

Characterization of the *Schistosoma* transcriptome opens up the world of helminth genomics

Karl F Hoffmann and David W Dunne

Address: Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK.

Correspondence: Karl F Hoffmann. E-mail: kfh24@cam.ac.uk

Published: 24 December 2003

Genome Biology 2003, **5**:203

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2003/5/1/203>

© 2003 BioMed Central Ltd

Abstract

Among the metazoan parasites that cause debilitating disease in man, schistosomes are the first group for which near-complete transcriptome complements have been described. This new genomic information will have an enormous impact on all future investigations into the biology, pathogenesis and control of schistosomiasis.

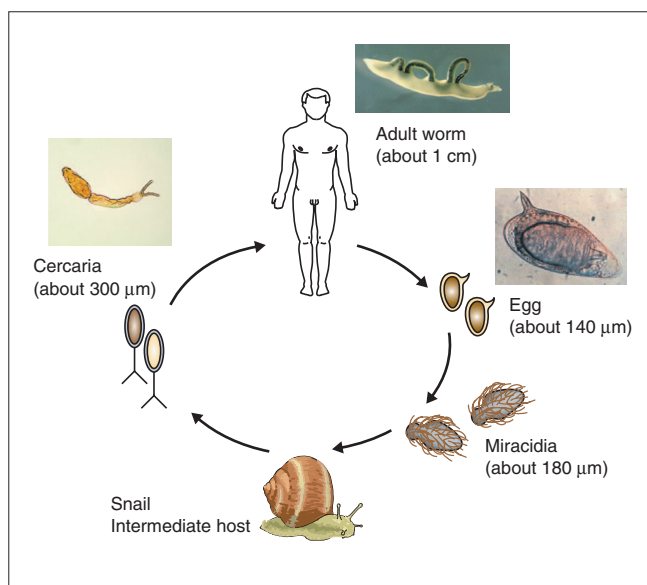
Schistosomes: the past and present

Schistosomes are members of the class Trematoda, a large, diverse and medically important group of parasitic worms within the phylum Platyhelminthes (flat worms). Ancestral trematodes were probably parasites of ancient marine molluscs. With the evolution of predatory vertebrate animals, however, the digenic forms of primitive trematodes developed complex life-cycles involving two or more different host species and both sexual and asexual reproduction. These parasites successfully diversified and expanded, accompanying the molluscs as they evolved into terrestrial and freshwater forms, and consequently all classes of vertebrate are now host to many trematode species. It is estimated that well in excess of 7,000 trematode species are parasitic on teleost fish, amphibians, reptiles, birds and mammals [1]. Trematodes of medical or veterinary importance include *Paragonimus westermani* (human lung fluke), *Clonorchis sinensis* (human liver fluke), *Fasciola hepatica* (cattle liver fluke) and, most importantly, the schistosomes (blood flukes).

We know that schistosomiasis (or bilharzia) has been a scourge of man for tens of thousands of years. Schistosome eggs have been found in ancient mummies in both Egypt and China. In 1851, working in Egypt, Theodore Bilharz first reported the presence of adult schistosome worms in human blood vessels. Several individual species that are infective to man were subsequently described, including *Schistosoma*

japonicum in South East Asia, by Katsurada in 1904, and *S. mansoni* and *S. haematobium* in Africa and the near East, by Sambon in 1907. These three species, of the five that cause human schistosomiasis, are responsible for the vast majority of human infections. Currently, it is estimated that at least 200 million human infections exist in 70 tropical and sub-tropical countries, leading to chronic debilitating disease and some 200-500 thousand human deaths per year [2]. Although an effective anti-schistosome drug, praziquantel, has been available for 25 years, the number of human infections worldwide stubbornly refuses to decline. This is partly because treatment is short-term efficacious, such that people remain susceptible to reinfection. Worryingly, praziquantel is the only drug widely available for treatment of human schistosomiasis; its exact mode of action is unknown, and the possible development of parasite drug-resistance is a concern.

Schistosomes have many features characteristic of the trematode group as a whole, yet are atypical in that the adult worms are dioecious, having both male and female forms. The schistosome life-cycle involves two distinct phases (Figure 1): asexual reproduction in the intermediate host, a freshwater snail, and sexual reproduction in the blood of the definitive host (man and other mammals). The miracidia and cercaria larvae are morphologically distinct, free-living, water-borne forms that traverse between the snail and mammalian hosts. Indeed, schistosomes reveal a truly astonishing

**Figure 1**

The schistosome life cycle. Schistosomes reproduce asexually in freshwater snails; a larval form, the cercaria, is released from the snail and can burrow into the skin of the definitive host, man. In humans, schistosomes migrate to the bloodstream where they mature into adult worms. Eggs produced by the female worm are released into the environment where they hatch into a second larval form, the miracidia, which can infect the snail. Adapted and reproduced with permission from [28].

morphological and physiological plasticity as they move from one highly specialized developmental stage to another.

Schistosomes are beautifully adapted to each of the many diverse environments encountered during their life-cycle, including the human host. As a result, they are especially difficult parasites to control. For example, not only are *S. mansoni* worm pairs able to produce hundreds of eggs daily in the immunologically hostile blood environment, they also provoke host immune responses that are essential for the passage of their eggs from the blood, through the host gut wall, to the environment outside. Thus, schistosomes modulate the human immune system to survive in the blood for years and subvert it to facilitate egg migration to continue their life-cycle. The combination of this complex, dioecious life-cycle and the inaccessible site of the worms within the mammalian host has prevented the use of classical genetics to investigate the role of individual genes in the highly sophisticated relationships between schistosomes and their intermediate and definitive hosts.

In the early 1990s a lack of genomic information prompted an international meeting to discuss the need for schistosome genome-related research. As a result, the *Schistosoma* Genome Network [3] was founded and global gene discovery projects for both *S. japonicum* and *S. mansoni* were initiated. Although redundancy in the generated sequence data remained consistently low, a comparison of expressed sequence tags (ESTs)

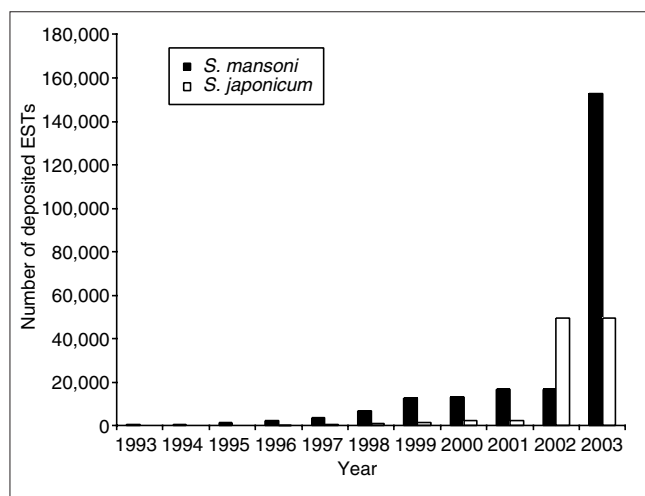
deposited annually since 1993 demonstrates that a numerical plateau of new entries funded by the *Schistosoma* Genome Network was quickly reached in 1999 for *S. japonicum* and in 2000 for *S. mansoni* (Figure 2). The work reported in two recent publications [4,5] has brought about a remarkable 20-fold increase in *S. japonicum* ESTs and an eight-fold increase in *S. mansoni* ESTs between July 2002 and October 2003, dramatically expanding both the diversity and coverage of these two parasite genomes (Figure 2). The wealth of structural information made available by these seminal sequencing efforts offers insight into dynamic processes related to developmental maturation, host-parasite interactions, signaling, tissue organization and sexual dimorphism. In addition to the molecular components associated with these important biological features, the newly generated transcriptome information offers an opportunity for advanced comparative and evolutionary investigations amongst the Platyhelminthes and across phyla.

Comparative transcriptome features of *Schistosoma* expressed sequence tags

The publication of the near-complete EST complements for *S. japonicum* [4] and *S. mansoni* [5] has added some 168,347 new schistosome sequences to GenBank [6]. The submitted sequences were derived from cDNA libraries created from diverse parasite source materials using different construction techniques. The *S. japonicum* project [4] used a directional phage vector for cDNA-library construction and generated ESTs from two different parasite life-stages (adult worm and egg). In contrast, the *S. mansoni* project [5] used both a normalized adult worm cDNA library and ORF expressed sequence tag (ORESTES) [7,8] mini-libraries from six life-cycle stages (adult, egg, miracidium, germ ball, cercaria, and cultured day-7 schistosomulum). Thus, stage-specific transcripts were more likely to be identified by the *S. mansoni* project, although a correlation cannot always be inferred between the number of EST reads for a given transcript in a specific life-cycle stage and differential expression [9].

The *S. mansoni* report [5] estimates, using two different methodologies, a total gene complement of between 13,960 and 14,205 unique elements. From the 30,988 unique assembled EST sequences (composed of 12,322 contigs and 18,666 singletons), the authors believe they have sampled roughly 92% of the *S. mansoni* transcriptome. These calculations suggest that approximately 12,880 distinctive *S. mansoni* open reading frames (ORFs) are reported in the article. As 13,131 unique clustered gene sets, representing approximately 93% gene coverage, were also identified in the *S. japonicum* report [4], similar expressed genomic complements are predicted for the two species.

The number of ESTs with significant database homology may also be similar in the two species. The *S. japonicum* project [4] reported that 65% of identified ESTs had significant database

**Figure 2**

Deposition of new *Schistosoma* EST structural information into public databases has dramatically increased the amount, diversity and coverage of both *S. mansoni* and *S. japonicum* genomes. Although the efforts of the *Schistosoma* Genome Network [3] led to an annual increase in the amount of non-redundant transcriptome information for both *S. mansoni* and *S. japonicum*, it was not until the EST sequencing projects of Hu *et al.* [4] and Verjovski-Almeida *et al.* [5] that a near-complete description of each parasite's genomic complement was obtained. Together, this transcriptome information will lead to exponential advances in the understanding of parasitic helminths and will provide vital structural components to be used in post-genomic research activities (see Table 1).

similarity to sequences already in GenBank, compared with 45% reported by the *S. mansoni* project [5]. Distinct BLASTx cut-off Expect (*E*) values ($E \leq 10^{-6}$ for *S. mansoni* and $E \leq 10^{-3}$ for *S. japonicum*) were used in the two studies for assigning sequence similarity and are therefore likely to be responsible for this apparent difference. Taken together, the observations indicate that a significant proportion of each species' genome (approximately 50%) is schistosome-specific and may be adapted for gene expression related to parasitism. The proportion of these 'schistosome-associated genes' shared between the two species is currently unknown, but careful comparative analyses will undoubtedly reveal much about common schistosome adaptations to parasitic lifestyles. Phylogenetic analyses of both species (using highly conserved genes) indicate an early and independent divergence of schistosomes from other metazoans, which may explain the high fraction of these ESTs that display no significant database match in GenBank. It is clear that the addition of these new schistosome sequences to the public databases will contribute substantially to our understanding of evolution and the parasitic lifestyle.

The diversity of expressed *Schistosoma* genes

Functional classification of both parasites' genomic complement, using Interpro [10], to identify protein motifs, and Gene Ontology [11], to describe the cellular processes, predicts that the genetic machinery responsible for most cellular

and physiological systems was established long before the divergence of the Platyhelminths. Although gene products associated with diverse biological functions - such as cell adhesion/tissue structure, antero-posterior axis differentiation, dorso-ventral patterning, epithelial interactions, motility, nervous system development and signaling - are present in both *S. mansoni* and *S. japonicum*, transcripts involved in metabolism and catabolism were among the most frequently characterized. The identification of these highly abundant transcripts strongly supports classical biochemical studies that have suggested that schistosomes may be completely dependent on host molecules (lipids, proteins and nucleic acids) for catabolically released energy and for the building blocks used during parasite development [12].

Some of the most interesting newly identified transcripts are those related to development, gender and host-parasite interactions. These transcripts may provide important insights into the evolution of morphologically distinct genders [13], the complete dependence of female parasites on male contact for sexual maturation [14,15], and the mechanisms by which schistosomes evade [16] or modulate [17-20] host immune responses. New information of this sort will undoubtedly push forward disease intervention and parasite-control strategies. Similarly, the use of comparative genomic strategies to determine the level of these transcripts in hermaphroditic trematodes should offer insights into the evolution of dioecy in this group of organisms. For example, the identification of orthologs of *Caenorhabditis elegans* sex-determination genes (*fox-1*, *mog-1*, *mog-4*, *tra-2*, *fem-1*, and *mag-1*) in *S. mansoni* [5], and the classification of definitive-host associated X- and Y-linked orthologs in *S. japonicum* [4], implicate these transcripts in schistosome sexual-development pathways.

There are schistosome ESTs showing sequence similarity to components of the endocrine system (for example, homologs of the insulin receptor, insulin-like growth factor 1 receptor and insulin signaling-pathway proteins), and of the immune system (for example, homologs of cytokines, cytokine-like proteins, major histocompatibility complex and allergens); these similarities strongly suggest that the parasites actively interact with host biochemical and defense mechanisms, rather than merely evading them. Modulation of host biomolecules and cellular systems has been documented [17-21], and the identification of novel ESTs that may function in an analogous manner points to further complex, as yet undiscovered, interactions between schistosomes and their hosts.

The schistosome transcriptome also contains EST homologs of stress-response proteins (novel heat-shock proteins, antioxidant enzymes, Daf16 and Toll-like molecules) and longevity-associated proteins (for example Sir-2.1, Sir-2.2, Sir-2.5, Sir-2.6 and Sir-2.7) suggesting that there are multiple genetic programs to ensure parasite survival. For example, if parasite immunomodulatory molecules fail to

Table 1**Areas in which *Schistosoma* transcriptome information will be applied during the post-genomic era**

Resource	Application
Fabrication of whole transcriptome DNA microarrays	Resource for expression-profiling experiments, for example, across life-stages, between genders and so on
Construction of RNA interference (RNAi) libraries	Resource for the functional characterization of each deposited transcript
Initiation of whole-genome sequencing projects	Blueprint/template resource for initiation/completion of whole-genome sequencing projects
Characterization of proteome	Blueprint/template resource for identification of differentially expressed proteins, for example, across life-stages, between genders and so on
Investigation of evolutionary relationship amongst the metazoa	Resource for advanced phylogenetic studies
Identification of orthologous transcripts within or outside the genera	Resource for comparative genomic studies

prevent the host's immune system from recognizing the parasite and activating a response, parasite stress-response transcripts may be induced to counteract the host immune-effector mechanisms. This has previously been proposed for the antioxidant enzymes glutathione peroxidase, superoxide dismutase and glutathione-S-transferase in *S. mansoni* [22]. Similarly, if the schistosome is prevented from using its host's endocrine system, the parasite may temporarily induce genetic programs associated with longevity. Further characterization of these and other transcripts associated with hormone signaling, neurotransmission and neurogenesis present excellent opportunities to investigate novel anti-schistosome chemotherapeutic strategies. For example, the schistosome has homologs of progesterone receptor membrane components, small androgen receptor-interacting proteins, retinoic acid receptors, thyroid hormone receptor family members and progestin-induced proteins that may have endocrine function; and homologs of the neuropeptide Y₂ receptor, nicotinic acetylcholine receptor, glutamate receptor, serotonin receptor, muscarinic acetylcholine receptor, Notch receptor and calcium channels that may have neuronal functions. Surprisingly, very few G-protein-coupled-receptor paralogs (rhodopsin, glutamate, serotonin, and muscarinic acetylcholine receptors) were reported in either transcriptome, suggesting that schistosomes may have a limited repertoire and may instead rely on other mechanisms to interact with environmental or self stimuli. Functional studies of these and other parasite molecules intimately involved in host-parasite interactions should shed light on the elegant schistosome survival strategies that enable them to live for years in the bloodstreams of their definitive hosts [23].

The future of schistosome post-genomic studies

Newly available schistosome transcriptome data will facilitate high-throughput, post-genomic activities in many laboratories and will revolutionize comparative studies of the Platyhelminthes as a whole (Table 1). Soon, RNA interference

(RNAi) libraries of individual schistosome genes will be constructed to allow the functional characterization of differentially expressed transcripts identified by complete-transcriptome DNA microarrays. As both technologies (post-translational gene silencing and gene-expression profiling) have been suitably adapted to the study of schistosome biology [24-26], this is now an attainable goal. In addition, linking the genome and the proteome, and the use of proteomic approaches for the identification of individual schistosome polypeptides, will be greatly accelerated by these newly deposited EST sequences.

The recent detailed transcriptome data will also provide an integral framework for the construction and completion of a full schistosome genomic map, which is currently underway at The Institute for Genomic Research (USA) and the Wellcome Trust Sanger Institute (UK), where it is expected that eightfold (8X) coverage of the *S. mansoni* genome will soon be obtained [27]. This will ultimately open the path to previously unknown and valuable information about the genomes of a relatively under-studied, but biologically diverse and fascinating, phylum that contains species causing some of the world's most important and obdurate parasitic diseases.

Acknowledgements

We thank Rhian Hayward and Jennifer Fitzpatrick for critical editorial comments.

References

1. Basch PF: *Schistosomes: Development, reproduction, and host relations*. New York: Oxford University Press, 1991.
2. Johnston DA, Blaxter ML, Degraeve WM, Foster J, Ivens AC, Melville SE: **Genomics and the biology of parasites**. *Bioessays* 1999, **21**:131-147.
3. **Schistosoma Genome Network** [http://www.nhm.ac.uk/hosted_sites/schisto/]
4. Hu W, Yan Q, Shen DK, Liu F, Zhu ZD, Song HD, Xu XR, Wang ZJ, Rong YP, Zeng LC, et al.: **Evolutionary and biomedical implications of a *Schistosoma japonicum* complementary DNA resource**. *Nat Genet* 2003, **35**:139-147.

5. Verjovski-Almeida S, DeMarco R, Martins EA, Guimaraes PE, Ojopi EP, Paquola AC, Piazza JP, Nishiyama MY Jr, Kitajima JP, Adamson RE, et al.: **Transcriptome analysis of the acelomate human parasite *Schistosoma mansoni***. *Nat Genet* 2003, **35**:148-157.
6. **GenBank** [<http://www.ncbi.nlm.nih.gov/Genbank/index.html>]
7. de Souza SJ, Camargo AA, Briones MR, Costa FF, Nagai MA, Verjovski-Almeida S, Zago MA, Andrade LE, Carrer H, El-Dorry HF, et al.: **Identification of human chromosome 22 transcribed sequences with ORF expressed sequence tags**. *Proc Natl Acad Sci USA* 2000, **97**:12690-12693.
8. Dias Neto E, Correa RG, Verjovski-Almeida S, Briones MR, Nagai MA, da Silva W, Jr., Zago MA, Bordin S, Costa FF, Goldman GH, et al.: **Shotgun sequencing of the human transcriptome with ORF expressed sequence tags**. *Proc Natl Acad Sci USA* 2000, **97**:3491-3496.
9. Ajioka JW, Boothroyd JC, Brunk BP, Hehl A, Hillier L, Manger ID, Marra M, Overton GC, Roos DS, Wan KL, et al.: **Gene discovery by EST sequencing in *Toxoplasma gondii* reveals sequences restricted to the Apicomplexa**. *Genome Res* 1998, **8**:18-28.
10. **InterPro home** [<http://www.ebi.ac.uk/interpro/>]
11. **Gene Ontology Consortium** [<http://www.geneontology.org/>]
12. Bueding E: **Physiology and biochemistry of schistosomes**. *J Parasitol* 1970, **56**:558-559.
13. Platt TR, Brooks DR: **Evolution of the schistosomes (Digenea: Schistosomatoidea): the origin of dioecy and colonization of the venous system**. *J Parasitol* 1997, **83**:1035-1044.
14. Shaw JR, Erasmus DA: ***Schistosoma mansoni*: an examination of the reproductive status of females from single sex infections**. *Parasitology* 1981, **82**:121-124.
15. Shaw MK: ***Schistosoma mansoni*: vitelline gland development in females from single sex infections**. *J Helminthol* 1987, **61**:253-259.
16. Pearce EJ, Sher A: **Mechanisms of immune evasion in schistosomiasis**. *Contrib Microbiol Immunol* 1987, **8**:219-232.
17. Falcone FH, Dahinden CA, Gibbs BF, Noll T, Amon U, Hebestreit H, Abrahamsen O, Klaucke J, Schlaak M, Haas H: **Human basophils release interleukin-4 after stimulation with *Schistosoma mansoni* egg antigen**. *Eur J Immunol* 1996, **26**:1147-1155.
18. Haisch K, Schramm G, Falcone FH, Alexander C, Schlaak M, Haas H: **A glycoprotein from *Schistosoma mansoni* eggs binds non-antigen-specific immunoglobulin E and releases interleukin-4 from human basophils**. *Parasite Immunol* 2001, **23**:427-434.
19. Ramaswamy K, Kumar P, He YX: **A role for parasite-induced PGE2 in IL-10-mediated host immunoregulation by skin stage schistosomula of *Schistosoma mansoni***. *J Immunol* 2000, **165**:4567-4574.
20. Rao KV, Chen L, Gnanasekar M, Ramaswamy K: **Cloning and characterization of a calcium-binding, histamine-releasing protein from *Schistosoma mansoni***. *J Biol Chem* 2002, **277**:31207-31213.
21. Herve M, Angeli V, Pinzar E, Wintjens R, Faveeuw C, Narumiya S, Capron A, Urade Y, Capron M, Riveau G, Trottein F: **Pivotal roles of the parasite PGD2 synthase and of the host D prostanoid receptor 1 in schistosome immune evasion**. *Eur J Immunol* 2003, **33**:2764-2772.
22. Mei H, LoVerde PT: ***Schistosoma mansoni*: the developmental regulation and immunolocalization of antioxidant enzymes**. *Exp Parasitol* 1997, **86**:69-78.
23. Fulford AJ, Butterworth AE, Ouma JH, Sturrock RF: **A statistical approach to schistosome population dynamics and estimation of the life-span of *Schistosoma mansoni* in man**. *Parasitology* 1995, **110**:307-316.
24. Boyle JP, Wu XJ, Shoemaker CB, Yoshino TP: **Using RNA interference to manipulate endogenous gene expression in *Schistosoma mansoni* sporocysts**. *Mol Biochem Parasitol* 2003, **128**:205-215.
25. Skelly PJ, Da'dara A, Harn DA: **Suppression of cathepsin B expression in *Schistosoma mansoni* by RNA interference**. *Int J Parasitol* 2003, **33**:363-369.
26. Hoffmann KF, Johnston DA, Dunne DW: **Identification of *Schistosoma mansoni* gender-associated gene transcripts by cDNA microarray profiling**. *Genome Biol* 2002, **3**:research0041.1-0041.12.
27. **The *Schistosoma mansoni* genome project** [http://www.sanger.ac.uk/Projects/S_mansoni/]
28. Dunne D, Vennervald B: **Trematodes - Schistosomiasis**. In *Oxford Textbook of Medicine. Volume 1*. 4th Edition. Edited by Warrell D, Cox TM, Firth JD. Oxford: Oxford University Press; 2003:830-837.