Helminth genomics: The implications for human health

Paul J. Brindley
George Washington University

Makedonka Mitreva
Washington University School of Medicine in St. Louis

Elodie Ghedin
University of Pittsburgh - Main Campus

Sara Lustigman
New York Blood Center, Laboratory of Molecular Parasitology, New York

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Part of the Medicine and Health Sciences Commons

Recommended Citation
https://digitalcommons.wustl.edu/open_access_pubs/779

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact engeszer@wustl.edu.
Helminth Infections—The Great Neglected Tropical Diseases

Helminth parasites are parasitic worms from the phyla Nematoda (roundworms) and Platyhelminthes (flatworms) (Figures 1 and 2); together, they comprise the most common infectious agents of humans in developing countries. The collective burden of the common helminth diseases—which range from the dramatic sequelae of elephantiasis and blindness to the more subtle but widespread effects on child development, pregnancy, and productivity—rivals that of the main high-mortality conditions such as HIV/AIDS or malaria [1]. For example, based on a recent analysis [2], 85% of the neglected tropical disease (NTD) burden for the poorest 500 million people living in sub-Saharan Africa (SSA) results from helminth infections. Hookworm infection occurs in almost half of the poorest people in SSA, including 40–50 million school-aged children and 7 million pregnant women, in whom it is leading a cause of anemia. Schistosomiasis (192 million cases) is the second most prevalent NTD after hookworm, accounting for 93% of the world’s number of cases of schistosomiasis and possibly associated with increased horizontal transmission of HIV/AIDS. Lymphatic filariasis (46–51 million cases) and onchocerciasis (37 million cases) are also widespread in SSA, each disease representing a significant cause of disability and reduction in the region’s agricultural productivity. The disease burden estimate in disability-adjusted life years (DALYs) for total helminth infections in SSA is 5.4–18.3 million in comparison to 40.9 million DALYs for malaria and 9.3 million DALYs for tuberculosis. Yet, research into helminth infections has not received nearly the same level of support. This is partly because helminthiases are diseases of the poorest people in the poorest regions, but also because these pathogens are difficult to study in the laboratory by comparison to most model eukaryotes and many other pathogens. Standard tools and approaches, including cell lines, culture in vitro, and animal models, are generally lacking. In addition, the genomes of helminths are generally much more complex than those of model organisms like yeast and fruit flies [2].

Whereas helminth diseases are ancient scourges of humanity, with some known from biblical times, most can also be considered as re-emerging diseases in the sense that new outbreaks are reported routinely in response to environmental and sociopolitical changes [3]. For example, schistosomiasis has reemerged many times in Africa in recent times in response to hydrological changes, e.g. construction of dams, irrigation canals, reservoirs, etc. that establish suitable new environments for the intermediate host snails that transmit the parasites. Schistosomiasis has also reemerged in mountainous and hilly regions in Sichuan, China, where it had been controlled previously by intensive interventions [4]. Furthermore, new strains of schistosomes are indeed emerging through natural hybridizations between human and cattle species of schistosomes [5].

Despite the difficulties with investigation of helminth parasites, new insights into fundamental helminth biology are accumulating through genome projects and the application of genome manipulation technologies including RNA interference and transgenesis (Figure 3). What’s more, research on immunology of helminth infections has contributed enormously to our understanding of Th2 immune responses, the function of regulatory T cells, generation of alternatively activated macrophages, and the transmission dynamics of infectious agents. It is hoped that this progress can be translated into new and robust drugs, diagnostics, and vaccines for the helminth diseases.
of humanity and those of our livestock and companion species [1,6–10].

Genomics Approaches to Investigating Helminths

Over the past decade, increasing numbers of helminth-specific genome sequences have become available due to ever-improving techniques for obtaining biological material, extracting RNA and DNA, constructing complimentary DNA (cDNA)/whole genome shotgun libraries, and, especially, major advances in the chemistry and instrumentation for DNA sequencing and its concomitant decreased cost. Helminth genomics began with the generation and analysis of transcribed sequences (expressed sequence tags [ESTs] [11]), which has proved to be a rapid and cost-effective route to discover genes in other eukaryotes. In April 2009, there were 550,000 nematode and 450,000 platyhelminth ESTs in the dbEST division of GenBank, excluding those from the model nematode Caenorhabditis rhabditis. Of these, 60% were from parasites of humans and closely related animal pathogens used to study human infections (Table 1). These ESTs have many applications. They can be used to annotate helminth genomes (see below) to determine alternative splicing, verify open reading frames, and confirm exon/intron and gene boundaries. They are valuable also, for example, in functional genomics to design probes for expression microarrays [e.g., [12]] and to provide putative protein sequence information for proteomics methods [e.g., [13]], to name but a few applications. Quantitative analysis of ESTs (transcriptomics), including serial expression of gene analysis, can identify transcripts that are either over- or under-represented by comparison to other transcripts in various helminth life cycle stages or tissues [e.g., [14,15]], and the subset of genes evaluated with gene ontology programs provide insights into cellular and metabolic pathway functioning in the parasite [e.g., [16]]. Furthermore, one can identify potential targets for interventions by applying a hierarchy of considerations including a matrix of biological, expression, and phenotypic data [17] or by performing a pan-phylum analysis to identify conserved parasite-specific genes whose selective targeting will have low or no toxicity to the host [18,19] or genes that have diverged enough from the host counterpart, resulting in altered or absent functions [20].

The first multicellular genome sequenced was that of the free-living roundworm C. elegans [21]; reported in 1998, it is still the only metazoan for which the sequence of every nucleotide is known with high confidence. Today, the genome sequences of 22 species of helminths that either infect humans or are closely related parasites are completed or underway (Table 1). A comprehensive genome analysis has been published for several of them, including the lymphatic filarial nematode Brugia malayi [22], the dog hookworm Ancylostoma caninum [23], and the blood flukes Schistosoma japonicum and S. mansoni [24,25] (Figure 1; Table 1).
Some of the main obstacles to research on human parasites are their life cycle complexity, tissue complexity, and the paucity of genetic and transgenic methods for manipulating genes of interest. Comparative genome analyses have also provided insights into the adaptations of various parasites to niches in their human (and vector) hosts as well as insights into the molecular basis of the mutualistic relationship between the filarial nematode *B. malayi* and its endosymbiont *Wolbachia* (see below).

The genomes of the schistosomes *S. japonicum* and *S. mansoni* are the first complete genomes reported for members of the Lophotrochozoa [24,25], a large taxon that includes about 50% of all metazoan phyla including the mollusks, annelids, brachiopods, nemerteans, bryozoans, playthelminths, and others [26]. These schistosome genome sequences revealed remarkable features of the host–parasite relationship. Among these, the schistosome genome has lost numerous protein-encoding domains. Whereas the total number (~6,000) of protein families is broadly similar among schistosomes, humans, *C. elegans*, and fruit fly, about 1,000 protein domains have been abandoned by *S. japonicum*, including some involved in basic metabolic pathways and defense, implying that loss of these domains could be a consequence of the adoption of a parasitic way of life. If so, the remaining molecular repertoire must have evolved in parallel with this extensive domain loss to permit the pathogen to locate and infect humans efficiently, nourish itself, and interact with the external environment as well as with the host. On the other hand, despite extensive gene and domain loss, a number of schistosome gene families have expanded and these provide insights into the requirements for a parasitic lifestyle. Among the expanded gene families, a metallo-protease called invadolysin (or leishmanolysin) has at least 12 putative family members in schistosomes compared to a single orthologue in the human, fruit fly, and *C. elegans* genomes and only three in the free-living flatworm *S. mediterranea*. This protease family may facilitate skin penetration and tissue invasion by the cercaria, the infective-stage larva of the schistosome [24,25].

Publication of genome sequences of *filaria* and schistosomes has underscored the pressing need to develop functional genomics approaches for these significant pathogens. Functional analyses—which use approaches such as RNA interference (RNAi) and translational studies—are essential to resolve uncertainties in the molecular physiology of helminths and to illuminate mechanisms of pathogenesis that may lead to development of new interventions to control and eliminate these parasites or the diseases. Progress in the functional genomics of helminths was reviewed recently [6,27,28]. In brief, RNAi has been used to inactivate the RNA products of several genes in schistosomes (e.g., [29–32]) and nematodes (e.g., [33]; reviewed in [8]). In addition, the recent genome sequences of *S. mansoni* and *S. japonicum* now make feasible genome-scale investigation of transgene integration into schistosome chromosomes. Gene therapy–like approaches to transform schistosomes include the use of the piggyBac transposon and pseudotyped murine leukemia retrovirus as transgene vectors.
(Figure 3A), both of which offer a means to establish transgenic lines of schistosomes, to elucidate schistosome gene function and expression, and to advance functional genomics approaches for these parasites. Notably, progress is also being made to express reporter transgenes in parasitic nematodes including *Strongyloides stercoralis* [37], in which transgene approaches developed for use in *C. elegans* have recently been used to demonstrate that morphogenesis of infective larvae requires the DAF-16 orthologue FKTF-1 (Figure 3B) [38]. Progress is also being made with systems for analysis of promoter sequences of genes of parasitic nematodes (e.g., [39]).

Many future discoveries resulting from the parasitic helminths genome information can be expected to emanate from the broader scientific community rather than by the laboratory originating a genome sequence project. For the specialized genome sequence labs, dissemination of sequence information in a way that is useful, consistent, centralized, and lasting has been therefore a key goal. Efforts have gone well beyond depositing raw data in public databases. Currently, helminthologists have available a number of specialized sites for sequence analysis. *C. elegans* information is easily accessible at http://www.wormbase.org [40]. Useful information about the organism includes genome sequence, genetic and physical maps, transcript data (EST, mRNA, SAGE, TEC-RED, ORFeome, expression patterns from reporter gene fusions, and microarrays), the developmental lineage of all cells, connectivity of the nervous system, mutant phenotypes and genetic markers, gene expression described at the level of single cells, 3D protein structures, NCBI Clusters of Orthologous Groups, and apoptosis and aging information. It also contains extensive information from large-scale genomics analyses, including pre-computed sequence similarity searches, protein motif analyses, protein-protein interactions, findings from systematic RNAi screens, single nucleotide polymorphisms (SNPs), orthologous and paralogous relationships, and the assignment of Gene Ontology (GO) terms to gene products. These resources greatly aid in the interpretation of much of the sequence data emerging from parasitic helminths.

However, accumulating evidence suggest that *C. elegans* is not a good model for all parasitic helminths, especially for the ones that are phylogenetically very distant such as the basic nematode and zoonotic parasite *Trichinella spiralis* (e.g., [41]). The other specialized site is Nematode.net (http://www.nematode.net) [42], developed with a primary aim to disseminate the diverse collection of information for parasitic nematodes to the broader scientific community in a way that is useful, consistent, centralized, and enduring. In addition to sequence data, the site hosts assembled NemaGene clusters in GBrowse views, characterizing composition and protein homology, functional Gene Ontology annotations presented via the AmiGO browser, KEGG-based graphical display of NemaGene clusters mapped to metabolic pathways, codon usage tables, NemFam protein families (which represent conserved nematode-restricted coding sequences not found in public protein databases), and a Web-based WU-BLAST search tool that allows complex querying and other assorted resources. Furthermore, Nematode.net, by connecting data across the entire phylum Nematoda, has made a substantial contribution toward integrating the historically separate fields of *C. elegans*, vertebrate parasitology, and plant parasitology research. Finally,
<table>
<thead>
<tr>
<th>Phylum or Class</th>
<th>Species</th>
<th>Common Name / Disease</th>
<th>Primary Host</th>
<th>Genome size, Mb</th>
<th>GenBank Project ID</th>
<th>cDNAs (37/30 ABI), 1,000 s</th>
<th>Genome Sequencing Status</th>
<th>Sequencing Institutea</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nematoda</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(roundworms)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nematoda (roundworms)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clade V</td>
<td>Necator americanus</td>
<td>Hookworm/necatoriasis</td>
<td>Human</td>
<td>—</td>
<td>20369</td>
<td>5</td>
<td>In progress</td>
<td>WUGC</td>
</tr>
<tr>
<td>Clade IV</td>
<td>Ancylostoma caninum</td>
<td>Model hookworm</td>
<td>Dog</td>
<td>344</td>
<td>12841</td>
<td>81</td>
<td>Improving draft</td>
<td>WUGC</td>
</tr>
<tr>
<td>Clade IV</td>
<td>Nippostrongylus brasiliensis</td>
<td>Model hookworm</td>
<td>Rat</td>
<td>—</td>
<td>20445</td>
<td>14.7</td>
<td>In progress</td>
<td>SI</td>
</tr>
<tr>
<td>Clade III</td>
<td>Strongyloides stercoralis</td>
<td>Threadworm/strongyloidiasis</td>
<td>Human</td>
<td>—</td>
<td>—</td>
<td>11.4</td>
<td>In progress</td>
<td>SI</td>
</tr>
<tr>
<td>Clade III</td>
<td>S. ratti</td>
<td>Model threadworm</td>
<td>Rat</td>
<td>—</td>
<td>—</td>
<td>27.4</td>
<td>In progress</td>
<td>SI/WUGC</td>
</tr>
<tr>
<td>Clade III</td>
<td>Ascaris lumbricoides</td>
<td>Large roundworm/ascariasis</td>
<td>Human</td>
<td>230</td>
<td>—</td>
<td>1.8</td>
<td>In progress</td>
<td>SI</td>
</tr>
<tr>
<td>Clade III</td>
<td>A. suum</td>
<td>Model large roundworm</td>
<td>Pig</td>
<td>230</td>
<td>—</td>
<td>55.7</td>
<td>Improving draft</td>
<td>WUGC/Sl</td>
</tr>
<tr>
<td>Clade III</td>
<td>Brugia malayi</td>
<td>Filaria/lymphatic filariasis</td>
<td>Human</td>
<td>96</td>
<td>9549</td>
<td>26.2</td>
<td>Improving draft</td>
<td>TIGR/University of Pittsburgh</td>
</tr>
<tr>
<td>Clade I</td>
<td>Loa Loa</td>
<td>Filaria/toasis (cutaneous filariasis)/African eye worm</td>
<td>Human</td>
<td>—</td>
<td>—</td>
<td>3.3</td>
<td>In progress</td>
<td>SI</td>
</tr>
<tr>
<td>Clade I</td>
<td>Onchoerca volvulus</td>
<td>Filaria/river blindness</td>
<td>Human</td>
<td>150</td>
<td>—</td>
<td>15</td>
<td>In progress</td>
<td>SI</td>
</tr>
<tr>
<td>Clade I</td>
<td>Acanthocheilemona vitae</td>
<td>Model filaria</td>
<td>Rodent</td>
<td>—</td>
<td>33239</td>
<td>0</td>
<td>In progress</td>
<td>UMIGS</td>
</tr>
<tr>
<td>Clade I</td>
<td>Trichinella spiralis</td>
<td>Trichina worm/trichinosis</td>
<td>Pig to human</td>
<td>71</td>
<td>12605</td>
<td>25.3</td>
<td>Draft completed</td>
<td>WUGC</td>
</tr>
<tr>
<td>Clade I</td>
<td>Trichuris trichiura</td>
<td>Whipworm/trichuriasis</td>
<td>Human</td>
<td>—</td>
<td>—</td>
<td>0</td>
<td>In progress</td>
<td>SI</td>
</tr>
<tr>
<td>Clade I</td>
<td>T. muris</td>
<td>Model whipworm</td>
<td>Mouse</td>
<td>96</td>
<td>—</td>
<td>7</td>
<td>In progress</td>
<td>SI</td>
</tr>
<tr>
<td>Clade I</td>
<td>T. suis</td>
<td>Model whipworm</td>
<td>Pig</td>
<td>—</td>
<td>—</td>
<td>0</td>
<td>In progress</td>
<td>WUGC</td>
</tr>
<tr>
<td>Clade I</td>
<td>Echinoccoccus multilocularis</td>
<td>Tapeworm/aloeveral hydatidosis</td>
<td>Rodent; larva infects humans</td>
<td>150</td>
<td>—</td>
<td>1</td>
<td>In progress</td>
<td>SI</td>
</tr>
<tr>
<td>Clade I</td>
<td>E. granulosus</td>
<td>Tapeworm/unilocular hydatidosis</td>
<td>Canids; larva infects humans</td>
<td>150</td>
<td>12620</td>
<td>10</td>
<td>In progress</td>
<td>SI</td>
</tr>
<tr>
<td>Clade I</td>
<td>Taenia solium</td>
<td>Pork tapeworm/taenisias/cysticercosis</td>
<td>Human</td>
<td>270</td>
<td>17815</td>
<td>25</td>
<td>Draft completed</td>
<td>Mexico City</td>
</tr>
<tr>
<td>Clade I</td>
<td>Schistosoma mansoni</td>
<td>Blood fluke/intestinal schistosomiasis</td>
<td>Human</td>
<td>390</td>
<td>12599</td>
<td>206</td>
<td>Draft completed</td>
<td>SI/TIGR</td>
</tr>
<tr>
<td>Clade I</td>
<td>S. haematobium</td>
<td>Blood fluke/urinary schistosomiasis</td>
<td>Human</td>
<td>—</td>
<td>12616</td>
<td>0</td>
<td>In progress</td>
<td>SI</td>
</tr>
<tr>
<td>Clade I</td>
<td>S. japonicum</td>
<td>Blood fluke/intestinal schistosomiasis</td>
<td>Human</td>
<td>400</td>
<td>29491</td>
<td>104</td>
<td>Draft completed</td>
<td>CNHGC</td>
</tr>
<tr>
<td>Clade I</td>
<td>Clonorchis sinensis</td>
<td>Liver fluke/clonorchiasis</td>
<td>Human</td>
<td>—</td>
<td>17975</td>
<td>3</td>
<td>In progress</td>
<td>SNUCM</td>
</tr>
</tbody>
</table>

*WUGC, Washington University's Genome Center.
bPhylogeny based on Blaxter et al. [47].

CI, Broad Institute; CNHGC, Chinese National HGC; SI, Sanger Institute; SNUCM, Seoul National University College of Medicine; TIGR, The Institute for Genomic Research (now JCVI). doi:10.1371/journal.pntd.0000538.t001
Evolution of Parasitism in Helminths

Genomics research has helped our understanding of the evolution of helminths of humans and other hosts, certainly with regard to roundworms of the phylum Nematoda. The first comprehensive study of the molecular evolution of helminths was a phylogenetic analysis of the small subunit ribosomal DNA (ss rDNA) sequences from 53 roundworms [47]. This study included both major parasitic and free-living taxonomic groups. It identified five major clades within the Nematoda and suggested that parasitism of animal and plants arose independently multiple times. A more recent study included 339 nearly full-length ss rDNAs and proposed subdivision of the phylum into 12 clades [48]. This revealed that nematodes that feed on fungi occupy a basal position compared to their plant parasitic relatives, confirming that the parasitic nematodes of plants arose from fungivorous ancestors. Phylogenetic methods are also being used to study evolution of parasitism-related protein-coding genes (such as the enzymes that degrade the plant cell wall in nematode parasites of plants [cellulases, pectate lyases, etc.]) to understand better the mechanisms underlying the evolution of parasitism [reviewed in [49]]. Recent genome-wide analysis of two plant parasitic nematodes [50,51] provided a more complete picture of the acquisition of these cellulase genes, apparently by horizontal gene transfer (HGT) from prokaryotes. The subsequent expansion and diversification of HGT genes in these nematodes allow inferences about the evolutionary history of these parasites, and in addition present potential targets for anti-nematode drugs. When the genome of the necromenic nematode *Pristionchus pacificus* was reported recently, it became was clear that cellulases were not restricted to plant parasitic nematodes; their presence in this species indicated preadaptation for parasitism of animals [52], consistent with the intermediate evolutionary position of *Pristionchus* between the microbivorous *C. elegans* and the animal parasitic nematodes. In like fashion to evolution of parasitism among nematodes, we can predict that additional analyses of parasitic and free-living flatworm genomes will provide deeper insights into how and when parasitism evolved in the phylum Platyhelminthes, particularly in comparison to the fresh-water planarian *S. mediterranea*, a non-parasitic flatworm for which a draft genome is available [53]. In addition to evolution of parasitism of humans and other vertebrate hosts, helminth parasite genome sequences will also facilitate evolutionary studies on the role of intermediate hosts/vectors such as the snail in schistosome infections and the mosquito in filarial infections in this evolution.

Host–Parasite Relationships

Investigations of regulatory networks involved in the embryonic development, organogenesis, development, and reproduction of helminths based on newly available genome sequences have revealed the presence of well-characterized signaling pathways, including those for Wnt, Notch, Hedgehog, and transforming growth factor β (TGF-β). These pathways can be recognized in the *B. malayi* and schistosome genomes [22,24,25]. These include endogenous hormones including epidermal growth factor (EGF)-like and fibroblast growth factors (FGF)-like peptides. Predicted components of the Ras–Raf–MAPK and TGF-β-SMAD signaling pathways (including FGF and EGF receptors), for example, encoded by these genomes, have components sharing high sequence identity with their mammalian orthologs, implying that schistosomes or filarial worms, in addition to utilizing their own pathways, might exploit host growth factors as developmental signals.

Immune regulation by helminth parasites includes suppression, diversion, and alteration of the host immune response, resulting in an anti-inflammatory environment that is favorable to parasite survival. For example, chronic infections induce key changes in host immune cell populations including dominance of the T-helper 2 (Th2) cells and selective loss of effector T cell activity, against a background of regulatory T cells, alternatively activated macrophages, and TH2-inducing dendritic cells [54,55]. With advances in genomics, numerous parasite-derived proteins, including cytokine homologs, protease inhibitors, and an intriguing set of novel products, as well as glycoconjugates and small lipid moieties, have been discovered with known or hypothesized roles in immune interference [56–61]. These studies suggest that secreted parasite products interfere with different arms of the immune system by influencing the cytokine network and signal transduction pathways or by inhibiting essential enzymes. Using bioinformatics to compare the predicted proteome of *B. malayi* to proteins implicated in the immune response (interleukins, chemokines, and other signaling molecules), potential immune modulators produced by the filarial have been identified, including genes encoding the macrophage migration inhibition family of signaling proteins [62]. Furthermore, the genome of the blood fluke *S. mansoni* encodes a large array of paralogues of fucosyl and xylosyltransferases [25] that are involved in generating novel glycans at the host–parasite interface and could have an important role in the subverting the host immune system. A recent comprehensive review summarizes our current understanding of the growing number of individual helminth mediators that target key receptors or pathways in the mammalian immune system [63].

Helminth infection can have a broad impact on the entire immune system. Infection with trematode and nematode parasites, for example, correlates with a reduced incidence of atopic, allergic-type disorders [64]. Thus, helminth infection might potentially be useful as a novel therapy for allergic or autoimmune diseases [65]. Recently, worms, eggs, or purified nematode parasite protein have been used in preclinical and clinical trials to protect humans from allergy and autoimmunity (reviewed in [66–70]), including Crohn’s disease and ulcerative colitis [71,72].
Other studies have shown that substances produced by helminths, for example *Ascaris suum*, *Nippostrongylus brasiliensis*, and *Acanthocheilonema viteae*, can directly interfere with allergic responses or with development of allergen-specific Th2 responses [73–75]. ES-62, a molecule secreted by the filarial nematode *A. viteae*, directly inhibits the FceRI-induced release of mediators from mast cells, protects against mast-cell–dependent hypersensitivity in skin and lungs [76] and inhibits collagen-induced arthritis [77]. Research is underway to develop molecules that mimic the activity of ES-62 as drugs for allergic and autoimmune diseases [66]. Other helminth-derived products have the potential to reduce allergic responses. These products include schistosomal lysophosphatidylserine (lyso-PS) [61] and thioredoxin peroxidase from the liver fluke *Fasciola hepatica* [78]. These findings demonstrated that helminths produce products that can interfere with both the development of allergic responses and the workings of host effector mechanisms.

**The “Dependent” Helminth**

As a consequence of evolution of an obligatory parasitic existence, helminth parasites are dependent upon their immediate and definitive hosts for many necessities including nutrients such as amino acids; filariae are dependent on insect vectors to transport them to the host. The newly available genome sequences for schistosomes and *B. malayi* have confirmed earlier biochemical studies that had revealed aspects of physiological/biochemical dependence of these parasites on the host. For example, schistosomes cannot synthesize fatty acids de novo, or sterols, purines, and nine human essential amino acids plus arginine or tyrosine, and must catabolize complex precursors obtained from their hosts. Loss or degeneracy of fatty acid, sterol, and purine synthesis pathways in schistosomes likely relates to the adoption of a parasitic lifestyle; it is notable that genes encoding all the key enzymes for both the de novo fatty acid and purine syntheses are complete in the (free-living) planarian *S. mediterranea*. To obtain essential lipid nutrients, the schistosome genome encodes transporters, including apolipoproteins, low-density lipoprotein receptor, scavenger receptor, fatty-acid-binding protein, ATP-binding-cassette transporters and cholesterol esterase, to exploit fatty acids and cholesterol from host blood [25,79].

Many species of filarial nematodes are themselves infected by the endosymbiotic bacterium *Wolbachia*. The genome sequence of the *Wolbachia* species that infects the roundworm nematode *B. malayi* [80] helped establish which metabolites the bacterium potentially provides to the nematode (riboflavin, flavine adenine dinucleotide, heme, and nucleotides, for example) and which are provided by the nematode to the endobacterium (notably, amino acids). This type of information has opened up the exciting possibility that drugs already registered for human use might inhibit key biochemical pathways in *Wolbachia* that could sterilize or kill the adult worms. Although the *Wolbachia* genome is even more degenerate than that of the related pathogen *Rickettsia*, it has retained more intact metabolic pathways than *Rickettsia*. This may be important in its biochemical contribution to host (i.e., filarial) viability and fecundity.

The wBm genome encodes many more proteases and peptidases than *Rickettsia*, which likely degrade host proteins in the extracellular environment. Other proteins encoded by wBm include a common type IV secretion system, as used by some pathogenic gram-negative bacteria to transfer plasmids and proteins into surrounding host cells, and an abundance of ankyrin domain-containing proteins, which might regulate host gene expression, as suggested for *Ehrlichia phagocytophila* AnkA [81], as well as several proteins predicted to localize on the cell surface. Ankyrin domain–containing proteins are noteworthy because of their roles in protein–protein interactions in a variety of cellular processes. A number of other wBm molecules are of interest as potential drug targets. For example, glutathione biosynthesis genes may provide glutathione for the protection of the filaria from oxidative stress or human immunological effector molecules. Heme produced from wBm (all five synthesis genes are present) could be vital to worm embryogenesis, as there is evidence that heme is produced and reproduction are controlled by ecdysteroid-like hormones [92], synthesis of which requires heme. Depletion of *Wolbachia* might therefore halt production of these hormones and block molting and/or embryogenesis in *B. malayi*. Most, if not all, nematodes, including *B. malayi*, appear to be unable to synthesize heme, but must obtain it from extraneous sources, such as the host, the food supply, or perhaps from endosymbionts.

**Challenges for the Future**

The filarial and schistosome genome sequences now available provide the vanguard for assembly of a genome sequence catalog of the numerous other neglected helminth parasites (Table 1). Comparative genomics will likely be a dominant approach to organize, interpret, and utilize the vast amounts of genomic information anticipated from the genomes of these parasites (e.g. [83,84]). In terms of sequencing tools, the new generation of “massively parallel” sequencing platforms commercially available today, (such as the Roche/454 pyrosequencer [85], Illumina/Solexa [86], and SOLID [87]) offer of the order of 100- to 1,000-fold increases in throughput over the Sanger sequencing technology [88] on capillary electrophoresis instruments. This rapid change to producing millions of DNA sequence reads in a short time will have a huge impact on research on NTDs. Each platform has a specific application: while the Roche/454 is optimal for in-depth analysis of whole transcriptomes and de novo sequencing of bacterial and small eukaryotic genomes, the Illumina and SOLID systems are highly attractive for resequencing projects aimed at identifying genetic variants (mutations, insertions, deletions), profiling and discovering noncoding RNAs (ncRNAs), and studying epigenetic modifications of histones and DNA. With the increased read length and improved error rate of massively parallel pyrosequencing technology, de novo sequencing of helminth genomes has become possible at a fraction of earlier costs. In the next five years, projects at the Washington University’s Genome Center (http://www.genome.gov/10002154) and the Wellcome Trust Sanger Institute (http://www.sanger.ac.uk/Projects/Helminths/) should increase the available sequence data on human helminths and their close relatives by an order of magnitude, adding more than 20 draft genomes to those listed in Table 1.

Once these reference genomes become available, sequencing of clinical isolates is expected to accelerate. Sequencing of the clinical strains and strain-to-reference comparisons can be performed using platforms such as Illumina/Solexa and SOLID to investigate genome-wide polymorphism and provide a comprehensive picture of natural helminth genome variation. These approaches should also be valuable for exploring genetic changes involved in resistance to anti-worm drugs and understanding the potential mechanisms of drug resistance in human parasites, and can be expected to facilitate development of genetic markers to monitor and manage any future appearance and spread of drug resistance. These phenomena are of tremendous importance, particularly since some major neglected helminth diseases are being targeted in mass drug treatment campaigns [89]. In addition, the new generation of sequencing technologies has also provided unpre-
edented opportunities for high-throughput functional genomic research (reviewed in [90]) that awaits application to helminth research.

Although some details of immunomodulation by helminth components have been characterized, we are just beginning to understand how these parasite products act on immune responses and to assemble fragmentary information on individual components into a comprehensive picture. Comparisons of helminth molecules with orthologues/paralogues from free-living relatives will strengthen efforts to decipher the strategies adopted by helminth parasites to evade and subvert their host immune responses. This information will be exploitable for development of drugs and vaccines against the parasites and potentially also novel therapeutic biologics for use in humans. Future studies might determine whether helminth proteins with unknown function might be the source for the intriguing regulatory effects helminth infections have on the host immune response.

Treatment for helminthic infections, responsible for hundreds of thousands of deaths each year, depends almost exclusively on just two or three drugs: praziquantel, the benzimidazoles, and ivermectin. Vaccines and new drugs are needed, certainly because drug resistance in human helminth parasites such as schistosomes, whipworms, and filariae, to these compounds would present a major problem for current treatment and control strategies. Pharmacogenomics with the new helminth genomes represents a practicable route forward toward new drugs. For example, chemogenomics screening of the genome sequence of S. mansoni identified >20 parasite proteins for which potential drugs are available approved for other human ailments [23], and indeed for which, in the case of the schistosome thioether oxidase glutathione reductase, auranofin (an anti-arthritis medication) was shown recently to exhibit potent anti-schistosomal activity [91]. Finally, the vast new sequence information would undoubtedly allow revision of our understanding of the host-parasite relationship, its evolution, vector–pathogen and helminth–symbiont interactions, unique aspects of cell biology and biochemistry, phylogenetic relationships, intervention targets, research approaches (e.g. [92]), and so forth.

Acknowledgments

We thank Victoria Mann, Geoffrey Gobert and Gabriel Rinaldi for access to their unpublished findings on schistosomes transduced with pseudotyped virions.


