7808 (59) | 7754–7809 (56) | 7755–7813 (59) | GTC | GTC | GTC | 528 | 528 | 528 | 528 | 528 | 528 |
| tRNA-Gly | 7808–7863 (56) | 7809–7868 (56) | 7813–7869 (57) | TCC | TCC | TCC | 528 | 528 | 528 | 528 | 528 | 528 |
| cox 4  | 7864–8577 (714) | 7867–8577 (711) | 7860–8580 (711) | GTT/TAG | GTT/TAA | GTA/TAG | 238 | 237 | 237 | 238 | 237 | 237 |
| trn-His | 8576–8631 (56) | 8577–8634 (58) | 8579–8637 (59) | TAG | TAG | TAG | 528 | 528 | 528 | 528 | 528 | 528 |
| trn-Val | 9590–9599 (55) | 9590–9599 (55) | 9590–9599 (55) | TTA | TTA | TTA | 112 | 112 | 112 | 112 | 112 | 112 |
| trn-Asn | 11590–11593 (55) | 11590–11593 (55) | 11590–11593 (55) | TGT | TGT | TGT | 145 | 145 | 145 | 145 | 145 | 145 |
| trn-Glu | 13170–13222 (53) | 13170–13222 (53) | 13170–13222 (53) | TGC | TGC | TGC | 78 | 78 | 78 | 78 | 78 | 78 |
| trn-Val | 14208–14263 (56) | 14208–14263 (56) | 14208–14263 (56) | TTA | TTA | TTA | 78 | 78 | 78 | 78 | 78 | 78 |

Tcan: *Toxocara canis*, Tcat: *Toxocara cati*, Tmal: *Toxocara malaysiensis*
toward A and T, with T being the most favored nucleotide and C the least favored, in accordance with mt genomes of other nematodes. The content of A+T is 68.57% for T. canis (21.9% A, 46.7% T, 22.0% G and 9.4% C), which is the lowest of the nematodes studied to date (Table 2). The content of A+T is 69.95% for T. cati (22.3% A, 47.7% T, 20.9% G and 9.1% C) and 68.86% for T. malaysiensis (21.7% A, 47.2% T, 22.0% G and 9.1% C), respectively (Table 2).

Protein-coding genes and codon usage patterns
The boundaries between protein-coding genes of mt genomes of T. canis, T. cati and T. malaysiensis were determined by aligning their sequences and by identifying translation initiation and termination codons with those of A. suum. For each of the three Toxocara species, the lengths of protein-coding genes cox3, nad1, nad2, nad3 and nad4 are the same as those of A. suum, whereas the lengths of genes atp6, nad4L, nad5 and nad6 are reduced, and the lengths of cox2 and cyt b are increased (Table 1) when compared to those of A. suum. The length of cox1 in T. canis and T. cati is the same as that of A. suum, but the length of cox1 in T. malaysiensis is increased (Table 3).

The inferred nucleotide and amino acid sequences for each of the 12 mt proteins of T. canis, T. cati and T. malaysiensis were compared with those of A. suum and A. simplex. The identity of the nucleotide and amino acid sequences is 71–90% and 72.2–95.8%, respectively (Table 3). Based on identity, cox1 is the most conserved protein-coding gene, while the cyt b is the least conserved. For all 12 proteins, the amino acid sequence identities are higher when compared between each of the three Toxocara species than between each of Toxocara specis and A. suum or A. simplex. The identities of the nucleotide and amino acid sequences among five ascaridoid species are higher than those among C. elegans, A. duodenale, N. americanus and Co. oncophora (data not shown). These findings reinforce the conclusion that the three Toxocara species are genetically more closely related to A. suum and A. simplex than to C. elegans, A. duodenale, N. americanus and Co. oncophora.

The predicted initiation and termination codons for the protein-coding genes of the three Toxocara species were compared with those of ascaridoid species (A. suum and A. simplex) and with selected species representing different nematode orders including the human hookworm A. duodenale, filarioid worm O. volvulus, and rhabditid species S. stercoralis. The most common start codon for three Toxocara species is TIG (four of 12 protein-coding genes), followed by ATI (three of 12 protein-coding genes for T. canis and T. malaysiensis, two for T. cati), ATA (two of 12 protein-coding genes for T. canis and T. cati, one for T. malaysiensis), and ATG, GTG, GTT and GTA are used as initiation codons. GTG, which is used in the cyt b of T. canis and T. malaysiensis, and cox3 of T. malaysiensis, is not used as a start codon in the other nematodes compared. GTA used in the cox2 of T. malaysiensis is also used as a start codon for nad4L in D. inmitis (data not shown). Seven of the 12 protein-coding genes were predicted to have a TAG or TAA translation termination codon. The remaining protein-coding genes were inferred to end with an abbreviated stop codon, such as TA or T. For the three Toxocara species, the 3’-end of most of these genes is immediately adjacent to a downstream tRNA gene (Table 1), which is

| Genes | Tcan | Tcat | Tmal | Asu | Asi | Ad | Ov | Ss |
|-------|------|------|------|-----|-----|----|----|----|
| atp6  | 68.7 | 64.5 | 61.1 | 68.4 | 66.1 | 74.0 | 72.9 | 72.7 |
| cox1  | 66.5 | 67.5 | 65.1 | 66.7 | 67.0 | 69.3 | 67.0 | 72.6 |
| cox2  | 65.0 | 67.3 | 66.0 | 68.8 | 66.3 | 70.8 | 69.2 | 75.5 |
| cox3  | 64.6 | 67.3 | 66.7 | 69.5 | 68.2 | 74.1 | 70.0 | 73.9 |
| cyt b | 65.4 | 67.0 | 65.3 | 70.4 | 67.9 | 74.3 | 71.8 | 75.9 |
| nad1  | 64.5 | 67.3 | 66.7 | 69.5 | 68.2 | 74.1 | 70.0 | 73.9 |
| nad2  | 70.1 | 71.1 | 71.7 | 73.0 | 73.3 | 81.2 | 74.3 | 81.3 |
| nad3  | 71.7 | 73.5 | 71.7 | 73.8 | 75.3 | 78.3 | 76.4 | 78.3 |
| nad4  | 68.0 | 68.2 | 67.3 | 71.0 | 69.9 | 78.5 | 73.2 | 76.6 |
| nad6  | 67.8 | 73.0 | 73.0 | 76.5 | 76.7 | 80.3 | 78.6 | 82.9 |
| rrL   | 69.4 | 71.2 | 70.1 | 72.9 | 72.1 | 77.1 | 72.9 | 79.6 |
| AT-region | 79.4 | 81.3 | 78.4 | 84.7 | 87.2 | 90.1 | 85.3 | 85.0 |

Tcan: Toxocara canis, Tcat: Toxocara cati, Tmal: Toxocara malaysiensis, Asu: Ascaris suum, Asi: Anisakis simplex, Ad: Ancylostoma duodenale, Ov: Onchocerca volvulus, Ss: Strongyloides stercoralis, EmrG: entire mitochondrial genome.
Table 3: Comparison of protein-coding genes in size, with identity of nucleotides and predicted amino acids for five ascaridoid species

| Gene | Sizes of protein-coding genes | Identity of nucleotides/amino acids |
|------|--------------------------------|------------------------------------|
|      | Tcan | Tcat | Tmal | Asu | Asi | Tcan/Tcat | Tcan/Tmal | Tcan/Asu | Tcan/Asi | Tcat/Tmal | Tcat/Asu | Tcat/Asi | Tmal/Asu | Tmal/Asi | Asu/Asi |
| atp  | 598  | 598  | 598  | 600 | 600 | 88.8/91.0 | 85.5/89.9 | 78.1/79.9 | 79.1/80.9 | 89.1/93.0 | 79.8/79.4 | 79.4/80.9 | 80.8/81.4 | 78.6/82.4 | 77.7/85.4 |
| cytb | 1107 | 1107 | 1107 | 1098| 1099| 82.2/87.2 | 83.2/88.9 | 77.2/76.6 | 71.3/72.0 | 83.0/88.6 | 75.8/75.8 | 72.1/72.0 | 74.8/77.4 | 72.7/73.9 | 73.7/75.8 |
| cox1 | 1578 | 1578 | 1581 | 1578| 1576| 89.7/94.1 | 89.3/93.9 | 86.1/92.2 | 83.3/90.7 | 90.0/94.7 | 86.4/91.4 | 83.7/90.5 | 86.4/92.4 | 83.1/91.6 | 83.3/92.6 |
| cox2 | 714  | 711  | 711  | 699 | 699 | 87.8/94.9 | 89.5/95.8 | 85.1/90.9 | 80.7/88.4 | 87.8/95.8 | 84.7/91.4 | 80.7/88.8 | 85.7/90.9 | 82.1/87.9 | 82.1/90.1 |
| cox3 | 768  | 768  | 768  | 768 | 766 | 86.3/93.7 | 85.8/92.9 | 81.0/88.6 | 79.2/87.5 | 87.0/94.5 | 81.8/85.0 | 80.2/87.1 | 81.9/87.8 | 79.0/87.1 | 79.2/89.0 |
| nad1 | 873  | 873  | 873  | 873 | 873 | 85.3/91.7 | 87.5/93.1 | 80.2/80.7 | 81.3/85.9 | 87.1/94.5 | 80.8/83.1 | 82.1/87.2 | 81.0/81.7 | 82.0/86.2 | 80.6/83.4 |
| nad2 | 844  | 844  | 844  | 844 | 846 | 84.5/82.9 | 84.7/85.4 | 75.9/76.9 | 72.2/75.8 | 86.4/88.6 | 77.6/75.4 | 72.3/76.9 | 76.4/75.8 | 72.4/76.9 | 75.8/80.1 |
| nad3 | 336  | 336  | 336  | 336 | 336 | 87.5/88.3 | 86.6/85.6 | 80.4/81.1 | 82.4/85.6 | 84.5/84.7 | 85.1/84.7 | 81.0/84.7 | 79.5/79.3 | 79.5/79.3 | 78.9/83.8 |
| nad4 | 1230 | 1230 | 1230 | 1230| 1230| 83.1/87.8 | 81.4/86.6 | 75.0/78.2 | 71.8/76.0 | 82.0/86.6 | 74.7/77.8 | 73.8/78.5 | 75.0/78.7 | 72.3/77.5 | 74.8/80.9 |
| nad4L| 232  | 232  | 232  | 234 | 232 | 85.0/90.9 | 85.8/88.7 | 78.1/85.7 | 76.7/80.5 | 85.4/89.6 | 80.7/81.8 | 79.7/81.8 | 79.4/79.2 | 78.0/80.5 | 80.2/87.0 |
| nad5 | 1582 | 1582 | 1582 | 1582| 1582| 83.3/87.3 | 84.3/86.5 | 76.7/77.6 | 75.3/77.0 | 85.8/88.2 | 76.7/79.1 | 75.8/79.3 | 77.7/78.7 | 77.4/78.1 | 76.2/82.0 |
| nad6 | 434  | 434  | 434  | 435 | 435 | 83.7/85.4 | 83.0/80.6 | 72.6/72.2 | 71.0/72.9 | 86.4/88.9 | 76.3/73.6 | 72.6/73.6 | 76.1/73.6 | 72.6/72.2 | 74.0/80.6 |
| rrnS | 697  | 696  | 696  | 701 | 699 | 89.9 | 90.4 | 82.8 | 81.9 | 92.2 | 82.3 | 84.2 | 81.5 | 82.3 | 83.4 |
| rrnL | 958  | 955  | 955  | 960 | 957 | 83.6 | 82.4 | 72.9 | 69.5 | 85.2 | 73.0 | 72.1 | 72.5 | 71.6 | 71.5 |
| EmtG | 14322| 14029| 14226| 14284| 13916| 85.3 | 85.7 | 80.2 | 78.5 | 87.1 | 81.1 | 80.2 | 80.5 | 79.5 | 80.2 |

Tcan: *Toxocara canis*, Tcat: *Toxocara cati*, Tmal: *Toxocara malaysiensis*, Asu: *Ascaris suum*, Asi: *Anisakis simplex*, EmtG: entire mitochondrial genome.
consistent with the arrangement in the mt genomes of A. suum and A. simplex [6,15], but in contrast to that of C. elegans where both the nad1 and nad3 genes terminate in T or TA, and are followed by the putative ATT translation initiation codon of their downstream protein-coding genes. The protein-coding gene nad6 ended with an abbreviated stop codon TA is followed by the putative ATT translation initiation codon of its downstream protein coding gene nad4L, which is similar to C. elegans.

In general, the nucleotides of metazoan mt genomes are not randomly distributed, and such nucleotide bias is often associated with unequal usage of synonymous codons. The mt genome nucleotide composition of nematodes is biased toward A and T. The A+T content of protein-coding genes ranged from 63.2% to 74.9% for all three Toxocara species (Table 2). This bias in nucleotide composition towards AT (Table 2) affects both the codon usage pattern and amino acid composition of proteins. In these three species examined, all 64 possible codons were identified by sequence comparison with those of other nematodes (Table 1). The A+T contents are 79.4% (T. canis), 81.3% (T. cati) and 85.2% between T. malayensis and A. simplex (AGN) in which the DHU-arm is lacking.

Ribosomal RNA genes
The rrnS and rrnL genes of the three roundworm species were identified by sequence comparison with those of A. suum. The rrnS is located between trnE and trnS (UCN), and rrnL is located between trnH and nad3. The two genes are separated from one another by the protein-coding genes nad3, nad5, nad6 and nad4L. The size of the rrnS gene is 697 bp for T. canis, 696 bp for T. cati and 696 bp for T. malayensis. The size of the rrnL gene is 958 bp for T. canis, 955 bp for T. cati and 955 bp for T. malayensis. The sizes of the two rRNA genes for the three Toxocara species are similar to those of other nematodes (Table 1). The A+T contents of the rrnS for T. canis, T. cati and T. malayensis are 66.9%, 68.8% and 66.4%, respectively, whereas those of the rrnL are higher (72.0%, 73.5% and 70.1%, respectively), and A+T contents of the two genes are the lowest among the nematodes studied to date (Table 2). Sequence identity in the rrnS and rrnL genes is 87.5% and 83.6% between T. canis and T. cati, 87.9% and 82.4% between T. canis and T. malayensis, and 91.4% and 85.2% between T. cati and T. malayensis, respectively.

Non-coding regions
Like A. suum and A. simplex, the longest non-coding region (AT-region) in the three Toxocara mt genomes is located between the trnS2 and trnN. Their sizes are 985 bp (T. canis), 711 bp (T. cati) and 936 bp (T. malayensis), and A+T contents are 79.4% (T. canis), 81.3% (T. cati) and 78.4% (T. malayensis), respectively, which are significantly lower than the comparable NCRs of nematodes studied to date (Table 2). Repeated sequence motifs (CR1-CR6) present in the C. elegans AT-rich region [6] are not found in Toxocara spp. However, there are some AT dinucleotide repeats in the AT-region of Toxocara mt genomes.
of which the longest consists of repeat units (34 base pairs). Similar AT dinucleotide repeats have been found in *A. suum* [6]. The function or role of these AT repeats is currently unknown [6,7]. Although nothing is yet known about the replication process(es) in the mtDNA of parasitic nematodes, the high A+T content and the predicted structure of the AT-rich non-coding region suggests an involvement in the initiation of replication [28].

For the three roundworm species, the second longest non-coding region is located between genes *cox1* and *nad4*, as in the mt genomes of *A. suum* [6]. Its length is 111 bp (*T. canis*), 116 bp (*T. cati*) and 112 bp (*T. malaysiensis*), with an A+T content of 86.2%, 74.1% and 74.1%, respectively, and is shorter than that of *A. suum* (117 bp). The non-coding region for the three *Toxocara* species could form a hairpin loop structure (AATTTTTAAAAATT).

**Phylogenetic analyses**

The final alignment of the amino acid sequences of 12 proteins for the six taxa (*T. canis, T. cati, T. malaysiensis, A. suum, A. simplex* and *O. volvulus*) yield 3516 characters (2079 variable, 339 parsimony-informative). In all three phylogenetic analyses, three *Toxocara* species were clustered together (Fig. 1). *T. malaysiensis*, the recently described *Toxocara* species from cat [26], was inferred to be the sister species of *T. cati* with high bootstrap values. This result was consistent with that of a previous study [25] which used sequences of internal transcribed spacers of nuclear ribosomal DNA, thus providing mt DNA evidence for the validity of *T. malaysiensis* as an ascaridoid of cats. *T. malaysiensis* is more closely related to *T. cati*, the common ascaridoid of cats, than to *T. canis*, the common ascaridoid of canids.

*Toxocara* species was resolved being more closely related to *A. suum* than to *A. simplex* with moderate support in the phylogenetic analyses, which was consistent with results of previous morphological and molecular studies [29,30]. But relationship between *A. suum* and *A. simplex* was poorly inferred in the MP and ML analyses (Fig. 1).

**Conclusion**

*Toxocara* species are the important socio-economic parasites because they have significant impact on human health. The determined mt genomes of the three roundworms, *T. canis, T. cati* and *T. malaysiensis*, add the mtDNA data to the order Ascaridida, which includes a broad range of parasites of major socio-economic importance. Determination of the complete mt genome sequences for three *Toxocara* species of human and animal health significance provides a foundation for studying the systematics, population genetics and ecology of these and other nematodes of socio-economic importance.

![Figure 1](http://www.biomedcentral.com/1471-2164/9/224)

**Figure 1**

Inferred phylogenetic relationship among the five ascaridoid species (*Toxocara canis, T. cati, T. malaysiensis, Ascaris suum, Anisakis simplex*) derived from neighbour joining (NJ), maximum parsimony (MP) and maximum likelihood (ML) analyses of amino acid sequences of 12 protein-coding genes from their mitochondrial genomes, using one filarioid species (*Onchocerca volvulus*) as outgroup. The numbers along branches indicate bootstrap values resulting from different analyses in the order: NJ/MP/ML. Values lower than 50 are given as "-".
Methods  
**Parasites and DNA extraction**

The ascaridoid nematodes used in the present study were from Zhanjiang city for *T. canis*, Changsha city for *T. cati* and Guangzhou city for *T. malaysiensis*, China, respectively. Adult nematodes of three *Toxocara* species were obtained from the intestines of dogs and/or cats, washed in physiological saline, identified primarily based on morphological characters to species, fixed in 70% (v/v) ethanol and stored at -20°C until use. Total DNA was isolated from individual nematodes using sodium dodecyl-sulphate/proteinsin K treatment, followed by spin-column purification (Wizard Clean-Up, Promega). The specific identity of each individual nematode was verified by species-specific PCR amplification using the sequences of the first and/or second internal transcribed spacers (ITS-1 and/or ITS-2) of ribosomal DNA (rDNA) as the species-specific genetic markers [31,32].

**Long-PCR amplification and sequencing**

Using primer set 5F/40R (~9 kb region) and 39F/42R (~6 kb region) [33], the entire mt genome of each *Toxocara* species was amplified in two overlapping fragments by long-PCR from approximately 20 ng of total genomic DNA purified from an individual nematode, respectively.

**Sequence analyses**

Sequences were assembled manually and aligned against the complete mt genome sequence of *A. suum* (GenBank™ accession number NC001327) using the program Clustal X to identify gene boundaries. The open-reading frames and codon usage profiles of protein-coding genes were analyzed using the program MacVector 4.1.4 (Kodak, version 4.0). Translation initiation and translation termination codons were identified based on comparison with those reported previously for *A. suum*. The amino acid sequences inferred for the mt genes of three ascaridoids were aligned with those of *A. simplex* (GenBank™ accession number AY994152) and *A. suum* using Clustal X. Based on pairwise alignments, amino acid identity (%) was calculated for homologous genes. Codon usage was examined using the relationships between the nucleotide composition of codon families and amino acid occurrence, where the genetic codons are partitioned into AT-rich (i.e. those which are AT-rich at the first two codon positions), GC-rich codons (which are GC-rich at the first two codon positions) and unbiased codons. For analyzing ribosomal RNA genes, putative secondary structures and anticodon sequences by eye by aligning sequences with those of *A. simplex* and *A. suum*.

**Phylogenetic analyses**

Phylogenetic analyses were performed using the five ascaridoid species (*T. canis, T. cati, T. malaysiensis, A. suum, A. simplex*) as ingroups, and one filarialoid species (*O. volvulus*) serving as outgroup (GenBank™ accession number AF015193), based on amino acid sequences of 12 protein-coding genes. Amino acid sequences for each gene were examined using the program MacVector 4.1.4 (Kodak, version 4.0).
were individually aligned using Clustal X under default setting, and then concatenated into single alignments for phylogenetic analyses. Standard unweighted maximum parsimony (MP) were performed in PAUP* 4.0b10 [35] using heuristic searches with tree-bisection-reconnection branch swapping and 1000 random-addition sequence replicates with 10 trees held at each step. The Dayhoff matrix model was utilized in the analyses of neighbour joining (NJ), implemented by MEGA 3.1 [36], and maximum likelihood (ML) implemented by PhyML 2.1 [37]. Branch supports were estimated by bootstrap analysis of 1000 replicates for NJ and MP trees, and 100 replicates for ML tree.

Authors’ contributions
MWL performed the majority of the study and analyzed the data, and contributed to drafting of the manuscript. RQL and IHQS performed part of the study, and provided technical assistance. XYW contributed to the analysis of the data and helped in revising the manuscript. XQZ conceived and designed the research plan, participated in all aspects of the study, provided funds, supervised the research, and took the lead on drafting the manuscript. All authors read and approved the final manuscript.

Acknowledgements
Project support was provided to XQZ in part by grants from the Program for Changjiang Scholars and Innovative Research Team in University (grant no. IRT0723), the Special Funds for the Training of PhD Students from the Ministry of Education, China (grant no. 20040564008), and the Scientific Research Foundation for the Returned Overseas Chinese Scholars, the Ministry of Education, China (grant no. [2003] 406). Funding to pay the Open Access publication charges for this article was provided by the College of Veterinary Medicine, South China Agricultural University.

References

1. Wolstenholme DR: Animal mitochondrial DNA, structure and evolution. Int Rev Cytol 1992, 141:173-216.
2. Boore JL: Animal mitochondrial genomes. Nucleic Acids Res 1999, 27:1767-1780.
3. Avise JC: Molecular Markers, Natural History and Evolution, New York and London, Chapman and Hall; 1994:1-511.
4. Hu M, Chilton NB, Gasser RB: The mitochondrial genomics of parasitic nematodes of socio-economic importance: recent progress, and implications for population genetics and systematics. Adv Parasitol 2004, 56:133-212.
5. Hu M, Gasser RB: Mitochondrial genomics of parasitic nematodes – progress and perspectives. Trends Parasitol 2006, 22:78-84.
6. Ökimoto R, Macfarlane JL, Clary DO, Wolstenholme DR: The mitochondrial genomes of two nematodes, Caenorhabditis elegans and Ascaris suum. Genetics 1992, 130:471-498.
7. Keddie EM, Higazi T, Unnasch TR: The mitochondrial genome of Onchocerca volvulus : sequence, structure and phylogenetic analysis. Mol Biochem Parasitol 1998, 95:111-127.
8. Lavrov DV, Brown WM: 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 metazoa. Genetics 2001, 157:621-637.
9. Hu M, Chilton NB, Gasser RB: The mitochondrial genomics of the human hookworms, Ankylostoma duodenale and Necator americanus (Nematoda: Secernentea). Int J Parasitol 2002, 32:145-158.
10. Hu M, Chilton NB, Gasser RB: The mitochondrial genome of Strongyloides stercoralis (Nematoda) – idiosyncratic gene order and evolutionary implications. Int J Parasitol 2003, 33:1393-1408.
11. Hu M, Gasser RB, Abs EL-Osta YG, Chilton NB: Structure and organization of the mitochondrial genome of the canine heartworm, Dirofilaria immitis. Parasitology 2003, 127:37-51.
12. Van der Vries E, A single nucleotide polymorphism map of the mitochondrial genome of the parasitic nematode Cooperia oncophora. Parasitology 2004, 128:421-431.
13. He Y, Jones J, Armstrong M, Lamberi F, Moens M: The mitochondrial genome of Xiphinema americanus sensu stricto (Nematoda: Enoplea): considerations on the homology of the length and structural features of encoded genes. J Mol Evol 2005, 61:819-833.
14. Montiel R, Lucena MA, Medeiros J, Simões N: The complete mitochondrial genome of the entomopathogenic nematode Steinernema carpocapsae: insights into nematode mitochondrial DNA evolution and phylogeny. J Mol Evol 2006, 62:211-225.
15. Kim KH, Eom KS, Park JK: The complete mitochondrial genome of Anisakis simplex (Ascaridida: Nematoda) and phylogenetic implications. Int J Parasitol 2006, 36:129-138.
16. Habluetzel A, Traldi G, Ruggieri S, Attili AR, Scoppa P, Marchetti R, Menghini G, Esposito F: An estimation of Toxocara canis prevalence in dogs, environmental egg contamination and risk of human infection in the Marche region of Italy. Vet Parasitol 2003, 113:243-252.
17. Minnaar WN, Kreek RC, Fourie LJ: Helminths in dogs from a peri-urban resource-limited community in Free State Province, South Africa. Vet Parasitol 2002, 107:343-349.
18. Oliveira-Sequeira TC, Amarante AF, Ferrari TB, Nunes LC: Prevalence of intestinal parasites in dogs from São Paulo State, Brazil. Vet Parasitol 2002, 103:19-27.
19. Martínez-Barbosa S, Vázquez Tsui O, Cabello RR, Cárdenas EM, Chasis OA: The prevalence of Toxocara canis in domestic cats in Mexico City. Vet Parasitol 2003, 114:43-49.
20. Calvez P, Lucientes J, Castillo JA, Estrada R, Gracia MJ, Peribáñez MA, Ferrer M: Gastrointestinal helmhinth parasites in stray cats from the mid-Ebro Valley, Spain. Vet Parasitol 1998, 75:235-240.
21. Barbathe N, Serrão ML, Ferreira AM, Almeida NK, Guerrero J: A survey of gastrointestinal helminths in cats of the metropolitan region of Rio de Janeiro, Brazil. Vet Parasitol 2004, 123:133-139.
22. Despommier D: Toxocariasis: clinical aspects, epidemiology, molecular ecology, and molecular aspects. Clin Microbiol Rev 2003, 16(2):265-272.
23. Fisher M: Toxocara canis : an underestimated zoonotic agent. Trends Parasitol 2003, 19:167-170.
24. Vidal JE, Sztabnok J, Seguro AC: Eosinophilic meningencephalitis due to Toxocara canis: case report and review of the literature. Am J Trop Med Hyg 2003, 69:341-343.
25. Zhu XQ, Jacobs DE, Chilton NB, Sani RA, Cheng NA, Gasser RB: Molecular characterization of a Toxocara variant from cats in Kuala Lumpur, Malaysia. Parasitology 1998, 117:155-164.
26. Gibbons LM, Jacobs DE, Sani RA: Toxocara malayensis n. sp. (Nematoda: Ascaridoidea) from the domestic cat (Felis catus, Linnaeus, 1758). J Parasitol 2001, 87:660-665.
27. Ikemura T: Correlation between the abundance of yeast transfer RNAs and occurrence of respective codons in protein genes: differences in synonymous codon choice patterns of yeast and Escherichia coli with reference to the abundance of isoaccepting transfer RNAs. J Mol Biol 1982, 158:573-597.
28. Zhang DX, Hewitt GM: Insect mitochondrial control region. A review of its structure, evolution and usefulness in evolutionary studies. Biochem Symp 1997, 25:99-120.
29. Hartwich G: Keys to genera of the Ascaridoidea. In CH keys to the nematode parasites of vertebrates Volume 2. Edited by: Anderson RC, Chabaud AG, Willmott S. Farnham Royal, Bucks, Commonwealth Agricultural Bureaus; 1974:1-15.
30. Zhu XQ, Gasser RB, Jacobs DE, Hung GC, Chilton NB: Relationships among some ascaridoid nematodes based on ribosomal DNA sequence data. Parasitol Res 2000, 86:738-744.
31. Li MW, Lin RQ, Chen HH, Sani RA, Song HQ, Zhu XQ: PCR tools for the verification of the specific identity of ascaridoid nematodes from dogs and cats. Mol Cell Probes 2007, 21:349-354.

32. Li MW, Zhu XQ, Gasser RB, Lin RQ, Sani RA, Lun ZR, Jacobs DE: The occurrence of Toxocara malaysiensis in cats in China, confirmed by sequence-based analyses of ribosomal DNA. Parasitol Res 2006, 99:554-557.

33. Hu M, Chilton NB, Gasser RB: Long PCR-based amplification of the entire mitochondrial genome from single parasitic nematodes. Mol Cell Probes 2002, 16:261-267.

34. Lowe TM, Eddy SR: tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 1997, 25:955-964.

35. Swoford DL: PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods). Sunderland, MA, Sinauer Associates; 2002.

36. Kumar S, Tamura K, Nei M: MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 2004, 5:150-163.

37. Guindon S, Gascuel O: A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 2003, 52:696-704.