The evolution of parasitism in Nematoda

Citation for published version:

Digital Object Identifier (DOI):
10.1017/S0031182014000791

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Parasitology

Publisher Rights Statement:
Copyright © Cambridge University Press 2014 The online version of this article is published within an Open Access environment subject to the conditions of the Creative Commons Attribution licence http://creativecommons.org/licenses/by/3.0/

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
The evolution of parasitism in Nematoda

MARK BLAXTER* and GEORGIOS KOUTSOVOULOS

Institute of Evolutionary Biology, The University of Edinburgh, Edinburgh EH9 3JT, UK

(Received 19 February 2014; revised 16 April 2014; accepted 16 April 2014)

SUMMARY

Nematodes are abundant and diverse, and include many parasitic species. Molecular phylogenetic analyses have shown that parasitism of plants and animals has arisen at least 15 times independently. Extant nematode species also display lifestyles that are proposed to be on the evolutionary trajectory to parasitism. Recent advances have permitted the determination of the genomes and transcriptomes of many nematode species. These new data can be used to further resolve the phylogeny of Nematoda, and identify possible genetic patterns associated with parasitism. Plant-parasitic nematode genomes show evidence of horizontal gene transfer from other members of the rhizosphere, and these genes play important roles in the parasite-host interface. Similar horizontal transfer is not evident in animal parasitic groups. Many nematodes have bacterial symbionts that can be essential for survival. Horizontal transfer from symbionts to the nematode is also common, but its biological importance is unclear. Over 100 nematode species are currently targeted for sequencing, and these data will yield important insights into the biology and evolutionary history of parasitism. It is important that these new technologies are also applied to free-living taxa, so that the pre-parasitic ground state can be inferred, and the novelties associated with parasitism isolated.

Key words: Nematoda, nematodes, parasitism, evolution, genome, symbiont, Wolbachia, phylogeny, horizontal gene transfer.

THE DIVERSITY OF THE NEMATODA

Nematoda is an ancient and biologically diverse phylum of moulted animals. They range in size from 0.2 mm to over 6 m, and can be found in most habitats, including within and on host animals and plants (Blaxter and Denver, 2012). In many marine and terrestrial sediments they are the most abundant group in terms of individuals (Platonova and Gal’tsova, 1976), and while only approximately 23,000 species have been described (J. Hallan, unpublished; https://insects.tamu.edu/research/collection/hallan/), the true species-level diversity may be 1 million or more (Lambshead, 1993). Most terrestrial plants and larger animals are associated with at least one species of parasitic nematode, and most of the human population experiences nematode parasitism during their lives (with perhaps one quarter to one third of the population experiencing nematode parasitism during their lives). Estimates of the number of species of parasitic nematode per host suggest that there may be of the order of 25,000 nematode parasites just of vertebrates, most of which remain undescribed (Dobson et al. 2008). Nematodes are thus important regulators of plant and animal production. Understanding the evolutionary origins of plant and animal parasitism, and the mechanisms by which parasites locate and invade their hosts, avoid host immunity, and acquire nutrition, are important goals for not only basic, but also for medical and veterinary science. In this paper we discuss the changes in our understanding of the diversity and relationships of nematodes, and of the biology of their parasitic habits, that have been brought about by study of their genes and, increasingly, genomes.

Nematoda are part of Ecdysozoa, a superphylum of animals first defined through analyses of molecular markers (Aguinaldo et al. 1997). Support for Ecdysozoa as distinct from other groupings of protostome taxa is less strong from analyses of morphological characters (Nielsen, 2001). Ecdysozoan phyla are characterized by the presence of a cuticle that is periodically moulted during the life cycle, though the specifics of the molecular nature of the cuticle and the orchestration of ecdysis differ between phyla. Other shared features adduced as evidence of relatedness between these phyla include an absence of cilia in adults, and in many members the presence of a triradiate pharynx. The Ecdysozoa in turn comprises two groups, the Panarthropoda (phyla Tardigrada, Onychophora and Arthropoda) and Cycloneuralia (Nematoda, Nematomorpha, Priapulida, Kinorhyncha and Loricifera). Within Cycloneuralia, which may be paraphyletic with respect to Panarthropoda, Nematoda are consistently placed as sisters to Nematomorpha in morphological and molecular analyses (Schmidt-Rhaesa, 1997; Dunn et al. 2008).

Nematomorpha are a fascinating group of obligate parasites of terrestrial (Gordioidea) and marine (Nectonematoidea) arthropods. These ‘horsehair worms’ have a parasitoid life cycle, with the larval
Fig. 1. The phylogenetic structure of the Nematoda and the origins of parasitism (A) A cartoon of the phylogenetic structure of the Nematoda, based on nuclear small subunit ribosomal RNA analyses and interpretation of taxon relationships derived from morphology (De Ley and Blaxter, 2004; Blaxter and Denver, 2012). Taxon systematic
The evolution of parasitism in Nematoda

The multiple origins of parasitism within the Nematoda

Nematodes exhibit a wide range of relationships with other species. Parasitism is a common way of life, and a large proportion of nematode species may be parasites. Poulin has usefully classified the different kinds of parasitic relationships between species into a spectrum of life-habit modes (Poulin, 2011; Poulin and Randhawa, 2013). Some relationships are phoretic: the nematodes use another species to aid dispersal to new sites (Bovien, 1937). Phoretic associations can be very specific, as in Rhabditis stammeri, associates of burying beetles (Nicrophorus spp.), which have a complex response to the beetle life cycle that assures their presence in emerging adults by entering and diapausing in the hind guts of mature larvae before they pupate (Richter, 1993).
Other phoretic relationships are less specific, and dispersal stages can be found attached to many different transport hosts. In most cases, the dispersal stage is a third stage juvenile (J3, or L3 for larva). The costs to the phoretic host are hard to measure, but may be significant in heavily colonized hosts, or where the associate extracts some nutrition from its carrier. Many plant parasitic nematodes, particularly the migratory endoparasites, could be classified as microherbivores (Poulin and Randhawa, 2013), as their ecology is similar to that of an ungulate browsing on bushy plants. However these nematodes do induce specific cell responses in host plants, and thus the relationship is more than just browsing. Some intestinal parasites, such as the Rhigonematida of millipedes (Hunt, 1996), feed on gut contents or other nematode parasites rather than on host tissue, and might even be classed as commensals.

Independent origins of the parasitic habit can be validated by molecular phylogenetic placement of parasitic taxa and their free-living relatives (Blaxter et al. 1998; Dorris et al. 1999). With the current available molecular data we can define three origins of plant parasitism, 10 of parasitism of a wide range of non-vertebrates and five of parasitism of vertebrate hosts across the three subclasses (Fig. 1 and Table 1). Many additional events of acquisition of parasitic lifestyles could be proposed. For example, Sudhaus has suggested at least 20 independent events of acquisition of parasitism of insects in Nematoda (Sudhaus, 2008). Enoplia has the fewest parasitic species, while Chromadia (and Rhabditida within Chromadia) has the most. All parasitic groups appear to have a terrestrial or limnic origin. The sole enoplian animal parasite, Ironus macrocephalum (Ironsidae), was described from the earthworm Pheretima wendessiana in New Guinea (Pierantoni, 1916). Other Ironidae are freshwater species.

There are some striking phylogenetic associations between non-vertebrate and vertebrate parasites. The Strongylomorpha (in Rhabditina), a major group of gut and airway parasites of vertebrates, are sisters to the Heterorhabditidae. Heterorhabditis species are entomopathogens that invade the haemocoel of insect larvae, and release a symbiotic bacterium that kills the host. Similarly, the mammal-parasitic Strongyloidoidae (Tylenchina; Panagrolaimomorpha) are related to the entomopathogenic Steinernematidae. Steinernema species also invade the haemocoel, and use a bacterial symbiont (a different group of species) to kill their hosts. The entomopathogens show convergence in life style, and are phylogenetically associated with a transition to vertebrate parasitism. However, there is no evidence that the vertebrate parasites utilize symbiotic bacteria during their life cycle. One model for the origin of the vertebrate-parasitic Strongylomorphs and Strongyloidoids is that they represent host capture by an ancestral terrestrial, entomopathogenic or entomoparasitic species, and subsequent radiation in the new host groups. Not all associations with arthropods necessarily lead to full parasitism, as there are several nematode groups (notably the Diplogasteromorpha in Rhabditina) where many species are phoretic or necromenic associates of arthropods but very few parasites have been found.

The Spirurina (Chromadia; Rhabditida) are all parasites, mostly in vertebrates. Many Spirurina utilize vector hosts to facilitate transit from one definitive host to another, but there are also groups (Ascaridomorpha, Oxyuridomorpha) that do not use intermediate hosts. Some groups utilize multiple intermediate hosts, with Gnathostomatomorpha passing through both a first, crustacean paratenic host and a second, fish host before establishing in carnivorous mammals. Surprisingly, in the new Spirurina phylogeny, groups with a simple, direct life cycle appear to be derived from a radiation of groups that have vector hosts. Gnathostomatomorpha arise basally, and Ascaridomorpha and Oxyuridomorpha have their origins as sisters to other vector-borne groups such as Spirurumorpha (Nadler et al. 2007; Laetsch et al. 2012). Loss and gain of vector hosts within groups are common, and radical shift of vector species while continuing to use similar definitive hosts is a common feature of related parasites. For example, the Onchoerocerciae, parasitizing rodents, ungulates and primates, utilize mosquitoes, tabanid flies, black flies and mites as vectors. Two groups of arthropod parasites are nested within the otherwise vertebrate parasitic Spirurina (Rhigonematomorpha are gut parasites of large millipedes, and many Oxyuridomorpha are parasites of arthropods), and thus there must have been at least two independent transitions from vertebrate to arthropod parasitism (possibly by neotenic development in vector species) in these groups. Interestingly, Philometra obturans, a dracunculoid parasite of pike, has been observed to grow to sexual maturity in copepod intermediate hosts when the fish definitive host is absent (Moravec and de Buron, 2013).

**COMMON THEMES**

There are few common themes in the parasitic lifestyles of nematodes. One is that the stage at which parasites transition from a free-living portion of the life cycle to the parasitic portion (or vice versa) is usually the J3/L3 stage, the same stage usually involved in phoretic associations (Sudhaus, 2010). Most rhabditid parasites first invade their hosts as J3, and mermithids exit their hosts as J3. The diapausing J3 stage of the free-living nematode Caenorhabditis elegans (not a parasite, but a species that has phoretic associations with terrestrial isopods) is particularly resistant to environmental insult, and is used as a model for the infective J3 of parasitic nematodes.

---

Mark Blaxter and Georgios Koutsovoulos
<table>
<thead>
<tr>
<th>Origin number&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Nematode group</th>
<th>Host group(s)</th>
<th>Parasitic types sensu Poulin&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-vertebrate hosts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (multiple events)</td>
<td><em>Heterorhabditis</em>, <em>Phasmarhabditis</em>, <em>Rhabditis</em> (and others)</td>
<td>Hexapoda, Mollusca, Clitellata</td>
<td>Parasitoid, Castrator, Directly transmitted parasite</td>
<td>Rhabditomorpha contains many taxa with phoretic relationships with arthropods and molluscs</td>
</tr>
<tr>
<td>2 (multiple events)</td>
<td>Diplogasteromorpha (e.g., <em>Parasitodiplogaster</em>, <em>Medhinema</em>, <em>Cephalobium</em>)</td>
<td>Directly transmitted parasite</td>
<td>Diplogasteromorpha contains many species with phoretic or necromenic associations with arthropods</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Drilonematomorpha</td>
<td>Clitellata</td>
<td>Directly transmitted parasite</td>
<td>Some species show ‘alternation of generations’ where the parasite can reproduce both within and outside the host; <em>Daudheyelia</em>, a parasite of snails, may be a member of <em>Tylenchomorpha</em></td>
</tr>
<tr>
<td>4</td>
<td>Steinernematidae, Allantonematidae</td>
<td>Hexapoda</td>
<td>Directly transmitted parasite</td>
<td></td>
</tr>
<tr>
<td>5 (multiple events)</td>
<td>Hexatyline, Aphelenchoidea, Sphaerularioidea (and others)</td>
<td>Hexapods</td>
<td>Castrator, Directly transmitted parasite</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Rhigonematomorpha</td>
<td>Myriapoda</td>
<td>Directly transmitted parasite</td>
<td>Rhigonematomorpha is nested within the otherwise vertebrate parasitic Spirurina; the group contains no vertebrate parasites</td>
</tr>
<tr>
<td>7</td>
<td>Oxyuridomorpha</td>
<td>Arthropoda</td>
<td>Directly transmitted parasite</td>
<td>Oxyuridomorpha is nested within the otherwise vertebrate parasitic Spirurina; the group also contains vertebrate parasites</td>
</tr>
<tr>
<td>8</td>
<td>Monhysterina (<em>Gammarinema</em>)</td>
<td>Crustacea</td>
<td>Directly transmitted parasite</td>
<td>May be classed as ‘commensals’ rather than parasites</td>
</tr>
<tr>
<td>9</td>
<td>Ironina</td>
<td>Annelida</td>
<td>Directly transmitted parasite</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Mermithida</td>
<td>Arthropoda</td>
<td>Parasitoid</td>
<td></td>
</tr>
<tr>
<td><strong>Vertebrate hosts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Strongyloidea</td>
<td>Vertebrates</td>
<td>Directly transmitted parasite, Trophically transmitted parasite, Vector transmitted parasite</td>
<td>Some species have non-vertebrate paratenic or vector hosts; <em>Heterorhabditis</em> is a sister group to the vertebrate parasites</td>
</tr>
<tr>
<td>2 (multiple events)</td>
<td>Rhabdiasidae, <em>Pelodera strongyloides</em></td>
<td>Anuran</td>
<td>Directly transmitted parasite</td>
<td>There may be additional independent he precise placement of Rhabdiasidae awaits resolution</td>
</tr>
<tr>
<td>3</td>
<td>Strongyloididae</td>
<td>Mammals</td>
<td>Directly transmitted parasite</td>
<td>‘Alternation of generations’ where the parasite can reproduce both within and outside the host</td>
</tr>
<tr>
<td>4</td>
<td>Spiruromorpha</td>
<td>Vertebrates</td>
<td>Directly transmitted parasite, vector transmitted parasite, trophically transmitted parasite</td>
<td>Includes the non-vertebrate parasitic oxyurids and rhigonematids</td>
</tr>
<tr>
<td>5</td>
<td>Trichinellida plus Dioctophymatida</td>
<td>Mammals</td>
<td>Directly transmitted parasite, trophically transmitted parasite</td>
<td></td>
</tr>
</tbody>
</table>

**Plant hosts**

<table>
<thead>
<tr>
<th>Origin number&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Nematode group</th>
<th>Host group(s)</th>
<th>Parasitic types sensu Poulin&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (multiple events)</td>
<td>Tylenchomorpha</td>
<td>Viridiplantae (and macroalgae)</td>
<td>Directly transmitted parasite, Micro-predator</td>
<td>Plant parasitic groups are associated with fungal-feeding groups, and there may be an association; plant parasitism may have arisen multiple times</td>
</tr>
<tr>
<td>2</td>
<td>Dipterophorina (Trichodoridae)</td>
<td>Viridiplantae</td>
<td>Directly transmitted parasite, Micro-predator</td>
<td></td>
</tr>
<tr>
<td>3 (multiple events)</td>
<td>Dorylaimida (<em>Xiphinema</em>, <em>Longidorus</em> and others)</td>
<td>Viridiplantae</td>
<td>Directly transmitted parasite, Micro-predator</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> There are many isolated additional descriptions of nematode associations with other taxa.  
<sup>b</sup> The numbering of events follows Fig. 1.  
<sup>c</sup> See Poulin (2011); Poulin and Randhawa (2013) for details.
(particularly in Rhabditomorpha). It has been suggested that adaptation to rich food sources low in available oxygen such as rotting vegetation might predispose some groups to the acquisition of parasitic habits as the guts of other animals might offer a similar set of environmental challenges.

Echoes of the ‘great chain of being’, a world view common since pre-Darwinian times but still expressed today, can be heard in many discussions of parasite evolution. Parasites are suggested to be likely to show regression, losing morphological and metabolic complexity as they come to rely on hosts for more and more of the details of life as a metazoan. However, while a gut parasite might rely on the host for metabolic provisioning (thus implying a scope for simplification), it may also be subject to new stresses and demands in terms of countering host immunity (implying evolution of novelty). Similarly, parasites that utilize multiple hosts, or have major transitions in lifestyle, might be expected to gain or at least retain metabolic complexity, as they require toolkits to thrive in several distinct environments. Morphological novelties are common in parasitic nematodes: buccal teeth in hookworms (Chromadoria; Strongylophora); tagmatization in Trichura (Dorylaimia; Trichinellidae); ornate spines, flanges and other cuticular decorations observed in arthropod-parasitic Oxyuridomorpha and Rhigonematomorpha (Chromadoria; Spirurina).

Even parasites with apparently simple life cycles can have very complex biology. Most Ascaridids infect new hosts through ingestion of embryonated eggs from the environment, and are often used as examples of simple, direct life cycles. Ascaris suum, which enters the host intestine as larvae arrested within a protective eggshell, but then invades the gut wall, enters the bloodstream, and reinfects the gut by exiting from the capillaries in the lungs, climbing the tracheal tree and returning to the gut by being swallowed, has a very complex in-host life cycle. This migration seems against reason, as it is costly to the parasite, and has been argued to be retention of an ancient trait of migration derived from a vector-borne ancestor. Retention of a costly phenotype in contemporary species requires explanation. The answer, derived by analysis of species pairs that differ in whether they migrate or not, is simply that migration is strongly associated with adult body size and with adult fecundity: migration is an adaptive trait maintained by selection (Skorping et al. 1991; Read and Skorping, 1995a, b). It may be that the ability to migrate was inherited from an ancestral species, but its presence in living taxa is because it is currently adaptive.

**Genomics and Parasitism**

The first metazoan genome to be sequenced was that of a nematode: *C. elegans* (Chromadoria; Rhabditomorpha) (The *C. elegans* Genome Sequencing Consortium 1998) and nematode parasitologists were early adopters of the genomics toolkit (Blaxter et al. 1996, 1997). A major effort over the past 20 years has delivered a large body of data on the transcribed genes of nematodes, initially through Sanger di-deoxy sequencing expressed sequence tag sampling approaches (Parkinson et al. 2004) but latterly using the new sequencing technologies to deliver whole transcriptomes (Blaxter et al. 2012). The roster of nematode species with significant transcriptome data publicly available now exceeds 100. These data have been critical in revealing genes associated with parasitic traits, and in identifying new drug targets and vaccine candidates (Elsworth et al. 2011).

The first parasitic nematode genome to be sequenced was that of *Brugia malayi* (Chromadoria; Spiruromorpha) (Ghedin et al. 2007). To date 19 nematode genome sequences have been published (see Fig. 1A and Table 2), and the majority of these genomes derive from parasites. As might be expected there is a bias in the species chosen for sequencing, and vertebrate and especially human parasites are over-represented. New sequencing technologies allow a new draft genome to be generated with very little resource, and more and more groups are now exploring the genomes of their target taxa. A loose alliance of nematode genome-interested biologists has been founded, and is nucleating the sequencing of as many genomes across the diversity of the phylum as is possible. The 959 Nematode Genomes initiative (Kumar et al. 2012a, b) mimics other community genome programmes such as the insect 5k project (i5K Consortium, 2013) (excepting that the headline goal is derived from the number of somatic cells in the adult female hermaphrodite *C. elegans*). It allows contributors to post information on the genomes they are pursuing and update the community as to their successes, and is a clearing-house for queries and collaborations (see http://www.nematodegenomes.org). By showing that genome sequencing, assembly, annotation and interpretation are within the reach of any research community, the initiative now has recorded well over 100 species for which genome data have been or are being generated.

We have been working to deliver genomic data for a number of species of key interest. These include free-living species (for example representatives of the under-sampled Enoplia, and additional Caenorhabditis species that will illuminate the biology of the model *C. elegans*) as well as parasites (Table 2). Using short-read next-generation sequencing technologies, and freely available assembly and annotation toolkits, we have been able to deliver important new genomes to the research communities. While generation of genomes of the quality and completeness of *C. elegans* would still require very significant
## Table 2. Nematode genome sequences

<table>
<thead>
<tr>
<th>Species</th>
<th>Systematic position</th>
<th>Genome size (Mb)</th>
<th>Contiguity of assembly N50 (kb)</th>
<th>Number of scaffolds</th>
<th>Number of nuclear protein-coding genes predicted</th>
<th>Web link for access (publication)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clade I: Dorylaimia</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Trichinella spiralis</em></td>
<td>Trichinellida</td>
<td>63·51</td>
<td>6 373 445</td>
<td>6 819</td>
<td>16 380</td>
<td><a href="http://www.wormbase.org/species/t_spiralis">www.wormbase.org/species/t_spiralis</a> (Mitreva et al. 2011)</td>
</tr>
<tr>
<td><strong>Clade III: Spirurina</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Acanthocheilonema viteae</em></td>
<td>Spiruromorpha</td>
<td>77·35</td>
<td>25 808</td>
<td>6 796</td>
<td>10 397</td>
<td><a href="http://acanthocheilonema.nematod.es">http://acanthocheilonema.nematod.es</a></td>
</tr>
<tr>
<td><em>Brugia malayi</em></td>
<td>Spiruromorpha</td>
<td>94·14</td>
<td>191 089</td>
<td>9 827</td>
<td>17 846</td>
<td><a href="http://badger.bio.ed.ac.uk/filarial/">http://badger.bio.ed.ac.uk/filarial/</a></td>
</tr>
<tr>
<td><em>Dirofilaria immitis</em> (and <em>Wolbachia wDi</em>)</td>
<td>Spiruromorpha</td>
<td>88·30 (0·92)</td>
<td>22 560 (919 954)</td>
<td>71 281 (2)</td>
<td>16 061 (871)</td>
<td><a href="http://dirofilaria.nematod.es">http://dirofilaria.nematod.es</a></td>
</tr>
<tr>
<td><em>Litomosoides sigmodonti</em> (and <em>Wolbachia wLs</em>)</td>
<td>Spiruromorpha</td>
<td>64·81 (1·05)</td>
<td>45 863 (605 213)</td>
<td>3165 (10)</td>
<td>10 246 (1042)</td>
<td><a href="http://badger.bio.ed.ac.uk/filarial/">http://badger.bio.ed.ac.uk/filarial/</a> (Comandatore et al. 2013)</td>
</tr>
<tr>
<td><em>Loa loa</em></td>
<td>Spiruromorpha</td>
<td>91·37</td>
<td>174 388</td>
<td>5 770</td>
<td>15 444</td>
<td><a href="http://www.broadinstitute.org/annotation/genome/filarial_worms">www.broadinstitute.org/annotation/genome/filarial_worms</a> (Desjardins et al. 2013)</td>
</tr>
<tr>
<td><em>Onchocerca ochengi</em> (and <em>Wolbachia wOo</em>)</td>
<td>Spiruromorpha</td>
<td>95·51 (0·96)</td>
<td>12 317 (957 990)</td>
<td>24 057 (1)</td>
<td>13 990 (664)</td>
<td><a href="http://onchocerca.nematod.es">http://onchocerca.nematod.es</a></td>
</tr>
<tr>
<td><em>Wuchereria bancrofti</em></td>
<td>Spiruromorpha</td>
<td>81·51</td>
<td>5 161</td>
<td>25 884</td>
<td>19 327</td>
<td><a href="http://www.broadinstitute.org/annotation/genome/filarial_worms">www.broadinstitute.org/annotation/genome/filarial_worms</a> (Desjardins et al. 2013)</td>
</tr>
<tr>
<td><strong>Clade IV: Tylenchina</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Panagrellus redivivus</em></td>
<td>Panagrolaimomorpha</td>
<td>65·06</td>
<td>267 941</td>
<td>867</td>
<td>24 249</td>
<td><a href="http://www.wormbase.org/species/p_redivivus">www.wormbase.org/species/p_redivivus</a> (Srinivasan et al. 2013)</td>
</tr>
<tr>
<td><em>Bursaphelenchus xylophilus</em></td>
<td>Tylenchomorpha</td>
<td>74·56</td>
<td>949 830</td>
<td>5 526</td>
<td>18 074</td>
<td><a href="http://www.wormbase.org/species/b_xylophilus">www.wormbase.org/species/b_xylophilus</a> (Kikuchi et al. 2011)</td>
</tr>
<tr>
<td><em>Meloidogyne hapla</em></td>
<td>Tylenchomorpha</td>
<td>53·02</td>
<td>37 608</td>
<td>3 452</td>
<td>14 420</td>
<td><a href="http://www.wormbase.org/species/m_hapla">www.wormbase.org/species/m_hapla</a> (Opperman et al. 2008)</td>
</tr>
<tr>
<td><em>Meloidogyne incognita</em></td>
<td>Tylenchomorpha</td>
<td>82·09</td>
<td>12 786</td>
<td>9 533</td>
<td>19 212</td>
<td><a href="http://www.wormbase.org/species/m_incognita">www.wormbase.org/species/m_incognita</a> (Abad et al. 2008)</td>
</tr>
<tr>
<td><strong>Clade V: Rhabdina</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Caenorhabditis angaria</em></td>
<td>Rhabditomorpha</td>
<td>99·01</td>
<td>8 7708</td>
<td>11 453</td>
<td>27 967</td>
<td><a href="http://www.wormbase.org/species/c_angaria">www.wormbase.org/species/c_angaria</a> (Mortazavi et al. 2010)</td>
</tr>
<tr>
<td><em>Caenorhabditis briggsae</em></td>
<td>Rhabditomorpha</td>
<td>108·42</td>
<td>17 485 439</td>
<td>12</td>
<td>21 850</td>
<td><a href="http://www.wormbase.org/species/c_briggsae">www.wormbase.org/species/c_briggsae</a> (Stein et al. 2003)</td>
</tr>
<tr>
<td><em>Caenorhabditis elegans</em></td>
<td>Rhabditomorpha</td>
<td>100·29</td>
<td>17 493 829</td>
<td>7</td>
<td>20 520</td>
<td><a href="http://www.wormbase.org/species/c_elegans">www.wormbase.org/species/c_elegans</a> (The C. elegans Genome Sequencing Consortium 1998)</td>
</tr>
<tr>
<td><em>Caenorhabditis sp 1</em></td>
<td>Rhabditomorpha</td>
<td>109·33</td>
<td>24 542</td>
<td>14 350</td>
<td>d</td>
<td><a href="http://caenorhabditis.nematod.es">http://caenorhabditis.nematod.es</a></td>
</tr>
<tr>
<td>Species</td>
<td>Systematic position</td>
<td>Contiguity of assembly (N50a)</td>
<td>Genome size (Mb)</td>
<td>Number of nuclear protein-coding genes</td>
<td>Number of scaffoldolds</td>
<td>web link for access (publication)</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------------------</td>
<td>-------------------------------</td>
<td>------------------</td>
<td>----------------------------------------</td>
<td>------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Caenorhabditis sp. 16</td>
<td>Rhabditomorpha</td>
<td>86.00</td>
<td>57,356</td>
<td>8,419</td>
<td>24,217</td>
<td><a href="http://caenorhabditis.nematod.es">http://caenorhabditis.nematod.es</a></td>
</tr>
<tr>
<td>Haemonchus contortus</td>
<td>Rhabditomorpha</td>
<td>77.01</td>
<td>312,328</td>
<td>129.0</td>
<td>24,217</td>
<td><a href="http://caenorhabditis.nematod.es">http://caenorhabditis.nematod.es</a></td>
</tr>
</tbody>
</table>

* Genome size is estimated from the span of genome assembly.  
  a Number of nuclear protein-coding genes.
  b The N50 is the weighted median scaffold length (the scaffold length at which 50% of the assembled genome is in scaffold of that length or greater).  
  c The A. suum genome undergoes chromatin diminution such that the somatic genome is ~40 Mb smaller than the germline genome (Wang et al. 2012).  
  d The gene predictions for sp. 5 are preliminary. The gene sets for the Caenorhabditis species are being re-predicted as part of a co-analysis across 10 Caenorhabditis genomes (E. Schwarz, M. Blaxter, unpublished).  
  e The Romanormermis culicivorax has a genome of ~322 Mb (Schiffer et al. 2013), A. suum 334 Mb (Wang et al. 2012) and Haemonchus contortus 370 Mb (Laing et al. 2013)). Within clades, genome sizes can vary widely (for example within the genus Meloidogyne sizes range from 54 Mb in M. hapla to 150 Mb estimated in Meloidogyne incognita (Abad et al. 2008; Lunt et al. 2014)), and this is associated with polyploidy. The knowledge of genome sizes of free-living nematodes outside the Rhabditomorpha is very sparse, and it may be that free-living, diploid species’ genomes regularly exceed the two-fold range known thus far. In general, the numbers of genes predicted from parasitic species are less than or equivalent to those predicted from free-living species. Caenorhabditis elegans has ~21 800 protein-coding genes, while H. contortus (Chromadoria; Strongylomorpha) has 20 600 (Laing et al. 2013). However, the mermithid R. culicivorax (Dorylaimia; Mermithida) has ~12 000 (Schiffer et al. 2013), Trichinella spiralis (Dorylaimia; Trichinellida) has only 15 800 (Mitreva et al. 2011), and the onchocercid nematodes (B. malayi and relatives; Chromadoria; Spiruromorpha) have between 12 000 and 14 000 (Godel et al. 2012; Desjardins et al. 2013). Whether this reveals a loss of genic complexity in some parasites, or an overall more complex genomic heritage in the Rhabditomorpha remains unclear. Changes in genome size are generally reflected in congruent changes in intron length, intergenic distance and repeat content.
Plant parasitic nematodes must overcome the formidable defences of the plant cell wall in order to extract nutrition from their hosts. In addition, many sedentary plant parasites induce galls of various forms, structures induced in the plant by a parasite that must ‘know’ some of the tricks of plant developmental biology. Plant parasitic species secrete cellulolytic enzymes, and cloning and sequencing of these effectors revealed that some had their closest homologues in rhizosphere fungi and bacteria rather than other animals (Smant et al. 1998). These enzymes became the first robustly supported instances of horizontal or horizontal gene transfer into nematode genomes. A diverse roster of plant cell wall-degrading enzymes is now known from cyst and root-knot nematodes (Bird et al. 2014), and supplemented by piecemeal understanding of the repertoires of other phytopathogenic species. Phylogenetic inference suggests that these horizontally transferred genes were acquired early in the evolution of the Tylenchomorpha (Danchin et al. 2010; Rybarczyk-Mydlowska et al. 2012). It is likely, given the inferred species of origin of the genes, that there were several independent events of gene acquisition (Mitreva et al. 2009). Surveying tylenchomorph transcriptomes and genomes for genes with sequence-similarity profiles similar to those of the validated horizontal gene transfers reveals a wide range of candidates that include nod factor homologues and genes of unknown function that have disjunct distributions in plants and plant-parasitic nematodes, or in root bacteria and plant parasitic nematodes (Elsworth et al. 2011).

Importantly, horizontally transferred plant cell wall-degrading enzymes have also been identified in parasites from the other independently evolved plant-parasitic groups (Diptherophorina in Enoplia and Dorylaimida in Dorylaimia), suggesting that while the transition to phytophagy was di
cult, it involved similar evolutionary trajectories in each case: the acquisition of the necessary toolkit from professional saprophytes in the root environment. It is interesting in this respect that similar cellulolytic enzymes have been identified in the free-living Pristionchus pacificus and related species (Rhadinotina; Diplodageteromorpha) (Mayer et al. 2011). Horizontally acquired genes are evident in other species, including C. elegans (Parkinson and Blaxter, 2003).

The close association between animal parasites and their hosts, and the requirement to specifically modulate especially the adaptive and anamnestic immune responses of vertebrate hosts, raises the question as to whether animal parasites also acquired novel genes from their hosts, or other commensals of these hosts. Surveys of expressed sequence tag datasets from plant parasitic species identified many potential horizontal transfer candidates, but very few candidates were identified in either transcriptome data from across the diversity of animal parasites, or in the genomes of animal parasites in Strongylomorpha, Ascaridomorpha, Sphaeromorpha or Trichinellida (Elsworth et al. 2011). One interesting horizontally acquired gene in the Sphaeromorpha is an alphaproteobacteria-like, second copy of ferrochelatase, an enzyme involved in the synthesis of haem (Elsworth et al. 2011). This gene is present in onchocercid nematodes (B. malayi and relatives), and suggests an additional or distinct requirement for haem in these tissue and blood parasites (Wu et al. 2013). The ferrochelatase is quite distinct from that found in the Wolbachia symbionts of onchocercids (see below).

**Bacterial Symbionts**

The long history of the tree of life is punctuated by many, highly significant events of symbiosis. In the Nematoda, several distinct types of symbiosis with bacteria have been recorded (Murfin et al. 2012). The free-living Stilbonematidae (Chromadorea) associate specifically with gammaproteobacteria that grow as a lawn on the nematodes’ cuticle. These sulphur-fixing bacteria act as a major food source for the nematode, and the nematodes ‘farm’ their bacterial associates by migrating to H2S-rich sediment layers (Bulgaresi et al. 2011; Murfin et al. 2012). In Anoplostoma (Enoplia), adult nematodes have neither mouth nor anus, and their guts are filled with a sulphur-fixing symbiont. Other similar trophic symbioses undoubtedly await discovery.

The entomopathogenic genera Heterorhabditis (Rhabditina, Strongylomorpha) (Bai et al. 2013) and Steinernema (Tylenchina, Panagrolaimomorpha) (Goodrich-Blair, 2007) share a life-cycle strategy that utilizes specific Entobacteriaceae bacterial symbionts (Photorhabdus with Heterorhabditis, Xenorhabdus with Steinernema) to kill newly invading insect larvae, and then to provide nutrition to the growing and reproducing nematodes. While the symbions used and the details of the interactions differ, the convergence of these two nematode genera on the same general strategy is remarkable. In the plant-parasitic Xiphinema (Dorylaimia; Dorylaimida), a verucomicrobial symbiont, Xiphinemato bacter, is found intracellularly (Vandekerckhove et al. 2000). Its role in the biology of the nematode is largely unknown, although the symbiont is maternally transmitted and may play a role in modification of the nematodes’ reproductive mode.

The alphaproteobacterium Wolbachia pipientis was first described from insects, where they are reproductive parasites, manipulating the reproductive status, gender or sexual compatibility of their hosts (O’Neill, 1995; Werren, 1997). Wolbachia have subsequently been found in a range of terrestrial...
arthropods, and from nematodes. Molecular phylogenetic data suggest the presence of supergroups of Wolbachia that have distinct biology and host distributions (Lo et al. 2002). Supergroup A and B Wolbachia are most widespread, and are found in insects. Nematodes are infected with supergroup C, D and F Wolbachia. Wolbachia have been described from the Spiruromorpha (in the Onchocercidae), and the Tylenchomorpha (Radopholus similis is the only species with infection described to date) (Jacob et al. 2008; Haegeman et al. 2009). General surveys using Wolbachia-specific PCR assays of many other nematode species have been negative (Bordenstein et al. 2003; Duron and Gavotte, 2007; Foster et al. 2008). However, in R. similis (Tylenchina, Tylenchomorpha) the identification of expressed sequence tags corresponding to likely Wolbachia genes led to the identification of an intracellular symbiont in this plant parasite (Jacob et al. 2008; Haegeman et al. 2009).

In supergroups A and B, the symbiont phylogeny does not match that of its hosts, and host species tend to include both infected and uninfected individuals, reflecting frequent loss and acquisition of the symbiont through phylogenetic time. This pattern reflects the parasitic nature of the symbiosis. In contrast, supergroup C and D Wolbachia from onchocercid nematodes show traits suggestive of long, and possibly essential, mutualist interactions. The Wolbachia and nematode host phylogenies are congruent, indicating little if any host switching (Bandi et al. 1998; Casiraghi et al. 2004). In infected species, all individuals are infected, and killing of the Wolbachia with antibiotics such as tetracyclines also affects the viability of the nematode host, with loss of fecundity and nematode death (Bandi et al. 1999; Hoerauf et al. 1999; Landmann et al. 2011, 2012). The exact nature of the mutualism remains unclear: the Wolbachia may assist the nematode metabolically (the distribution of bacteria in adult nematodes is reminiscent of the distribution of essential nutritive Buchnera bacteria in aphids) or in evading the vertebrate host’s immune system (by confusing T-helper cell polarization with bacterial and metazoan signals at the same time) (Fenn and Blaxter, 2004, 2007; Darby et al. 2012). Genome sequencing of filarial Wolbachia has permitted culling of the possible hypotheses for essentiality, but has not yielded data that definitively support specific metabolic vs immunoprotective roles (Darby et al. 2012).

It is also possible that the interference of the Wolbachia with oogenesis and development (Landmann et al. 2011, 2012) makes it difficult, in evolutionary terms, for the nematode to rid itself of the symbionts. The symbiosis is not essential in the phylogenetic long term, as there are onchocercid species, such as Loa loa, Onchocerca flexuosa and Acanthocheilonema viteae, which have lost the infection and are now Wolbachia-free. Genomically, filarial Wolbachia display the expected phenotypes of mutualist endosymbionts: the genomes are reduced compared with the insect-parasitic supergroups A and B, with fewer protein coding genes and a lack of mobile elements such as phage (Comandatore et al. 2013).

The onchocercid nematodes that lack living Wolbachia still retain a signature of past infection in the form of horizontally transferred fragments of Wolbachia-like DNA in their nuclear genomes (McNulty et al. 2013). Species that have live Wolbachia also have horizontally transferred Wolbachia-like fragments in their genomes (Dunning-Hotopp et al. 2007; Ioannidis et al. 2013). Horizontal transfer of organellar DNA to the nucleus is common, and thus the presence of these Wolbachia fragments could simply be a product of non-functional, stochastic incorporation of Wolbachia fragments into the genome (Blaxter, 2007). More excitingly, these inserted fragments could be being used by the nuclear genome to express new, Wolbachia-derived proteins. While some Wolbachia fragments are expressed at low levels (Ioannidis et al. 2013), most are not, and most are gene fragments that also have disabling mutations that render them inactive. Comparisons between the nuclear genomes of onchocercid species with and without Wolbachia has identified few shared insertions and no smoking gun of a constrained, conserved transfer that might substitute for a live Wolbachia.

The other Wolbachia found in nematodes are much less well-studied. Some onchocercid nematodes carry a Wolbachia that is placed in supergroup F, alongside Wolbachia from termites, fleas and bedbugs (Bordenstein et al. 2003; Duron and Gavotte, 2007; Foster et al. 2008; Jacob et al. 2008; Haegeman et al. 2009; Comandatore et al. 2013). Initial analyses suggested that the R. similis Wolbachia was distantly related to any other supergroup (Jacob et al. 2008; Haegeman et al. 2009), but this result is questionable (Koutsovoulos et al. 2014). The bovine lungworm Dictyocaulus viviparus (Rhabditina, Strongylophora) was not known to have any association with Wolbachia until its genome was sequenced (Koutsovoulos et al. 2014). Within the nuclear genome contigs were ~1 Mb of DNA fragments that had highest similarity to Wolbachia genomes (Koutsovoulos et al. 2014). These fragments bore all the hallmarks of being non-functional horizontal transfers from a once-present Wolbachia. Using these horizontally transferred fragments, the likely source of the transfer was identified as a supergroup F Wolbachia. The D. viviparus data allow resolution of the relationship of supergroup F (and the Wolbachia from R. similis) as sisters to supergroups C and D. The fragments in the D. viviparus genome included remnants of bacteriophage, suggesting that the source genome might have been more like that of the parasites of supergroup A and B than the reduced C and D symbionts. It will be important to survey
other emerging nematode genomes for evidence of past association with Wolbachia, and perhaps other bacteria, and thus reveal the extent of the interactions between these symbionts and nematodes, and perhaps even identify particular associations with parasitism.

A GENOME-BASED TREE OF NEMATODA

One critical issue that molecular phylogenetic analyses now face is that more data are needed. To date, most analyses have used a single locus, the nuclear small subunit ribosomal RNA gene (nSSU). The nSSU is a good gene for deep phylogenetics, but the phylogenetic history that can be extracted from its ~1800 bases is limited. There are now over 8000 nematode nSSU sequences in the public sequence databases (many fragmentary) from over 4000 nominal species. It is not possible to derive a robustly supported tree from this many sequences, and many internal nodes that were unresolved in the earliest analyses remain unresolved in the most recent ones (Holterman et al. 2006; van Megen et al. 2009; Bik et al. 2010), probably because of a lack of unambiguous signal in the single, short nSSU locus. The mitochondrial genome is a readily accessed source of data for phylogenetic inference, and complete genomes are available for over 40 species. These have yielded phylogenies that are well-resolved but at odds with nSSU phylogenies (Park et al. 2011; Sultana et al. 2013). Specifically, neither Spirurina (Blaxter et al. 1998) nor Tylenchina are recovered as monophyletic, and the sister relationship between Heterorhabditis bacteriophora and Strongylomorpha is not recovered (Park et al. 2011; Sultana et al. 2013). Whether these differences arise from biases or errors in the nuclear or mitochondrial data that have not been mitigated remains to be clarified.

One key utility of the new genome-wide data from a wide range of nematode species is that we have a much larger set of data to draw on when compiling matrices for phylogenetic inference (Jones et al. 2011). A major issue is the selection of loci that are orthologous (i.e. where representatives in different species have their origin in a single instance in an ancestral species) and where data coverage is relatively complete. Using gene sets inferred from complete genome sequences, and also complete or high-density transcriptome data, it is possible to infer a set of orthologous genes across the breadth of the phylum. One approach to achieving this is to use a tool such as Core Eukaryotic Genes Mapping Approach (CEGMA) (Parra et al. 2007), which identifies a set of 248 genes known to be present in six model eukaryote species. These genes tend to be highly conserved in sequence, and one limitation resulting from their use may be that there is not enough variation to record closely spaced, or recent branching patterns. An alternative approach is to generate a sequence-based clustering of all genes from all the species under study, using a tool such as orthoMCL (Li et al. 2003), and to query the resultant data for putative sets of orthologues. Using these approaches it is possible to build datasets that include many hundreds of thousands of aligned nucleotides (and several hundred thousand aligned codons or amino acids). These datasets can then be used to address questions left unanswered by the nSSU datasets. First attempts to explore resolution of the deeper phylogeny of Nematoda with multiple nuclear genes derived from whole-genome sequencing projects have largely supported the existing nSSU-derived phylogeny, albeit with limited taxon representation (Desjardins et al. 2013; Laing et al. 2013).

The true power of this phylogenomic approach will only be realized when many hundred nematode genomes representing known diversity are sequenced, but already large datasets can be collated and explored. We have used a combination of whole-genome-derived and transcriptomics-derived gene sets to, firstly, attempt to replicate the findings from the nSSU analyses performed previously, and secondly to resolve some of the remaining unresolved polytomies and conflicts between analyses (Fig. 1B) (Blaxter et al. 2014). Analyses were performed with data from 181 genes from 23 nematode taxa including representatives of the Dorylaimia, Enoplia and Chromadorea. Taxon sampling remained most limited in Enoplia (a single representative, Enoplus brevis), and in the comb-like series of ordinal taxa subtending the Rhabditida in Chromadorea (only Laxus oneistus from Chromadorida). With these taxa the major clades (I–V) that were identified using nSSU (Blaxter et al. 1998) were recovered, the branching order of clades III (Spirurina), IV (Tylenchina) and V (Rhabditina) was resolved as (III, (IV,V)), and the Enoplia were robustly resolved as arising basal to Dorylaimia plus Chromadorea. While E. brevis has a relatively short branch length in these analyses, we caution that its placement may be artefactual due to phylogenetic artefacts elsewhere in the tree, or outgroup problems.

PROSPECTS

The 959 Nematode Genomes initiative notes nearly 100 genomes in progress for the phylum. We have heard anecdotally of many more taxa where researchers are approaching their research goals through genome sequencing, or deep transcriptome sequencing. Improved sequencing technologies such as long single-molecule reads will improve the contiguity of genomes, and improved algorithms will enable assembly even in the presence of high levels of heterozygosity. Careful sampling, and methods for unbiased amplification of genomic DNA from single specimens will fill in the diversity of the tree, and multi-locus phylogenies provide deep resolution
of relationships. The next few years will also see the development of rich collated resources for nematode genomes, including shared genome browsing environments, robust inferences of gene orthology and gene family evolution, and identification of genes and gene families that show particular patterns of evolution associated with distinct clades in the phyllum. In addition, with the development of robust protocols for RNA-based interference in many species, and the development of specific genome editing methods that can be applied to any organism, we expect that questions of the specific roles of many genes to be elucidated. As ever, the questions remain biological: which traits and which genomic features are associated with parasitism, what selective forces maintain them, and how do these change through the ongoing struggle between host and parasite?

ACKNOWLEDGEMENTS

We thank our colleagues in Edinburgh and elsewhere who have been working with us on sequencing, assembly and annotation of nematode genomes, in particular Sujiay Kumar. We thank Andrew Jackson, Matt Berriman and James Cotton for inviting MB to participate in the Wellcome Trust workshop on ‘The evolution of parasite genomes and the origins of parasitism’. We apologize to colleagues whose work we were unable to reference due to limitations of space.

REFERENCES


Blaxter, M. (2007). Systematic roles of many genes to be elucidated. As ever, the questions remain biological: which traits and which genomic features are associated with parasitism, what selective forces maintain them, and how do these change through the ongoing struggle between host and parasite? Parasitology. 10.1017/s0031182013001263.


