Published in final edited form as:

Parasite Immunol. 2012 ; 34(2-3): 121–129. doi:10.1111/j.1365-3024.2011.01344.x.

Filarial and *Wolbachia* genomics

A. L. SCOTT¹, E. GHEDIN², T. B. NUTMAN³, L. A. McREYNOLDS⁴, C. B. POOLE⁴, B. E. SLATKO⁴, and J. M. FOSTER⁴

¹Department of Molecular Microbiology and Immunology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MA, USA

²Department of Computational & Systems Biology, Center for Vaccine Research, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

³Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, MD, USA

⁴New England Biolabs, Ipswich, MA, USA

SUMMARY

Filarial nematode parasites, the causative agents for a spectrum of acute and chronic diseases including lymphatic filariasis and river blindness, threaten the well-being and livelihood of hundreds of millions of people in the developing regions of the world. The 2007 publication on a draft assembly of the 95-Mb genome of the human filarial parasite *Brugia malayi* – representing the first helminth parasite genome to be sequenced – has been followed in rapid succession by projects that have resulted in the genome sequencing of six additional filarial species, seven nonfilarial nematode parasites of animals and nearly 30 plant parasitic and free-living species. Parallel to the genomic sequencing, transcriptomic and proteomic projects have facilitated genome annotation, expanded our understanding of stage-associated gene expression and provided a first look at the role of epigenetic regulation of filarial genomes through microRNAs. The expansion in filarial genomics will also provide a significant enrichment in our knowledge of the diversity and variability in the genomes of the endosymbiotic bacterium *Wolbachia* leading to a better understanding of the genetic principles that govern filarial–*Wolbachia* mutualism. The goal here is to provide an overview of the trends and advances in filarial and *Wolbachia* genomics.

Keywords

Brugia; filarial; genome; Nematode; noncoding RNA; proteomics; transcriptomics; *Wolbachia*

INTRODUCTION

Filarial nematode parasites are vector-borne pathogens of substantial medical and veterinary importance. While it is difficult to estimate the current prevalence of filarial infections in humans owing to ongoing eradication efforts (1), in 2006 it was estimated that well over 1 billion were at risk and approximately 150 million people were infected with the major species of filarial nematodes: *Wuchereria bancrofti*, *Brugia malayi*, *Onchocerca volvulus* and *Loa loa* (2). In humans, filarial parasites typically result in persistent infections that

cause some of the most debilitating disease states recorded: lymphatic blockage resulting in elephantiasis, ocular pathology leading to river blindness and severe dermatitis manifesting as sowda. While current drugs are effective (predominantly against the larval forms of the parasite), development of resistant strains has recently been reported (3). There is no vaccine to protect against filarial infection or to modulate disease.

There are eight filarial species of medical importance for humans and a number of additional species of veterinary and scientific interest (Table 1). All of the filarial parasites have common elements to their life cycle; most notably, all species are diecious, undergo four larval moults, are transmitted by hematophagous arthropod vectors and maintain developmental pauses at the first and third larval stages. The developmental pauses are critical for successful transmission. The nature and severity of pathology varies across filarial species with most of the tissue pathology and immunological changes associated with the adult parasites and the transmissible larval stage.

The Phylum Nematoda contains five clades (4) and an estimated 1 million species (5). Filarial species reside in clade III (Spiruria), which is comprised of animal parasites including ascaridid, spirurid and oxyurid parasites of vertebrates and oxyurid and rhigonematid parasites of arthropods (6). Since 1998, when it became the first multicellular organism to be fully sequenced (7), the terrestrial, bacteriovorus, clade V species *Caenorhabditis elegans* has dominated as a prototype system for advancing understanding in the molecular genetics and developmental biology of nematodes. It also represents a powerful model that continues to contribute to a variety of important fields including medical genetics, ageing, cancer and infectious diseases [reviewed in (8–10)]. Early on, it was presumed, given the common body plan and developmental progression of all nematode species, that the structure, organization, gene content and regulation of the *C. elegans* genome could serve as an apt general model for nematodes. With the publication of the genome from the first parasitic nematode, *B. malayi*, with its unique genome structure, large number of unique genes and the discovery that many filarial species harbour an endosymbiont (11), the idea was sown that no single nematode genome could serve as a general model. This idea has been supported by subsequent reports on the composition and structure of other filarial and nonfilarial nematode genomes (12–14). It appears that diversity in nematode genome structure and content parallels the extensive diversity found in the Nematoda itself.

One of the notable advances to emerge from the *B. malayi* sequencing project was an appreciation that studies on the genomics and biology of many filarial species are incomplete without a consideration of the genome of their endosymbiotic bacterium *Wolbachia* (Table 1). The current idea that *Wolbachia* and filarial parasites have evolved a mutualistic relationship in which the bacterium exerts control over the development (moulting), fecundity and viability of its host (15,16) underscores the importance of integrated ensemble studies of these two organisms.

The advances associated with next-generation deep sequencing have significantly reduced the costs of sequencing, assembling and annotating genomes. These technical and informatics advances have already made an impact on the field of nematode genomics through increasing the corpus of nematode (including filarial) sequenced genomes and by providing support data that refine and expand assembly and annotation. The ability to produce sequence free of the traditional bias associated with conventional sequencing library construction has facilitated the filling of gaps and the generation of higher-order assemblies. The enhanced transcriptomic capabilities conferred by next-generation sequencing approaches are providing unprecedented insights into the full spectrum of stage-, sex- and tissue-associated gene expression that is refining and expanding the complement of genes

for each species and enhancing annotation through the identification of new exons, spliced variants and gene boundaries. In addition, the next-generation sequencing platforms have facilitated the initial studies in defining gene regulation in filarial species through the identification of micro RNAs. The goal of this review is to provide an outline of the advances that are taking place in filarial nematode and *Wolbachia* genomics and how these advances are expanding and refining our understanding of the structure and composition of filarial genomes, the regulation of gene expression and the molecular interactions that define endosymbiosis between filarial parasites and *Wolbachia*.

THE *BRUGIA MALAYI* GENOME

The reported draft of the *B. malayi* 95-Mb nuclear genome was based on a traditional shotgun sequencing approach that accounted for an estimated 90% of the genome at approximately nine-fold coverage (11). This resulted in an assembly distributed on >8000 scaffolds that ranged in size from <2 kb to more than 6.5 Mb. It was estimated that 14% of the genome was made up of repeated DNA comprised mainly of the major A+T-rich 322-bp *HhaI* repeat family, the 62/53-bp *MboI* repeat family and large number of simple and low complexity repeats. The number and distribution of the tandemly arrayed repeat elements presented a significant challenge to the full assembly of the *B. malayi* genome into chromosomal units (1N = 5 autosomal and 1X).

Deep sequencing approaches have recently been applied to extend and enhance the resolution of the *B. malayi* genome and to resolve some of the content and assembly issues encountered in the conventional shotgun-based sequencing. While the analysis of these data is in progress, it is clear that the combined short read (Illumina) and long read (pyrosequencing, 454 Life Sciences; Roche, Branford, CT, USA) strategy will have a significant impact on the resolution of the *B. malayi* genome. The combined Illumina/454 sequence data have increased the coverage to 20X resulting in the annotation of an additional approximately 6 Mb into scaffolds. In addition, the new sequencing efforts have resulted in a significant increase in the N50 of the scaffolds from 93 to 357 kb and decreased the number of scaffolds from 8236 to under 2900 (this work is in progress). The expectation is that this next build of the *B. malayi* genome will also contain enhanced annotation of genes and regulatory elements provided by recently completed transcriptomic and proteomic studies (17–20).

TRANSCRIPTOMICS

Analyses of the gene complements of the three free-living (*C. elegans*, *Caenorhabditis briggsae* and *Pristionchus pacificus*) and the four parasitic (*B. malayi*, *Trichinella spiralis*, *Meloidogyne incognita* and *Meloidogyne hapla*) published nematode genomes indicate that as many as 45% of the predicted protein-coding genes for each organism represent novel, species-specific entries into the major databases (7,12–14,21,22). Given the challenges that this level of protein sequence diversity space presents for informatics-based gene finding and annotation of nematode genomes, transcriptomic and proteomic approaches play especially important roles in defining and verifying gene content. Indeed, the >25 000 expressed sequence tags (EST) derived from over a dozen larval and adult cDNA libraries (representing over approximately 10 000 clusters; Table 1) (23) provided critical support data during the annotation of the *B. malayi* genome (11). Although EST databases have proven to be a rich source for gene discovery and genome annotation, the EST approach is relatively inefficient and fraught with cloning bias and limitations associated with single-pass sequence data and dealing with partial sequences. The recent ability to apply massively parallel sequencing strategies provides a more efficient and ostensibly less biased approach to define full-length transcripts and detail gene expression. The value of this complementary

approach is exemplified by the modENCODE project for *C. elegans*, where RNAseq and proteomics were employed along with other methods to identify 1650 heretofore unrecognized protein-coding genes, redefine the structure of known genes, reveal the scope of alternative splicing and increase the estimate of diversity of noncoding RNAs 20-fold (24).

The initial use of this RNAseq approach for filarial transcriptomics is a project in which the Illumina deep sequencing platform was employed to produce over 10^8 paired end reads to gain a high-resolution expression profile of *B. malayi* eggs/embryos, immature (3 days of age) and mature microfilariae, third- and fourth-stage larvae and adult male and female worms (in annotation; S. Michalski and B. Christensen, personal communication). It is anticipated that this analysis will, in addition to expanding on the 11 508 predicted genes (11), provide enhanced detail on the genomic organization of previously identified *B. malayi* genes including mispredicted and unpredicted exons, the nature of splice variants and 5' and 3' UTR sequences.

PROTEOMICS

The identification of a complement of 11 508 protein-coding genes in the initial draft of *B. malayi* genome has provided the baseline data needed to apply high-throughput proteomic approaches to augment genome annotation including identification of translational start and stop sites, frame shifts and the verification of the translation of those gene models designated as hypothetical/predicted. Bennuru *et al.* (18) have recently published a high-density proteome map of *B. malayi* using reverse-phase liquid chromatography–tandem mass spectroscopy to define the excretory/secretory (ES) and somatic proteins (17) produced by the adults, microfilariae and infective stage larvae. Taken together, these analyses identified 7103 (>60%) of the proteins predicted in the genome including 2336 of the 4956 genes that had been assigned hypothetical/predicted status (11). A noteworthy observation from the analyses of the ES and somatic extracts is that a significant percentage (approximately 30%) of the proteins are synthesized in a developmentally regulated manner. While many of these stage-associated proteins are unique to the databases and have no discernable functional motifs, this regulated pattern of translation implicates these molecules in processes important for filarial survival. As other filarial species attain the level of annotation that will allow the application of similar proteomics approaches, it will be fascinating to determine whether this developmental regulation of protein production holds true and whether integrated comparative genomic/transcriptomic/proteomic analysis will provide the power to assign putative roles to the many hypothetical/predicted filarial genes.

ADDITIONAL FILARIAL GENOMES

As noted above, next-generation sequencing technologies have made it possible to rapidly increase the number of nematode genomes available for analysis of structure and content. Recently, sequence data for the genomes of three major human filarial pathogens, *W. bancrofti*, *O. volvulus* and *L. loa*, have been made publically available (http://www.broadinstitute.org/annotation/genome/filarial_worms/MultiHome.html). In addition, draft genome sequence will soon be available for a number of filarial pathogens of animals including *Dirofilaria immitis*, *Litomosoides sigmodontis* and *O. ochengi* (http://www.nematodes.org/nematodegenomes/index.php/Species_in_sequencing).

In the very near future, this rich substrate of filarial genomes will enable an unprecedented opportunity to identify the biochemical and regulatory pathways that are common and unique to filarial species and thus generate rational models for filarial physiology that point to targets for drug- or small molecule-based control measures. Comparative analysis of closely related species pairs such as *B. malayi* and *W. bancrofti* and *O. volvulus* and *O.*

ochengi will provide data for better models of filarial evolution and the adaptations that resulted in certain species being restricted to a single vertebrate host. It will be of great interest to define the degree of short- and long-range synteny between the genomes of filarial species. The genomes may also provide evidence of the distinct and common mechanisms required for nematode survival and transmission by mosquitoes and flies.

SMALL NONCODING RNAS

The classical view of RNAs [tRNA, mRNA and rRNA] as molecules restricted to protein synthesis has been significantly modified over the last decade to include small noncoding RNAs that function as vital regulatory molecules. Nematodes have played a key role in seminal observations in this arena. The first 22-nucleotide microRNA (miRNA) was discovered by Ambros and Ruvkun in a search for the cause of a developmental timing defect in *C. elegans*. They discovered that a mutation in a miRNA, *lin-4*, altered its ability to bind to the 3' untranslated region and inhibit the expression of LIN-14, a heterochronic regulatory nuclear protein (25,26). The discovery of RNA interference (RNAi) in *C. elegans* (27) provided the basis for defining the mechanisms that govern the generation and action of miRNAs. From these initial nematode-derived observations, it has been demonstrated in eukaryotes from sponges to mammals that, in addition to the regulation of gene expression, small noncoding RNAs such as miRNAs (microRNAs), siRNAs (small interfering RNAs) and piRNAs (Piwi-interacting RNAs) are critical for the control of cellular metabolism, growth and differentiation, to maintain genome integrity and as a defence mechanism against viruses and mobile genetic elements (28,29).

Recently, noncoding RNAs have emerged as a major component of the eukaryotic transcriptome. Transcription of miRNAs results in the formation of a dsRNA hairpin that is sequentially cleaved by the double-stranded (ds) RNA-specific endoribonucleases Drosha and Dicer to generate a 22-nt dsRNA. In association with proteins of the Argonaute family, the dsRNA is unwound and directed to its cognate RNA and/or DNA molecules in the context of the RNA-induced silencing complex (RISC) (30). Part of the miRNA binding specificity for target mRNAs is determined by a 6- to 7-nt 'seed' sequence at the 5' end of each miRNA (31). The *B. malayi* genome contains many of the components required for miRNA processing including orthologues of Drosha, Dicer and Argonaute (11). Recently, deep sequencing has been employed to explore the diversity of small RNAs expressed by mature male, female and microfilariae of *B. malayi* (32).

Employing approximately 30 million deep sequencing reads, 145 miRNAs were identified in the adults and larvae of *B. malayi*. The miRNAs represent 99 families each defined by a unique seed sequence. Sixty of the miRNA families found in *B. malayi* are conserved in other species. Examples of conserved miRNAs include the let-7, miR-1 and bantam families (32, C. B. Poole, W. Gu, P. Davis, S. Kumar, J. Jin, D. Conte, Mello C. & L. A. McReynolds, unpublished data). These highly conserved miRNAs are likely involved in the regulation and maintenance of conserved developmental rather than species-specific functions.

Only 11/99 *B. malayi* miRNA families are restricted to helminths. Of this group, nine are found only in filarial parasites, including *L. loa*, *W. bancrofti* and *O. volvulus*. The sequences of the filarial-specific miRNAs are quite similar in all the filarial species examined, suggesting that they have recently evolved by gene-duplication and/or single base substitutions from existing miRNA families (C. B. Poole, W. Gu, P. Davis, S. Kumar, J. Jin, D. Conte, Mello C. & L. A. McReynolds, unpublished data). Additional studies are needed to identify the mRNA targets of the *B. malayi* miRNAs and to determine whether they are involved in the maintenance of parasitism.

About 30% of the *B. malayi* miRNAs are preferentially expressed in one stage over another. For example, miR-71 and miR-34 are 15-times more highly expressed in microfilariae compared to adult male and female parasites. Together, miR-71 [27%] and miR-34 [13%] represent 40% of the total miRNAs found in microfilariae. A similar differential increase in miR-71 and miR-34 has been noted in *C. elegans* diapause L1s compared to embryos (33). The prolonged pause in development of blood-stage microfilariae is thought to be analogous to diapause in *C. elegans*. The role of miR-71 and miR-34 extends beyond the control of developmental arrest. For both *C. elegans* adults and mammalian cells, miR-34 enhances radiation resistance (34). In *C. elegans*, the loss of miR-71 results in a two-fold reduction in life span possibly through blocking members of the DAF-2, insulin-like growth factor pathway (35). It will be of interest to determine whether miR-71 and miR-34 govern stress resistance and maintenance of arrested development in filarial nematodes.

Inspection of the *B. malayi* miRNA families reveals several examples of miRNA evolution. Current models propose that miRNAs evolve by a combination of gene amplification, mutation and arm switching mechanisms (36). The *B. malayi* miR-2 family is an example of gene amplification. This family has 11 members and has expanded compared to the five-member *C. elegans* miR-2 family. Six of the *B. malayi* miR-2 family members are more closely related to each other than to *C. elegans* miR-2, suggesting a recent expansion of this family in filarial parasites. Several of the filarial-specific miRs likely arose by mutation from conserved miR families. For example, alignment of *B. malayi* miR-57 with filarial-specific miR-A identified a single mutation in the seed sequence of miR-A compared to miR-57. With the overall identity of the two miRNAs at 67%, these two miRNAs likely arose from a common ancestor by duplication of miR-57 followed by nt substitutions to create the two different miRNAs (C. B. Poole, W. Gu, P. Davis, S. Kumar, J. Jin, D. Conte, Mello C. & L. A. McReynolds, unpublished data).

As noted above, the repeat content of the *B. malayi* genome is approximately 14%. The dominant repeat, the 322-bp *HhaI* DNA repeat family (37), encodes a small RNA. The *HhaI* RNA is present in both the adult and mf stages and is transcribed from one strand of the DNA. A computational search did not find any sequences that could form a stable dsRNA hairpin, suggesting that this RNA is not generated by Drosha or Dicer cleavage. The function of this small noncoding RNA is not known.

WOLBACHIA

Most species of filarial nematodes harbour within their tissues an α -proteobacterial endosymbiont belonging to the genus *Wolbachia*. These Rickettsia-like bacteria are present in all filarial species that infect humans with the exception of *L. loa* and *M. streptocerca*. The maternally transmitted endosymbionts are found predominantly in the hypodermal cells of the lateral cord as well as the ovaries, oocytes and developing embryos within the uteri of adult female worms [see (38) for review]. *Wolbachia* were first described in insect hosts and are widespread in arthropods with close to 70% of species believed infected (39). In arthropod hosts, the endobacteria are considered reproductive parasites because they induce a variety of phenotypes such as cytoplasmic incompatibility, parthenogenesis, feminization of males and the killing of male embryos, which serve to promote the fitness of infected females and the spread of endosymbiont through populations (40). Furthermore, *Wolbachia* can be cleared from their arthropod hosts by antibiotic treatment with little or no consequence (41).

In contrast, the *Wolbachia* infecting filarial nematodes display the features of an obligate mutualist – neither organism can survive long term without the other. Attempts to culture the *Wolbachia* endosymbiont from *B. malayi* (wBm) or to adoptively transfer wBm to an alternate host species have not been successful. Treatment with tetracycline family

antibiotics clear the *Wolbachia* from filarial parasites, and this clearance precipitates a block in embryogenesis followed eventually by the death of the worm. Similar treatment of *Wolbachia*-free filarial species shows no such effects on worm viability. Clinical trials with tetracyclines have validated the *Wolbachia* endosymbionts as an anti-filarial target, thus demonstrating a novel effective treatment, which, importantly, has both microfilaricidal and macrofilaricidal activity. However, the required long courses of tetracycline treatment and the contraindications for young children and pregnant women preclude its widespread implementation.

The genome sequence of the *Wolbachia* endosymbiont from *B. malayi* (*wBm*) was determined to facilitate understanding of the symbiosis between the bacterium and its filarial host and to assist in the identification of *Wolbachia* biochemical processes that could serve as novel antifilarial drug targets. Typical of most endosymbionts, the 1.08-Mb-circular *wBm* genome shows loss of many metabolic processes and is predicted to encode only 806 protein-coding genes. Mapping the genes to metabolic pathways indicated that the biosynthesis of riboflavin, flavin adenine dinucleotide (FAD), haeme and nucleotides might be important for the symbiotic relationship with the *B. malayi* host (42). These biosynthetic capabilities appear not to be encoded by the *B. malayi* genome, although nucleotide salvage pathways exist (11). Conversely, nematode-derived amino acids appear important for *wBm* growth (42). Interestingly, among the few *wBm* genes showing positive selection in a genome-wide screen for the presence of diversifying selection were several genes involved in biosynthesis of haeme, riboflavin and nucleotide biosynthesis (43). This observation lends weight to the notion that these pathways may be key components of the mutualistic association.

Additional bioinformatic-based discovery and experimentation have predicted the essential gene repertoire of *wBm* and identified several *Wolbachia* biochemical processes that represent candidate drug targets. These include enzymes of lipid II biosynthesis, lipoprotein biosynthesis, haeme biosynthesis, and the glycolytic enzymes pyruvate phosphate dikinase and cofactor-independent phosphoglycerate mutase. (44–49). Several of these are undergoing further evaluation as potential targets or are already subject to high-throughput inhibitor screens.

A recent proteomic analysis of extracts from whole adult and larval *B. malayi* also identified 577 of the 806 *wBm* proteins (17). This analysis demonstrated that many of the *wBm* proteins were differentially transcribed in the various parasite stages. In addition, 96 of the 166 hypothetical/predicted genes (42) were validated as producing a protein during one or more stages of filarial development. Proteomic-level identification of all of the proteins required for nucleotide and haeme biosynthesis in *wBm* supports the hypothesis that the endosymbiont provides critical factors required for parasite survival.

If *Wolbachia* provides metabolites to its filarial hosts as part of the endosymbiotic relationship, an obvious question is how do those *Wolbachia*-free filarial species fulfil their metabolic needs? To address this question, low-coverage genomic sequencing of *O. flexuosa* and *Acanthocheilonema viteae*, filarial species uninfected by *Wolbachia*, was undertaken (50). In both species, fragments of DNA originating from *Wolbachia* were evident in the nuclear genomes, indicating an ancestral *Wolbachia* infection followed by secondary loss of endosymbiont. Although some fragments were transcribed, none appeared capable of encoding a full-length functional protein. The laterally transferred *Wolbachia* DNA was present as short degenerated gene fragments that typically contained frame shifts and stop codons. Interestingly, this low-coverage sequence data also indicated that the *Wolbachia*-free filarial species, like *B. malayi*, were deficient in the biosynthesis of haeme, nucleotides, riboflavin and FAD [J. M. Foster and B. E. Slatko, unpublished data]. The absence of

nematode- or *Wolbachia*-derived genes encoding these biosynthetic capabilities raises questions regarding how the *Wolbachia*-free species fulfil these functions and the relevance of these *Wolbachia*-derived metabolites in species that do harbour *Wolbachia*. It is possible that the low-coverage genomic sequencing of *A. viteae* and *O. flexuosa* simply missed functional genome- or *Wolbachia*-encoded versions of these molecules. An alternative explanation is that filarial nematodes can meet their requirements for these molecules from exogenous sources and that metabolic provisioning by *Wolbachia* in infected filarial species is perhaps simply beneficial under steady state and only essential either at critical points in the filarial life cycle or under certain stressful conditions. Indeed, although *Wolbachia* within arthropods are typically considered ‘parasites’, there is growing evidence that they can also confer benefits to their hosts – including metabolic provisioning of molecules such as haeme and vitamin B (e.g. riboflavin) – under certain conditions (51,52).

While *Wolbachia*-infected filarial species may have become dependent on their endosymbiont for metabolites over evolutionary time, uninfected filarial species may have evolved to be more proficient at securing essential components from their surroundings. Despite uncertainties about the contribution of metabolic provisioning to the mutualistic association, the *Wolbachia* within filarial nematodes are clearly essential as demonstrated by the outcome of antibiotic treatment [see (38) for review].

Sequencing of *Wolbachia* genomes from additional filarial hosts together with more nematode genomes from both *Wolbachia*-infected and uninfected filarial species will facilitate a greater understanding of the nature of the symbiosis. Phylogenetic analyses resolve *Wolbachia* into about seven distinct supergroups and a number of additional lineages (53). *Wolbachia* from filarial nematodes are found almost exclusively in Supergroups C and D, while the other supergroups are comprised of arthropod *Wolbachia*. Supergroup F is an interesting case because it contains *Wolbachia* from both arthropods and filariae (*Mansonella perstans*) and probably reflects a relatively recent transfer of *Wolbachia* between these invertebrate phyla (54). Apart from *wBm* (supergroup D), only three other completed circular *Wolbachia* genomes exist – those from *Drosophila melanogaster* (*wMel*) and *Drosophila simulans* (*wRi*) from supergroup A and *Culex quinquefasciatus* (*wPip*) from supergroup B (55–57). The *Wolbachia* genomes from these different supergroups are highly scrambled relative to each other and show little synteny (42,55). A few areas of short-range synteny exist and probably reflect functionally related gene products such as those comprising the Type IV secretion system. Comparison of the two supergroup A genomes also shows a high degree of mosaicism caused by extensive recombination facilitated by the insertion elements and repeat blocks that populate arthropod *Wolbachia* genomes, but in this case there is considerably more long-range synteny (56).

A number of additional *Wolbachia* genomes have been sequenced to high coverage by next-generation sequencing technologies, but they remain incomplete or are undergoing finishing. These include *Wolbachia* genomes from the filarial nematodes *O. volvulus*, *O. ochengi*, *D. immitis* (all supergroup C), *W. bancrofti* and *L. sigmodontis* (both supergroup D). [see http://www.broadinstitute.org/annotation/genome/filarial_worms/GenomesIndex.html and <http://xyala.cap.ed.ac.uk/downloads/959nematodegenomes/blast/blast.php> for details].

Analysis of a collection of bacterial artificial chromosomes prepared from genomic DNA from *D. immitis* identified a subset that was derived from the *Wolbachia* endosymbiont (*wDi*; supergroup C) [J. M. Foster and B. E. Slatko, unpublished data]. DNA sequences were determined for both ends of these *Wolbachia* inserts (typically about 60 kb long) and mapped onto both the *wBm* genome (supergroup D) and, where possible, the *wOvo* (supergroup C) contigs that represent about 10% of the *Wolbachia* genome from *O. volvulus* (available at NCBI). No syntenic arrangement of adjacent genes from one end of a *wDi*

BAC or genes from opposite ends of a *w*Di BAC (about 60 kb apart) were found when mapped to *w*Bm (supergroup D). However, examples of synteny between *w*Di and *w*Ovo (both supergroup C) were apparent. This analysis predicts that future comparisons of new *Wolbachia* genome sequences from various filarial nematode hosts will reveal greatest structural conservation when members of the same *Wolbachia* supergroup are compared, similar to the situation observed for *Wolbachia* genomes from arthropod hosts. The new insight into the symbiotic relationship between *Wolbachia* and their filarial nematode hosts that might be afforded by the completion of the ongoing *Wolbachia* genome projects and subsequent comparative genomics is eagerly awaited.

WHERE DO WE GO FROM HERE?

Filarial nematodes impose an appalling toll on human development and well-being. While *C. elegans* will maintain its importance as a model for understanding certain aspects of nematode biology, the genome-level data from the increasing number of filarial and nonfilarial nematode species make it clear that the *C. elegans* genome cannot be used as a prototype for gene content or genome structure in Nematoda. It will soon be possible to employ large-scale comparative genomics within Nematoda to aid in identifying the core genetic elements that define a nematode. In addition, detailed genomics combined with transcriptomics, proteomics, metabolomics, biochemistry, physiology, etc. will serve as the bases for a systems biology approach to establish a high resolution understanding the cellular and molecular foundation for filarial nematode development in their vertebrate and arthropod hosts, the requirements to maintain the mutualistic relationship with their endosymbionts and the immunobiology of filarial infections.

Acknowledgments

We thank our colleagues of the Filarial Genome Project and the filarial research community for their continued support and encouragement. We thank Shelly Michalski (University of Wisconsin, Oshkosh) and Bruce Christensen (University of Wisconsin, Madison) for their input on *B. malayi* transcriptomics.

REFERENCES

1. Chu BK, Hooper PJ, Bradley MH, McFarland DA, Ottesen EA. The economic benefits resulting from the first 8 years of the Global Programme to Eliminate Lymphatic Filariasis (2000–2007). *PLoS Negl Trop Dis*. 2010; 4:e708. [PubMed: 20532228]
2. Ottesen EA. Lymphatic filariasis: treatment, control and elimination. *Adv Parasitol*. 2006; 61:395–441. [PubMed: 16735170]
3. Osei-Atweneboana MY, Awadzi K, Attah SK, Boakye DA, Gyapong JO, Prichard RK. Phenotypic evidence of emerging ivermectin resistance in *Onchocerca volvulus*. *PLoS Negl Trop Dis*. 2011; 5:e998. [PubMed: 21468315]
4. Blaxter ML, De Ley P, Garey JR, et al. A molecular evolutionary framework for the phylum Nematoda. *Nature*. 1998; 392:71–75. [PubMed: 9510248]
5. Lamshead PJ, Brown CJ, Ferrero TJ, Hawkins LE, Smith CR, Mitchell NJ. Biodiversity of nematode assemblages from the region of the Clarion-Clipperton Fracture Zone, an area of commercial mining interest. *BMC Ecol*. 2003; 3:1. [PubMed: 12519466]
6. Blaxter ML. Nematoda: genes, genomes and the evolution of parasitism. *Adv Parasitol*. 2003; 54:101–195. [PubMed: 14711085]
7. Consortium, C. e. S. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science*. 1998; 282:2012–2018. [PubMed: 9851916]
8. Olsen A, Vantipalli MC, Lithgow GJ. Using *Caenorhabditis elegans* as a model for aging and age-related diseases. *Ann N Y Acad Sci*. 2006; 1067:120–128. [PubMed: 16803977]
9. Poulin G, Nandakumar R, Ahringer J. Genome-wide RNAi screens in *Caenorhabditis elegans*: impact on cancer research. *Oncogene*. 2004; 23:8340–8345. [PubMed: 15517014]

10. Pradel E, Ewbank JJ. Genetic models in pathogenesis. *Annu Rev Genet.* 2004; 38:347–363. [PubMed: 15568980]
11. Ghedin E, Wang S, Spiro D, et al. Draft genome of the filarial nematode parasite *Brugia malayi*. *Science.* 2007; 317:1756–1760. [PubMed: 17885136]
12. Mitreva M, Jasmer DP, Zarlenga DS, et al. The draft genome of the parasitic nematode *Trichinella spiralis*. *Nat Genet.* 2011; 43:228–235. [PubMed: 21336279]
13. Opperman CH, Bird DM, Williamson VM, et al. Sequence and genetic map of *Meloidogyne hapla*: a compact nematode genome for plant parasitism. *Proc Natl Acad Sci U S A.* 2008; 105:14802–14807. [PubMed: 18809916]
14. Abad P, Gouzy J, Aury JM, et al. Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*. *Nat Biotechnol.* 2008; 26:909–915. [PubMed: 18660804]
15. Fenn K, Blaxter M. Are filarial nematode *Wolbachia* obligate mutualist symbionts? *Trends Ecol Evol.* 2004; 19:163–166. [PubMed: 16701248]
16. Taylor MJ, Bandi C, Hoerauf A. *Wolbachia* bacterial endosymbionts of filarial nematodes. *Adv Parasitol.* 2005; 60:245–284. [PubMed: 16230105]
17. Bennuru S, Meng Z, Ribeiro JMC, et al. Stage-specific proteomic expression patterns of the human filarial parasite *Brugia malayi* and its endosymbiont *Wolbachia*. *Proc Natl Acad Sci U S A.* 2011; 108:1234–1235. [PubMed: 21220304]
18. Bennuru S, Semnani R, Meng Z, Ribeiro JM, Veenstra TD, Nutman TB. *Brugia malayi* excreted/secreted proteins at the host/parasite interface: stage- and gender-specific proteomic profiling. *PLoS Negl Trop Dis.* 2009; 3:e410. [PubMed: 19352421]
19. Hewitson JP, Harcus YM, Curwen RS, et al. The secretome of the filarial parasite, *Brugia malayi*: proteomic profile of adult excretory-secretory products. *Mol Biochem Parasitol.* 2008; 160:8–21. [PubMed: 18439691]
20. Moreno Y, Geary TG. Stage- and gender-specific proteomic analysis of *Brugia malayi* excretory-secretory products. *PLoS Negl Trop Dis.* 2008; 2:e326. [PubMed: 18958170]
21. Dieterich C, Clifton SW, Schuster LN, et al. The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism. *Nat Genet.* 2008; 40:1193–1198. [PubMed: 18806794]
22. Stein LD, Bao Z, Blasiar D, et al. The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *PLoS Biol.* 2003; 1:E45. [PubMed: 14624247]
23. Elsworth B, Wasmuth J, Blaxter M. NEMBASE4: the nematode transcriptome resource. *Int J Parasitol.* 2011; 41:881–894. [PubMed: 21550347]
24. Gerstein MB, Lu ZJ, Van Nostrand EL, et al. Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. *Science.* 2010; 330:1775–1787. [PubMed: 21177976]
25. Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell.* 1993; 75:843–854. [PubMed: 8252621]
26. Wightman B, Ha I, Ruvkun G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. *Cell.* 1993; 75:855–862. [PubMed: 8252622]
27. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature.* 1998; 391:806–811. [PubMed: 9486653]
28. Wheeler BM, Heimberg AM, Moy VN, et al. The deep evolution of metazoan microRNAs. *Evol Dev.* 2009; 11:50–68. [PubMed: 19196333]
29. Tabara H, Sarkissian M, Kelly WG, et al. The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. *Cell.* 1999; 99:123–132. [PubMed: 10535731]
30. Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. *Nat Rev Genet.* 2009; 10:94–108. [PubMed: 19148191]
31. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell.* 2005; 120:15–20. [PubMed: 15652477]

32. Poole CB, Davis PJ, Jin J, McReynolds LA. Cloning and bioinformatic identification of small RNAs in the filarial nematode, *Brugia malayi*. *Mol Biochem Parasitol*. 2010; 169:87–94. [PubMed: 19874857]
33. Karp X, Hammell M, Ow MC, Ambros V. Effect of life history on microRNA expression during *C. elegans* development. *RNA*. 2011; 17:639–651. [PubMed: 21343388]
34. Kato M, Paranjape T, Muller RU, et al. The mir-34 microRNA is required for the DNA damage response in vivo in *C. elegans* and in vitro in human breast cancer cells. *Oncogene*. 2009; 28:2419–2424. [PubMed: 19421141]
35. de Lencastre A, Pincus Z, Zhou K, Kato M, Lee SS, Slack FJ. MicroRNAs both promote and antagonize longevity in *C. elegans*. *Curr Biol*. 2010; 20:2159–2168. [PubMed: 21129974]
36. de Wit E, Linsen SE, Cuppen E, Berezikov E. Repertoire and evolution of miRNA genes in four divergent nematode species. *Genome Res*. 2009; 19:2064–2074. [PubMed: 19755563]
37. McReynolds LA, DeSimone SM, Williams SA. Cloning and comparison of repeated DNA sequences from the human filarial parasite *Brugia malayi* and the animal parasite *Brugia pahangi*. *Proc Natl Acad Sci U S A*. 1986; 83:797–801. [PubMed: 3003750]
38. Foster, JM.; Hoerauf, A.; Slatko, BE.; Taylor, MJ. The molecular biology, immunology and chemotherapy of *Wolbachia* bacterial endosymbionts of filarial nematodes. In: Kennedy, M.; Harnett, W., editors. *Parasitic Nematodes: Molecular Biology, Biochemistry and Immunology*. Wallingford, UK: CABI; 2011. (in press).
39. Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH. How many species are infected with *Wolbachia*? – A statistical analysis of current data. *FEMS Microbiol Lett*. 2008; 281:215–220. [PubMed: 18312577]
40. Werren JH, Baldo L, Clark ME. *Wolbachia*: master manipulators of invertebrate biology. *Nat Rev Microbiol*. 2008; 6:741–751. [PubMed: 18794912]
41. Werren JH. Biology of *Wolbachia*. *Annu Rev Entomol*. 1997; 42:587–609. [PubMed: 15012323]
42. Foster J, Ganatra M, Kamal I, et al. The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode. *PLoS Biol*. 2005; 3:e121. [PubMed: 15780005]
43. Brownlie JC, Adamski M, Slatko B, McGraw EA. Diversifying selection and host adaptation in two endosymbiont genomes. *BMC Evol Biol*. 2007; 7:68. [PubMed: 17470297]
44. Foster JM, Raverdy S, Ganatra MB, Colussi PA, Taron CH, Carlow CK. The *Wolbachia* endosymbiont of *Brugia malayi* has an active phosphoglycerate mutase: a candidate target for anti-filarial therapies. *Parasitol Res*. 2009; 104:1047–1052. [PubMed: 19043737]
45. Henrichfreise B, Schiefer A, Schneider T, et al. Functional conservation of the lipid II biosynthesis pathway in the cell wall-less bacteria *Chlamydia* and *Wolbachia*: why is lipid II needed? *Mol Microbiol*. 2009; 73:913–923. [PubMed: 19656295]
46. Holman AG, Davis PJ, Foster JM, Carlow CK, Kumar S. Computational prediction of essential genes in an unculturable endosymbiotic bacterium, *Wolbachia* of *Brugia malayi*. *BMC Microbiol*. 2009; 9:243. [PubMed: 19943957]
47. Johnston KL, Wu B, Guimaraes A, Ford L, Slatko BE, Taylor MJ. Lipoprotein biosynthesis as a target for anti-*Wolbachia* treatment of filarial nematodes. *Parasit Vectors*. 2010; 3:99. [PubMed: 20946650]
48. Raverdy S, Foster JM, Roopenian E, Carlow CK. The *Wolbachia* endosymbiont of *Brugia malayi* has an active pyruvate phosphate dikinase. *Mol Biochem Parasitol*. 2008; 160:163–166. [PubMed: 18539347]
49. Wu B, Novelli J, Foster J, et al. The heme biosynthetic pathway of the obligate *Wolbachia* endosymbiont of *Brugia malayi* as a potential anti-filarial drug target. *PLoS Negl Trop Dis*. 2009; 3:e475. [PubMed: 19597542]
50. McNulty SN, Foster JM, Mitreva M, et al. Endosymbiont DNA in endobacteria-free filarial nematodes indicates ancient horizontal genetic transfer. *PLoS ONE*. 2010; 5:e11029. [PubMed: 20543958]
51. Brownlie JC, Cass BN, Riegler M, et al. Evidence for metabolic provisioning by a common invertebrate endosymbiont, *Wolbachia pipientis*, during periods of nutritional stress. *PLoS Pathog*. 2009; 5 e1000368.

52. Hosokawa T, Koga R, Kikuchi Y, Meng XY, Fukatsu T. *Wolbachia* as a bacteriocyte-associated nutritional mutualist. *Proc Natl Acad Sci U S A*. 2010; 107:769–774. [PubMed: 20080750]
53. Bordenstein SR, Paraskevopoulos C, Hotopp JC, et al. Parasitism and mutualism in *Wolbachia*: what the phylogenomic trees can and cannot say. *Mol Biol Evol*. 2009; 26:231–241. [PubMed: 18974066]
54. Casiraghi M, Bordenstein SR, Baldo L, et al. Phylogeny of *Wolbachia pipientis* based on *gltA*, *groEL* and *ftsZ* gene sequences: clustering of arthropod and nematode symbionts in the F supergroup, and evidence for further diversity in the *Wolbachia* tree. *Microbiology*. 2005; 151:4015–4022. [PubMed: 16339946]
55. Klasson L, Walker T, Sebaihia M, et al. Genome evolution of *Wolbachia* strain wPip from the *Culex pipiens* group. *Mol Biol Evol*. 2008; 25:1877–1887. [PubMed: 18550617]
56. Klasson L, Westberg J, Sapountzis P, et al. The mosaic genome structure of the *Wolbachia* wRi strain infecting *Drosophila simulans*. *Proc Natl Acad Sci U S A*. 2009; 106:5725–5730. [PubMed: 19307581]
57. Wu M, Sun LV, Vamathevan J, et al. Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: a streamlined genome overrun by mobile genetic elements. *PLoS Biol*. 2004; 2:E69. [PubMed: 15024419]

Table 1

Filarial and *Wolbachia* genomes

	Vertebrate host	Status of genome sequencing	Number EST's (clusters) ^a	<i>Wolbachia</i>	Status of genome
<i>Brugia malayi</i>	Human	Build 1 complete build 2 in progress	25 844 (9961)	Yes	Completed Draft
<i>Brugia timori</i>	Human	–	–	Yes	–
<i>Wuchereria bancrofti</i>	Human	Assembly and annotation in progress	4871 (2279)	Yes	Assembly and annotation in progress
<i>Onchocerca volvulus</i>	Human	Assembly and annotation in progress	14 791 (5144)	Yes	Assembly and annotation in progress
<i>Loa loa</i>	Human	Assembly and annotation in progress	27 (26)	No	–
<i>Mansonella perstans</i>	Human	–	–	Yes	–
<i>Mansonella streptocerca</i>	Human	–	–	No	–
<i>Mansonella ozzardi</i>	Human	–	–	Yes	–
<i>Brugia pahangi</i>	Feline	–	28 (28)	Yes	–
<i>Dirofilaria immitis</i>	Canine, feline	Assembly and annotation in progress	3762 (1788)	Yes	Assembly and annotation in progress
<i>Onchocerca ochengi</i>	Bovine	Assembly and annotation in progress	59 (37)	Yes	Assembly and annotation in progress
<i>Setaria digitata</i>	Bovine	–	–	No	–
<i>Onchocerca cervicalis</i>	Equine	–	–	Yes	–
<i>Onchocerca flexuosa</i>	Cervidae	Genome survey ^b	2063 (1527)	No	–
<i>Litomosoides sigmodontis</i>	Rodent	Assembly and annotation in progress	2699 (1631)	Yes	Assembly and annotation in progress
<i>Acanthocheilonema viteae</i>	Rodent	Genome survey ^b	–	No	–

^aNumber of EST's and clusters compiled in the most recent release of NEMBASE 4 (<http://www.nematodes.org/nembase4/>);

^bMcNulty et al. 2010 (50). ESTs, expressed sequence tags.