Evolutionary Biology of Parasitic Platyhelminths: The Role of Molecular Phylogenetics

D. Blair, A. Campos, M.P. Cummings and J.P. Laclette

As our appreciation of the diversity within the flatworms has grown, so too has our curiosity about the ways in which these varied creatures are related to one another. In particular, the parasitic groups (trematodes, cestodes and monogeneans) have been the focus of enquiry. Until recently, morphology, anatomy and life histories have provided the raw data for building hypotheses on relationships. Now, ultrastructural evidence, and most recently, molecular data from nucleic acid sequences, have been brought to bear on the topic. Here, David Blair, Andrés Campos, Michael Cummings and Juan Pedro Laclette discuss the ways in which molecular data, in particular, are helping us recognize the various lineages of flatworms.

The phylum Platyhelminthes (flatworms) is a large (>15,000 species) and diverse one, and appears to be an early diverging lineage within the Metazoa. Flatworms show great diversity in anatomy, habitat, life history and size, varying from the 20 m long eucestode Diphyllolothrium latum, to species barely visible to the naked eye. They also exhibit considerable genetic diversity among 18S rRNA gene sequences when compared to higher taxa of vertebrates. The phylum includes free-living, commensal, mutualist and parasitic species, of which many are significant for medical and economic reasons. All parasitologists know of the threat to human health posed by trematodes such as members of the genus Schistosoma, or cestodes such as Echinococcus granulosus. Despite its importance, many controversies exist concerning the systematics of the phylum and, in particular, the origin and interrelationships of the major parasitic groups. The evolutionary history of flatworms is still unclear because of the fragmentary information on their morphology, physiology and life cycles, and because these soft-bodied worms have left few fossil remains. A robust phylogeny will also provide insights into the origins of parasitism in the phylum and the shared evolutionary history of parasitic flatworms and their hosts.

In this review, we explore recent findings about phylogenetic relationships in the flatworms, especially the parasitic groups. We emphasize conclusions drawn from DNA sequence data (see Boxes 1,2) and focus on a number of questions that have emerged from the literature. For consistency, we will mostly follow the systematic scheme and names of the major groups provided in Ehlers3-4 (Fig. in Box 1).

Three main questions have emerged from the literature. (1) Do the major groups of parasitic flatworms [i.e. Trematoda (including Digenea and Aspidobothrea), Monogenea and Cestoda (including Gyrocotylidea, Amphilinidea and Eucestoda)] form a single clade? The name Neodermata was recently given to this proposed clade. (2) How are these parasitic groups related to one another? (3) What are the relationships of the major parasitic groups to other flatworm taxa?

Do the major parasitic groups form an exclusive lineage?

Members of the main parasitic groups differ considerably from one another in structure, habits and life cycles. As a result, several proposals to explain their origins assumed that they did not form a monophyletic group. For example, it has been suggested that trematodes, monogeneans and cestodes all originated independently from parasitic rhabdocoelan ancestors5-7, or that the monogeneans, from which descended the cestodes, originated from a free-living rhabdocoelan ancestor and had separate origins from the trematodes8.

Ultraceutural studies were the first to provide evidence contrary to such arguments. All the major parasitic groups resemble one another in structural and developmental aspects of epidermal cilia, sensory receptors, flame bulbs and sperm. In addition, in these taxa the epidermis of the larva is replaced in the adult by a syncytial 'neodermis'5,9, a feature that led Ehlers5 to propose the name Neodermata for the group. Studies based on molecular data from the nuclear 18S rRNA gene also strongly support a single origin for these major parasitic groups5-12.

Relationships among the main groups of Neodermata

Two main clades are recognized within the Neodermata on the basis of ultrastructure and aspects of ontogeny. The Trematoda (Aspidobothrea and Digenea) forms one lineage, and the Cercomorphome (Monogenea and all groups allied with the Cestoda) the other. The Cercomorphome is characterized by common possession of a cercomere, i.e. a posterior attachment structure bearing hooks that is present at some stage in the life cycle.

Monophyletic Trematoda. On the basis of morphological and life cycle characteristics2,3,13-16, the Aspidobothrea (Fig. 1a) has been proposed as the sister group of the digeneans. Molecular studies, using 18S rRNA sequences, have also supported this view5,11. However, the possibility that Aspidobothrea is the sister group to the remaining Neodermata has not been ruled out11,17.
Digenea. In the Digenea (Fig. 1b), most groupings at low levels (species, genus) are widely accepted. Unfortunately, there is still considerable dispute concerning the higher-level (family, order) classification. Information gleaned from morphology (including ultrastructure) and from life cycles has been the mainstay of digenean systematics. In the view of some workers, the amount of homoplasy in the existing data sets is such as to render impossible the accurate reconstruction of the phylogenetic history of the trematodes. Such claims have spurred molecular studies on the group: currently, more molecular data are available for Digenea than for any other flatworm taxon (see Fig. in Box 1).

The entire 18S rRNA gene (1970 bp) was evaluated as a source of phylogenetic information within the Digenea using sequences from eight taxa. Very little phylogenetic information was found in the conserved regions of the molecule, which make up only 20% of its length. Most information came from the variable domains and especially from the V4 domain, which comprises only about 16% of the molecule. Data from this region provide strong support for a number of familial groupings. For example, lepocreadiids and gyliarchaeids are closely related to the exclusion of the paramphistomes despite the similar anatomy of the last two.

Although the V4 region is likely to be the target of many future studies, it (and perhaps the 18S gene as a whole) does not seem to be informative about the earliest divergences among the digeneans. This is unfortunate because there has been lively debate on this topic. Several families have been proposed as the sister group to the rest of the digeneans (reviewed in Ref. 16). This question of the sister family has also been approached by molecular analysis, but with the data available so far, no single family emerges as sister to the remaining Digenea.

The 5' end of the 28S rRNA gene, including the D1 variable domain, has been the subject of a number of studies looking at deep branches among metazoaans (see, for example, Ref. 22). Analysis of the D1 region from 13 digeneans was found to be suitable for inferring relationships among species, genera and closely

Box 1. Gene Regions Used to Infer Relationships

There are several different classes of genes that can provide data about phylogenetic relationships. Nuclear ribosomal RNA genes, which encode structural RNA of ribosomes, tend to occur in eukaryotes as a series of tandemly repeated units, each containing three coding regions (5.8S, 150 bp; 18S, 2000 bp; 28S, 4500 bp) separated by various spacers. The rate of evolution differs both among and within these genes and spacers, providing information that is useful at a range of divergence levels. The internal transcribed spacers, located between the 18S and 28S genes and separated by the 5.8S gene, are the least-conserved regions used in studies, to date, on flatworms. Mainly for historical reasons, the gene encoding the 18S subunit (the most conserved of the ribosomal genes), has been widely used to study relationships of Platyhelminthes (see Fig. below). A short 35 ribosomal RNA gene occurs elsewhere in the genome. Studies using this gene for phylogenetic inference have been reviewed and dismissed.

One useful feature of ribosomal repeat units is that they tend to be ‘homogenised’ within an individual and among individuals within a species or population; a process called ‘concerted evolution’. As a result, intraspecific variation among ribosomal genes is very low compared to interspecific variation. A sequence from one individual should therefore be representative of its species and sample sizes need not be large.

Another source of phylogenetic information is the mitochondrial genome. This clonally inherited circular molecule usually encodes 37 genes, and is typically less than 20000bp long in animals. Mitochondrial genes often accumulate changes at higher rates than the nuclear genome, and therefore data from this genome are expected to be more useful for relatively recent divergences. Among the genes encoded by the mitochondrial genome are those for ribosomal subunits, which are smaller than their nuclear counterparts and are derived from the eubacterial ancestor of the mitochondrial.

Nuclear protein coding genes have rarely been used in systematic studies of Platyhelminthes (see Fig. below).

Phylogeny of the Platyhelminthes according to the widely accepted scheme of Ehlers. Modified after Ref. 9. Numbers in parentheses following each taxon name are numbers of species in each taxon for which partial or complete sequences of the following genes are available: 18S nuclear rRNA (2-29, 31-37, 41-46), 28S nuclear rRNA (25-57, 59-63), nuclear ribosomal internal transcribed spacers (24-26, 31, 33-36), nuclear protein-coding genes (46-67) and mitochondrial genes (24-26, 34-36, 46-66), respectively. Note that many nuclear protein coding genes have been sequenced for members of the Neodermata, but are not enumerated here because they are not available for sufficient numbers of taxa to permit phylogenetic analysis. The urodelids (no sequences available) and the feampulds (1,0,0,0,0) are not placed in the tree (see text). Trematocorina and pterostericorids belong with the Dalyyellidae ribodocel.
Box 2. Analysis of Molecular Data and its Problems

For each taxon, raw sequence data comprise a string of nucleotides (A,C,G,T/U), which are aligned to permit recognition of homologous sites. The alignment process is crucial; if homology is not properly established for each position, incorrect trees can be inferred. Alignment can be done either by ‘eye’ (if the data set contains few taxa), or by using computer programs. However, even computer-based methods frequently produce results which the human eye finds unacceptable and that require adjustment. Alignment is progressively more difficult as phylogenetic distance increases among taxa, and sequences become more dissimilar. Ribosomal sequences often pose particular alignment problems because homologous regions may differ in both length and nucleotide frequencies among taxa. Two recent studies differed concerning relationships between the Monogenea and Gyrocotylidea using similar ribosomal sequences but different alignments. Consideration of putative secondary structural features may help with alignment, but these techniques are relatively untried as yet. Protein encoding genes are often much easier to align because of their organization into codons. Insertions and deletions usually occur in multiples of three, and the inferred amino acid sequences can aid alignment.

Several methods exist for inferring trees, all of which embody implicit and explicit assumptions. Some methods, such as parsimony and maximum likelihood, evaluate the character states (nucleotides present) at each position. Another class of methods requires reduction of the data to a matrix of genetic distances among all pairs of taxa. Any of a number of algorithms (eg, the neighbour-joining method) is then used to infer a tree. (For a major review of these and other methods, see Ref. 52.)

Computer programs will produce a tree from virtually any data set, but often give no indication of how well the data fit the inferred tree. Advances have been made in methods for testing the robustness of phylogenetic trees. It is possible to assess the degree of support for particular groupings using various tests of which the bootstrap is the best known and one of the most controversial.

related families, but at deeper levels relationships were poorly resolved. The internal transcribed spacers are less conserved than the coding regions, and therefore informative at a lower (species/genus) level. For the schistosomes, including the human blood flukes, sequences from these spacers yield phylogenies that are very robust and consistent with conclusions based on anatomical data. Mitochondrial sequences have also been used to infer relationships within Schistosoma and the results are congruent with those found using data from the nuclear ribosomal region.

Monogenea. The Monogenea (Fig. 1c) has long been considered a monophyletic group. The most commonly cited shared features of the group are the possession of two pairs of eyes and three rows of epidermal ciliary bands in the oncomiracidium. However, recent ultrastructural studies of sperm have been persistent in pointing out that no relevant synapomorphies exist for the Monogenea as a whole, suggesting that the group may not be monophyletic. Indeed, the Monogenea, and especially the Monopisthocotylea, differ from other groups of neodermatans in exhibiting a considerable degree of ultrastructural variation. Molecular data also suggest non-monophyly of the Monogenea.

The affinities of problematical taxa that are morphologically distinct can be investigated using DNA sequence data. An example from the Monogenea is the genus Anoploiscus. Members of this genus are found on the external surfaces of fish. They lack hooks on the opisthostomum as adults but otherwise resemble monogeneans (two pairs of eyes and a typical neodermis in the adults). However, Anoploiscus exhibits spermatogenesis unlike other monogeneans even though the mature sperm resembles that seen in monopisthocotylean monogeneans. It also exhibits atypical ultrastructure of the flame bulbs and capillaries. Molecular phylogenetic studies, using 18S rRNA sequences, strongly support the placement of Anoploiscus in the Monopisthocotylea.

Another interesting genus that may be close to the Monogenea is Udonella (Fig. 1d). Members of this genus are found on the external surface of caligid copepods parasitizing marine fish. The genus may be allied with the Monogenea and Cestoda, or may be quite distinct from other neodermatans. No DNA sequences are yet available for this taxon.

Cestoda. Like monogeneans, cestodes (Fig. 1e) have received relatively little attention to date from molecular systematists. Relationships within some genera in the Taeniidae have been examined using nuclear ribosomal and mitochondrial sequences. The only other studies have concentrated on relationships among higher taxa of cestodes: Gyrocotylidea, Amphilimidea and Eucestoda. The Gyrocotylidea emerged as the sister group of the rest of the Cestoda, which is consistent with several recent proposals based on morphological data. The first molecular study to examine the position of Gyrocotyle placed this taxon in the Monogenea, but this may have been an artifact of the alignment used.

What is the sister group to the Neodermata?

Schemes resembling that shown in Box 1, placing rhabdocoel groups as sisters to the Neodermata, have been increasingly accepted in recent years (but see Refs 36–38 for an alternative phylogeny that separates the rhabdocoels and the neodermatans). It is not surprising, therefore, that recent theories on the affinities of the Neodermata have focused on parasitic or commensal rhabdocoels. In particular, the trematocelids, the feccampids and the pterasterocelids are among the taxa that have recently been proposed as sisters to the Neodermata.

The family Fasciidae (Fig. 1f) consists of species parasitic in crustaceans, but which leave their host to spawn inside a secreted cocoon. On ultrastructural grounds, this group initially appeared close to the Neodermata. However, a study using partial 18S rRNA sequences from 19 flatworms found the Fasciidae to be remote from both the Neodermata and the Rhabdocoela, where they are customarily placed. Parallelisms and convergences must have produced the similar ultrastructural features observed in the fasciids and neodermatans. Certainly, the same features used to infer a close relationship between these groups also appear, in part, in totally unrelated species scattered throughout the phylum. A new class, Fasciida, has been proposed for the family.

The trematocelids (Fig. 1g) are quaint creatures with a posterior sucker used to attach to the external
surface of a host (usually a freshwater decapod crustacean), and the anterior of the body bears two or more tentacles (except in Didymorchis). In at least one major systematic scheme for the parasitic flatworms\textsuperscript{2,13}, the trematodephalids were placed as sister group to the Neodermata, but recent anatomical\textsuperscript{39} and molecular\textsuperscript{10–12} studies do not support this view. Molecular data suggest that the trematodephalids are close to the dalyellids and the pterastericoids\textsuperscript{6}. Ultrastructure of protonephridia also suggests that the trematodephalids are close to other rhabdocoels\textsuperscript{37}.

The pterastericoids (Fig. 1h) are common inhabitants of the digestive tract of echinoderms\textsuperscript{6}. On the basis of sperm ultrastructure and other characteristics, it has been proposed that this group, together with the fecampids, could be sister to the Neodermata\textsuperscript{30–42}. However, additional study\textsuperscript{43,44} led to the conclusion that the pterastericoids are not, after all, close to the Neodermata; a result also supported by molecular data\textsuperscript{6}.

**Conclusions and prospects**

When congruent, robust trees are inferred from several different data sets, we can have a considerable amount of faith in the conclusions. The most extensive molecular data sets to date for any flatworm group are from digenetics of the family Schistosomatidae and come from several regions of the nuclear and mitochondrial genomes. Phylogenies inferred have been robust, congruent with one another and congruent with those inferred from other kinds of data. This demonstrates that robust phylogenies can be obtained for flatworms using molecular data.

For deep-level phylogenies of flatworms, the situation is not quite as clear. In 1988, at a time when much ultrastructural data had already been brought to bear on the problem, but before the advent of DNA sequence data, Rohde\textsuperscript{45} could say "... the sister group of the Neodermata cannot be identified at present ...". By 1994, following several studies using 18S rRNA sequences, he was scarcely more definite\textsuperscript{37}: "... the
sister group of the Neodermata [may be] a ... taxon comprised of the Proseriata, Rhabdocoela and Tricladida ... Scenarios of relationships among the rest of the non-parasitic groups are, of course, much more confused.

The 18S data sets used to date have given clear answers (outlined earlier) about some relationships within the flatworms, but have provided only weak suggestions about others. There could be many reasons for this. Some are outlined in Box 2 and others include: lack of data for key taxa and the possibility that all flatworm taxa arose almost simultaneously during an explosive ancient radiation. Additional data, hopefully involving the entire 18S gene, will certainly provide more answers: our knowledge has already been increased greatly thanks to this gene.

Protein-encoding genes, especially from the nuclear genome, are likely to be the focus of future phylogenetic studies on flatworms at all levels. Such genes are far easier to align, even at great phylogenetic depths, than are ribosomal genes. However, care is required in selection of an appropriate candidate gene. Only single-copy genes should be used, otherwise it might be hard to confirm that gene regions involved are strictly orthologous. Many nuclear genes occur as families, members of which may differ among themselves (e.g., eggshell protein genes of the trematodes). Suitable genes for phylogenetic studies on flatworms remain to be identified.

While continuing to expand molecular data sets for flatworms, every effort should be made to enhance the quality and quantity of morphological data sets. Hypotheses generated from one type of data can be tested using the other. Perhaps in this way the one true tree of platyhelminth evolution will eventually be traced!

Acknowledgements

The work was supported in part by grants from Fundación Miguel Aleman and CONACYT 1131-N9202 and 5103N (JPL) and from the Australian Research Council (DB). We wish to thank Klaus Rohde, Nikki Watson and Steve Barker for comments on the manuscript and for allowing us to cite papers in press.

References

8 Llewelyn, J. (1965) in Evolution in Parasites (Taylor, A.E., ed.), pp 47–78, British Society for Parasitology
In Search of the Most Effective Ways of Implementing Insecticide-treated Bednets: TDR Invites Operational Research Proposals

Insecticide-treated (bed)nets (ITNs) have been considered one of the most promising malaria control interventions. Results of large-scale trials evaluating the efficacy of ITNs in reducing morbidity and mortality in well-defined epidemiological settings in Africa will become available early in 1996. Preliminary results from these trials are very encouraging.

In order to match the ongoing field work on efficacy assessment, the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR) and the International Development Research Center of Canada (IDRC) have developed a joint initiative to encourage and support operational research on the most efficient means of promoting, implementing, and sustaining ITN interventions in different settings. A first round of proposals has been approved in September 1995 by the TDR Task Force on bednets. A second round of funding is planned for September 1996. Research groups are invited to submit detailed letters of intent for work that falls within the topics of the initiative.

Further details can be obtained from Dr J. Cattani, Secretary, Bednet Task Force, WHO/TDR, 1211 Geneva 27, Switzerland. Tel: +41 22 791 3737; Fax: +41 22 791 4854. Letters of intent should be sent by 15 May 1996 to the same address.