Helminth genome projects: all or nothing

Julius Lukeš, Aleš Horák and Tomáš Scholz

Institute of Parasitology, Czech Academy of Sciences and Faculty of Biology, University of South Bohemia, Branišovská 31,
370 05 České Budějovice, Czech Republic

Genomes of parasites that infect humans are being sequenced at a growing pace. Thanks to collaborative efforts and strong funding, the whole genomes of several protozoans have been sequenced, with *Plasmodium* spp., trypanosomes and *Leishmania major* leading the field [Box 1(i)]. In fact, the genome sequences of all protozoan parasites that cause the most serious infections in humans will be completed soon. It is hoped that these sequences, combined with functional genomics, will reveal novel drug targets, and open new and exciting lines of basic and applied research.

However, the situation with parasitic helminths is much less straightforward. It was announced recently that the genome sequence of *Brugia malayi* was completed [1] and that the sequences of two species of *Schistosoma* would soon be completed [2]. However, several medically important species have received little attention from molecular biologists, and some have not received any. While attempting to place the guinea worm (*Dracunculus medinensis*) into the 18S-rRNA-based phylogenetic tree, we noted that this nematode – which is supposed to represent the snake on the staff of Aesculapius, an ancient symbol of medicine – does not feature in the sequence databases. As the causative agent of dracunculosis, the guinea worm is responsible for the suffering of tens of thousands of people (Table 1) in an area ranging from Senegal to Eastern India. Moreover, not a single complete gene is available in public databases for any representative of the nematode superfamily Dracunculoidea, which encompasses more than 160 described species from all classes of vertebrate.

Box 1. Websites of interest

(i) Parasite-genome (http://www.ebi.ac.uk/parasites/paratable.html).
(ii) Graphic Images of Parasites (http://www.biosci.ohio-state.edu/~parasite/enterobius.html).
(iii) WHO Essential Medicines Library (nematodes – trichostrongyliasis) (http://mednet3.who.int/eml/disease_factsheet.asp?diseaseId=410).

Such a finding prompted us to inspect the presence in sequence databases of helminth parasites that infect humans (Table 1). Apart from the tapeworm *Echinococcus*

Table 1. The most important helminth parasites that infect humans and their presence in the GenBank database[a]

Species	Number of people infected (millions)	Number of nucleotide entries	Number of protein entries	Mitochondrial genome sequenced?
Tapeworms				
Taenia saginata	77 [4]	58	45	No
Hymenolepis nana	75 [4]	11	6	No
Taenia solium	10 [4]	247	244	Yes
Diphyllobothrium latum	9 [4]	5	1	No
Echinococcus spp.	2.7 [4]	10 758	628	Yes
Roundworms				
Ascaris lumbricoides	1472 [4]	1929	40	No
Ancylostoma and *Necator*	1298 [4]	44 083	216	Yes
Trichuris trichiura	1050 [4]	40	9	No
Enterobius vermicularis	400[b]	21	0	No
Wuchereria bancrofti	107 [4]	4896	38	No
Strongyloides stercoralis	70 [4]	11 434	48	Yes
Onchocerca volvulus	18 [4]	15 505	292	Yes
Brugia malayi	13 [4]	44 468	270	Yes
Loa loa	13 [4]	39	7	No
Trichinella spp.	11 [5]	10 852	98	Yes
Dracunculus medinensis	0.08 [4]	0	0	No
Trichostrongylus orientalis	Several[c]	0	0	No
Flukes				
Schistosoma spp.	200 [4]	264 602	1707	Yes
Paragonimus westermani	20.7 [4]	670	110	Yes
Opisthorchis spp.	10.3 [4]	22	8	No
Clonorchis sinensis	7 [4]	345	49	No
Fasciola hepatica	2.4 [4]	144	154	Yes
Fasciolopsis buski	0.2 [4]	2	0	No

[a] GenBank (http://www.ncbi.nlm.nih.gov/Genbank/index.html).
[b] See Box 1(ii).
[c] See Box 1(iii).

Corresponding author: Lukeš, J. (jula@paru.cas.cz).

Available online 26 April 2005

multilocularis – a rare parasite that infects humans, whose mitochondrial genome has been sequenced and that has 476 nucleotide entries in the databases – the information available for several nematodes (e.g. *Trichostrongylus orientalis* and *Trichuris trichiura*), flukes (e.g. *Fasciolopsis buski* and *Opisthorchis felineus*) and tapeworms (e.g. *Diphyllobothrium latum* and *Hymenolepis nana*) is negligible and is exemplified by just a few partial gene sequences. Even more striking is the case of the pinworm *Enterobius vermicularis*, probably the helminth that infects humans most frequently in industrialized countries [3] [Box 1(ii), Table 1]. An almost total absence of its sequences reveals that there is not a single laboratory that continuously studies the pinworm using molecular biology approaches. This is particularly surprising for a parasite that is known to have a major impact on the wellbeing of infants. Which factors or trends are behind such a gap in knowledge of the molecular biology of helminth parasites? Even if they have not been the subject of whole-genome-sequencing efforts, protozoan parasites that infect humans infrequently or cause less-serious infections have always attracted the interest of molecular parasitologists. Consequently, since the 1980s, at least a trickle of sequence information kept flowing into the databases. 18S rRNA genes are a must for every protozoan parasite that infects humans, and other phylogenetically informative sequences are also usually available.

Because of their geographical distribution, filariae such as *Loa loa* and *D. medinensis* are hard to obtain nonfixed and in sufficient quantity, thus hampering their suitability for molecular studies. However, this problem certainly does not apply to the pinworm and other globally present and frequent helminth parasites that remain underrepresented in, or even absent from, the databases. Is it the lack of a model host, legal aspects, ethical problems, technical difficulties, the large size of some helminth genomes or, simply, a matter of prioritization of limited funding that deters molecular biologists from touching these important organisms? Alternatively, is it the absence of a scientific community that would lobby for action and be able to exploit the information when it became available? It remains a challenge for parasitologists not only to identify the reasons for this unsatisfactory situation but also to draw the attention of molecular biologists towards the pinworm, guinea worm and other parasites.

References

1 Ghedin, E. *et al*. (2004) First sequenced genome of a parasitic nematode. *Trends Parasitol.* 20, 151–153

2 El-Sayed, N.M.A. *et al*. (2004) Advances in schistosome genomics. *Trends Parasitol.* 20, 154–157

3 Muller, R., ed. (2001) *Worms and Human Disease* (2nd edn), CABI Publishing

4 Crompton, D.W.T. (1999) How much human helminthiasis is there in the world? *J. Parasitol.* 85, 397–403

5 Dupouy-Camet, J. (2000) Trichinellosis: a worldwide zoonosis. *Vet. Parasitol.* 93, 191–200

1471-4922/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.pt.2005.04.004

The International Commission on Zoological Nomenclature (ICZN)

ICZN is dedicated to achieving standards, sense and stability for animal names in science.

ICZN is supported by the International Trust for Zoological Nomenclature (ITZN).

ICZN's work

- Produce, publish and revise the *International Code of Zoological Nomenclature*
- Consider and rule on specific cases of nomenclatural uncertainty
- Publish the *Bulletin of Zoological Nomenclature*

ICZN's targets

- Produce a user-friendly online version of the International *Code of Zoological Nomenclature*
- Transfer the *Bulletin of Zoological Nomenclature* online
- Produce online lists of all generic and specific animal names
- Develop a system for registration of valid new animal names
- Celebrate the 250th anniversary of Linnaean binominal nomenclature (in 2008)

The Commission's work directly affects

- Biodiversity studies
- Farming, fisheries and horticulture
- Medical, veterinary and parasitological research
- Palaeontology and biostratigraphy
- Entomology
- Conservation

For further information:
E-mail: iczn@nhm.ac.uk
Website: www.iczn.org