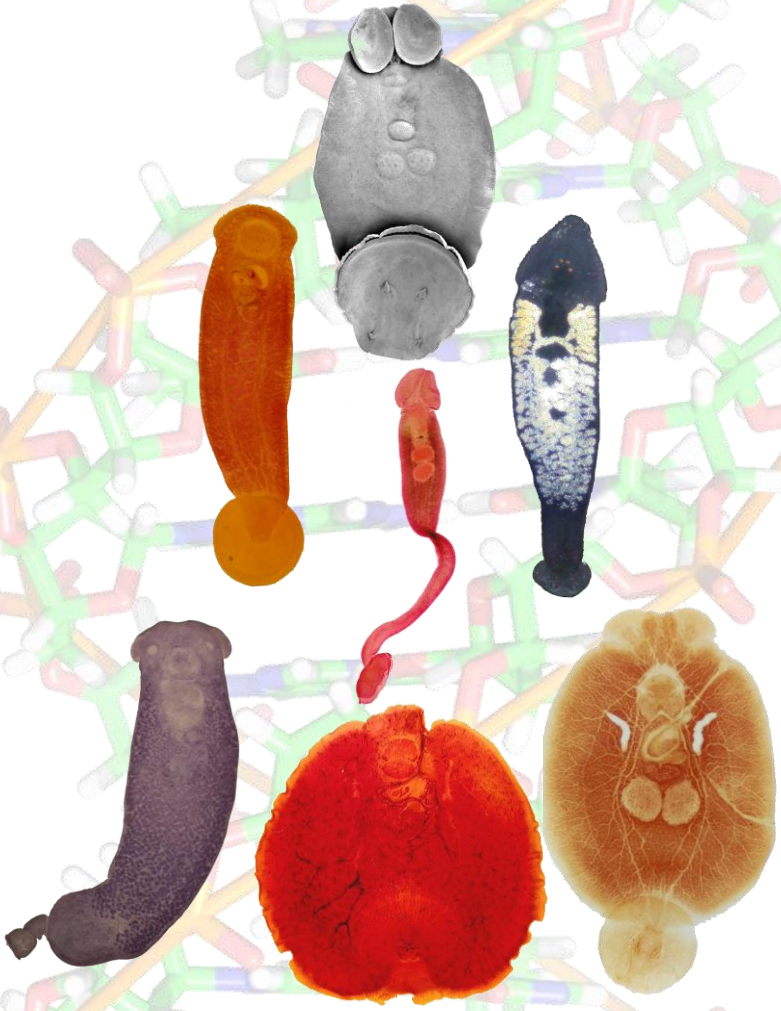
FAMILY TIES: MOLECULAR PHYLOGENETICS, EVOLUTION AND RADIATION OF FLATWORM PARASITES (MONOGENEA: CAPSALIDAE)



ELIZABETH PERKINS

Presented for the degree of Doctor of Philosophy

School of Earth and Environmental Sciences

The University of Adelaide, South Australia

February, 2010

CHAPTER I

GENERAL INTRODUCTION

GENERAL INTRODUCTION

Introduction

Parasitism is one of the most common and successful modes of life displayed amongst living organisms (Poulin and Morand 2000). A parasite can be defined as an organism that lives in close association for a significant period of its life on or in its host from which it derives nutritional or metabolic benefit (Whittington and Chisholm 2003). Parasitism has evolved independently at least 60 times in the animal kingdom and in some instances, it is the parasitic lineages that have diversified far more than their free-living relatives, such as in the Platyhelminthes (Brooks and McLennan 1993). Given that many parasite species still await discovery their true number is likely to be vast. Every free-living organism potentially hosts a parasite at some stage in its life (Whittington and Chisholm 2003), yet there is no parasite that is „universal" and can infect all available host species in an environment. The true diversity of parasites can only, at this stage, be imagined. Parasitic organisms are diverse and problematic (Brooks and McLennan 1993).

Parasitologists have faced many problems in correctly identifying and then inferring the relationships of parasites (Noble *et al.* 1989). Robust phylogenies are the basis for interpreting and understanding biological variation in the light of evolution. Homologous characters are critical in the construction of phylogenies. A character is homologous in two or more organisms if the character is present in their most recent common ancestor, but the character need not look or function alike. In fact, a phylogenetically informative character does not need to be functionally important (Brooks and McLennan 1993). As some idea of relationships between taxa is necessary to determine homology, homologous characters are usually hypothesised by developmental, structural and positional similarity. Such assumptions have posed significant problems in determining truly homologous characters in parasites (Brooks and McLennan 1993).

Some parasites tend to have simplified body plans in comparison to free-living relatives with some consequent reduction in the number of morphological characters (Brooks and McLennan 1993). This is well demonstrated by the highly modified parasitic copepods (Ho 2001), where a reduction in morphological characters, such as loss of body segmentation, makes character analyses especially challenging (Noble *et al.* 1989). Once characters are identified, a decision must be made about homology. Although parasitism has evolved independently on numerous

occasions, all parasites face similar problems in life. A parasite must find its host, attach to it, and then derive nutrition from their host. In general, because all parasites face these common challenges convergence in morphology is a frequent occurrence (Brooks and McLennan 1993). Characters may appear the same in two species but are not, in fact, derived from a common ancestor. These characters are, therefore, not homologous but analogous. Analogous characters do not reflect common ancestry, are not informative phylogenetically and can confound phylogenetic analyses. An example of character convergence is seen in the suckers of flatworm parasites, the monogeneans, cestodes and digeneans. Firm attachment to a host is vital for parasites and represents a strong selection pressure. While the suckers of these parasitic flatworms appear similar in all these groups, they are structurally very different and not derived by common ancestry (Littlewood *et al.* 1999a) despite assertions to the contrary (e.g. Brooks 1989; Brooks and McLennan 1993).

Parasite morphology can also be highly conserved, i.e. shows little variation within a group. Despite similar structures, different parasites may use these structures disparately depending on what host and/or site they attach to. In contrast there can also be significant intraspecific variation. Biological and environmental variables such as parasite and host age, host species and water temperature can also induce changes in some morphological structures (Brooks and McLennan 1993). These changes do not have a genetic basis and are not phylogenetically applicable. Insufficient knowledge about parasite speciation has also contributed to difficulties in the discrimination of parasite species. Due to the extent of problems faced with morphological characters as detailed above, molecular genetics is proving useful in resolving parasite relationships at many different levels in the phylum Platyhelminthes.

The Platyhelminthes

Identifying the basal bilaterian group is extremely important to our understanding of the evolutionary radiation of the major animal phyla (Littlewood *et al.* 2004). The Platyhelminthes was originally believed to be monophyletic and the most basal branch of the Metazoa (see Littlewood *et al.* 2001). Recently the phylum has been found to be paraphyletic and a single clade of free-living flatworms, the acoels, was considered as the most basal extant bilaterian lineage, distinct from other Platyhelminthes (Egger *et al*. 2009). Another study, however, has contradicted the

basal position of the acoels and considers them to be flatworms (Carranza *et al.* 1997). Whether the Platyhelminthes is indeed paraphyletic or monophyletic, the phylum still holds a key position in many theories about metazoan origins (Litvaitis and Rohde 1999; Egger *et al.* 2009). Relationships within the Platyhelminthes, especially the parasitic representatives, have attracted considerable attention.

The Platyhelminthes is a diverse phylum of aquatic and terrestrial organisms (Carranza *et al.* 1997). This phylum is divided into two groups; the „Turbellaria" and the Neodermata (see Kearn 1998). „Turbellaria" is a collective term for platyhelminths with a mostly free-living lifestyle (some symbionts) and traditionally consists of the acoels, rhabdocoels, triclads and polyclads. They are primarily epifaunal or infaunal inhabitants of marine and freshwater benthos but some pelagic and terrestrial forms exist. Defining features of „Turbellaria" are their mostly free-living lifestyle and a body covered in a ciliated epidermis. Neodermata are wholly parasitic and comprise three classes, the tapeworms (Cestoda), internal flukes of most vertebrates (Trematoda) and ectoparasitic flukes of fish (Monogenea).

Some of the most medically and economically important parasites are platyhelminths (Littlewood *et al.* 2004) including schistosomes (blood flukes) and *Echinococcus* (tapeworms causing hydatid disease), both of which can infect humans. Currently no morphological synapomorphy unites the Platyhelminthes. Resolving a stable phylogeny for the phylum has remained difficult due to the limited number of morphological characters and difficulty establishing character homology. Studies focusing on ultrastructural characters have helped resolve some of these problems, though none has resulted in a definitive phylogeny (Justine 1997). Two major points have been shown through ultrastructure: 1) „Turbellaria" may be paraphyletic and the term should be used with caution (hence the quotation marks); 2) three clearly defined clades have been identified: the Acoelomorpha; Catenulida; and the Rhaditophora (including the Neodermata). Again, lack of convincing homology between proposed characters has prevented further relationships from being determined confidently (Justine 1997).

The Neodermata

It is thought that the Neodermata evolved from a free-living rhabdocoel-like ancestor. The Neodermata is considered to be monophyletic with the character „replacement of larval epidermis by a neodermis (new skin) with sunken nuclei"

uniting the group (Baverstock *et al.* 1991; Littlewood *et al.* 1999a). The common ancestor to the Neodermata may have been initially endoparasitic, with only the Monogenea moving towards ectoparasitism, but retaining the neodermis (most parsimonious assumption) (Littlewood *et al.* 1999a). However, molecular phylogenetic analyses using complete mitochondrial genomes suggest that the Neodermata have moved from ectoparasitism to endoparasitism with vertebrate hosts acquired first (Park *et al.* 2007). The ability to infect a vertebrate host is believed to have led to the large number of species in the Neodermata.

The neodermis may play a role in nutrient acquisition through increased surface area from microvilli, microridges and pits and a highly active glycocalyx involved with active nutrient uptake and transport (Littlewood *et al.* 1999a). Other synapomorphies for the Neodermata currently include: electron dense collars of sensory receptors; axonemes of sperm incorporated into sperm body by proximo-distal fusion; protonephridial flame bulbs formed by two cells; incorporation of a vertebrate host in the lifecycle as either a single host (Monogenea; see Whittington 2004), facultative host (some Aspidogastrea; see Rohde 2001) or obligate final host (all others) (Munoz *et al*. 2006). While it is possible that these characters may be coincidental and retained from ancestral forms that adopted parasitism, they are currently considered synapomorphies for the group. Along with studies to resolve higher-level platyhelminth relationships, investigations have also pursued phylogenetic analyses within the major parasitic classes. My project also delves within a major parasitic class by focusing specifically on a family in the Monogenea.

Monogenea

Species of Monogenea primarily infect the external surfaces and gills of freshwater and marine fish (Whittington 2004). Some monogeneans, however, have exploited other aquatic vertebrates such as amphibians, turtles and even the hippopotamus (Whittington 1998). They are as diverse as the other obligate flatworm parasites despite having a single host lifecycle (Littlewood *et al.* 2004). Monogenea also tend to be highly host specific (i.e. some species commonly infect a single host species). The most recent phylogenetic review of this class was based on morphology and included 53 families (Boeger and Kritsky 2001). Ten families were omitted from the analyses of Boeger and Kritsky (2001) due to uncertainties regarding origins and validity. The Monogenea is supported by several synapomorphies including: larvae

and adults with two pairs of eye spots; three bands of ciliary patches and tapering epidermal cilia; reduced number of microtubules in apical parts of sperm; and similarity in gross protonephridial morphology in some species (Littlewood *et al.* 1999b). When fish are kept under stressed and crowded conditions, such as aquaria and sea cages, the host specificity of Monogenea can break down (Thoney and Hargis 1991). Monogeneans from several higher taxa have been implicated in causing disease and mortality in intensive aquaculture (Whittington and Chisholm 2008). There are few cases of monogeneans causing disease in natural host populations.

Morphological phylogenies tend to suggest monophyly for the Monogenea (e.g. Boeger and Kritsky 1993, 2001) while molecular phylogenies tend to suggest paraphyly (e.g. Mollaret *et al.* 1997; Olson and Littlewood 2002). Phylogenies based on sperm morphology also challenge monophyly of the group (Justine *et al.* 1985; Justine 1991). In molecular analyses, paraphyly may be an artefact of gene choice and hypotheses based on different or more genes may support monophyly (Lockyer *et al.* 2003; Littlewood *et al.* 2004). As molecular analyses have been unable to show paraphyly consistently, monophyly is still widely accepted for the Monogenea. Whether Monogenea is ultimately found to be monophyletic or paraphyletic, it seems that members radiated very rapidly from their ancestral stock (Littlewood *et al.* 2004). Assuming monophyly, the Monogenea is divided into two subclasses, the Monopisthocotylea and Polyopisthocotylea, though debate surrounds this nomenclature (Boeger and Kritsky 2001). It is primarily the posterior attachment organ (haptor) and diet that delineates the two subclasses. The epithelial feeding adult Monopisthocotylea have hooks and hooklets on their haptor whereas the haptor of the blood feeding adult Polyopisthocotylea is characterised by clamps (Boeger and Kritsky 2001).

A solution to the individual problems of morphological and molecular analysis is a total-evidence approach (Littlewood *et al.* 1999b; Olson and Littlewood 2002) where molecules and morphology are used in conjunction with each other to recover phylogenetic hypotheses. This can be done by either analysing each data set separately and in some manner constructing a consensus view of the resulting trees, or by combining the data in a single analysis that overcomes issues of hidden branch support not apparent in the separate analyses (Littlewood *et al.* 1998). This reduces the effects of bias and produces more robust hypotheses, perhaps less influenced by *a*

priori assumptions. Molecular analyses have only been used relatively recently in parasite phylogenetics. Currently there is a limit to the number and suitability of genes available and choice of genes is often conservative, limited to ribosomal RNAs and *Cytochrome Oxidase 1*. In the future, when many more genes have been assessed, morphological characters may be used more valuably by mapping them onto molecular hypotheses to examine character evolution and to delineate taxonomically diagnostic character states. My project focuses on the Capsalidae; its evolution and radiation and position within the Monogenea.

Capsalidae

Capsalids are ectoparasites of marine fish and some are important pathogens of fish in aquaculture and aquaria. According to Whittington (2004) the Capsalidae (Monogenea: Monopisthocotylea) comprises nine subfamilies (Figure 1), approximately 200 species in 48 genera, but the number of subfamilies has varied. The family is exceptional because while most species generally parasitise „modern" teleosts, representatives from five genera can also infect sharks and rays and species in one genus infect acipenserids (Whittington 2004). The general morphology of capsalids is conserved. They range in size from 1 mm to 3 cm and at 2-3 cm long, *Capsala martinierei* and *Entobdella hippoglossi* are among the largest monogeneans known.

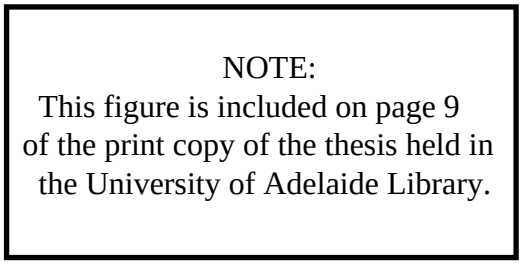


Figure 1. Diagrammatic representation (not to scale) of the nine capsalid subfamilies: **A.** Capsalinae; **B.** Benedeniinae; **C.** Dioncinae; **D**. Encotyllabinae; **E**. Entobdellinae; **F**. Interniloculinae; **G**. Nitzschiinae; **H**. Pseudonitzschiinae; **I**. Trochopodinae. (From Deveney 2002)

In general capsalids have a leaf-like body (e.g. Figure 1A). Encotyllabines are a notable exception where the body edges fold ventrally to create a tube-like body that terminates posteriorly in a bell-shaped haptor (Figure 1D), at the end of a muscular peduncle (Kearn and Whittington 1992). The nine subfamilies are characterised by different combinations of haptor morphology, anterior attachment organ morphology and testis number and arrangement (two or multiple, Whittington 2004). The haptor morphology of capsalids is conserved and may be subject to some convergence across the family. In general it is saucer-shaped (Figure 2A) with three pairs of median sclerites comprising a central pair of accessory sclerites and two pairs of ventrally-directed hamuli (Figure 2B). Small hooklets at the periphery and a

thin marginal valve (Figure 2B) are critical to maintain suction. Although haptor morphology is conserved, capsalids can still parasitise a diverse range of sites including: epithelium-covered lamina of teleost scales; smooth ventral epithelium of batoids; gill lamellae, arches and rakers; fins; branchiostegal membranes; lip folds and pharyngeal tooth pads (Whittington 2004).

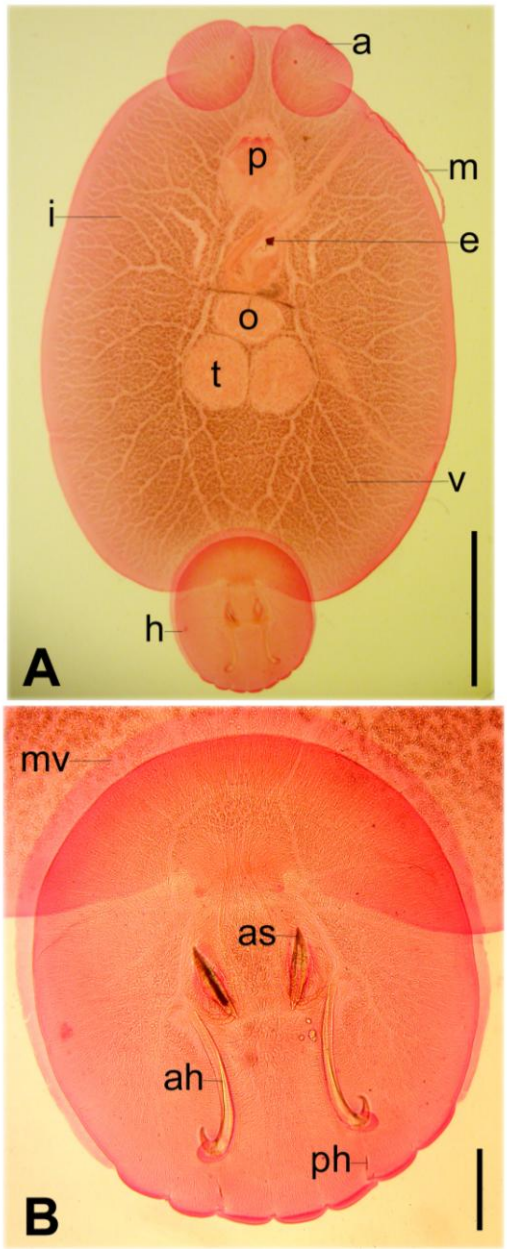


Figure 2: *Benedeniella posterocolpa* (Capsalidae: currently in Benenedeniinae, but my analyses indicate it is a member of the Entobdellinae; see Perkins *et al*. 2009) from ventral skin of the cownose ray, *Rhinoptera bonasus* (Myliobatidae) from the New York Aquarium (originally from Virginia, USA). **A**. Whole parasite, ventral view, showing paired anterior attachment organs (a), egg (e) in the ootype, posterior haptor (h), intestine (i), ovary (o), everted male copulatory organ (m), pharynx (p), testis (t) and vitellarium (v). **B**. Enlargement of haptor, the principal attachment organ, showing the three pairs of median sclerites (anterior hamuli, ah; accessory sclerites, as; posterior hamuli, ph) and the thin, flexible marginal valve (mv). There are also 14 peripheral hooklets (approx. 15 µm long) which are not clearly visible in this image. Scale bars: **A**, 2 mm; **B**, 400 µm.

Despite approximately 230 years of study, the classification, systematics and biology (for most species) of the Capsalidae remain unresolved. The current capsalid classification appears phenetic and is not based explicitly on cladistic principles (Whittington 2004). Such classifications can include arbitrary groups based on subjective opinion. Monophyly of the Capsalidae is currently supported morphologically by the presence of accessory sclerites (possibly modified hooklets according to Kearn 1963) on the haptor, providing a synapomorphy for the family (Whittington *et al.* 2004). This character is only absent in two capsalid species and it is thought that studies of larvae will show these characters to be present and indicate that they are secondary losses in adults (Whittington 2004).

A preliminary phylogenetic study of the family using nucleotide sequences of the *28S rRNA* gene included 17 species in seven genera and five of the nine subfamilies (Whittington *et al.* 2004). Hypotheses from this study showed the Benedeniinae to be paraphyletic. This study reinforced some interesting relationships about the evolution of the family. In particular, members of the Entobdellinae parasitise elasmobranchs and teleosts and the phylogenetic hypothesis proposed by Whittington *et al.* (2004) suggested that capsalids evolved on teleost hosts and switched to elasmobranch hosts recently. Boeger and Kritsky (1997) also suggested that capsalids had evolved on „modern" teleosts and secondarily dispersed to sturgeons, sharks and rays. A more comprehensive phylogeny is required using an increased number of representatives from genera and subfamilies to draw further conclusions.

Molecular Phylogenetic Techniques

Multi-locus phylogenetic analyses

Molecular phylogenetic analyses of parasitic groups typically use a single gene or a combination of linked ribosomal genes (e.g. *28S ribosomal RNA* and *18S ribosomal RNA*) (Campos *et al.* 1998; Cable *et al.* 1999). Single genes have limitations with analyses of one gene reflecting the gene tree and not necessarily the species tree (Maddison 1997). Combining multiple unlinked genes in analyses is an important step forward in constructing robust phylogenetic hypotheses. Multi-locus analyses have inherent difficulties. Combining data can overlook conflict between genes whereas separate analyses may not show underlying congruent signals

(Dolman and Hugall 2008). This can be overcome through various hypothesis testing methods and incongruence tests (Lee and Hugall 2003). While there are significant amounts of ribosomal data for many parasitic groups readily available on GenBank, there are limitations to these data. Ribosomal RNA genes are linked and present in multiple copies in the genome which can introduce problems of paralogy. A shift towards developing new, informative genes for phylogenetic analyses is needed.

With the second generation of sequencing well under way and now the third generation soon to be embraced, the ability to produce vast amounts of nuclear data for phylogenetic analyses is becoming more and more achievable (Meyer *et al.* 2007; Rusk 2009). In third generation sequencing, costs to obtain a complete nuclear genome may be as little as $1000. Access to such vast amounts of data will provide many informative genes for phylogenetic analyses that would have once required extensive work. These advances are without doubt the way forward for molecular phylogenetic analyses.

The mitochondrial genome

The mitochondrial (mt) genome presents a genome that is small enough in size that it can be readily sequenced using current technology but also large enough to provide a useful amount of informative data. The mt genomes of parasitic platyhelminths are similar to other metazoan mt genomes in gene composition, tRNA and rRNA structure but can be characterised by lacking ATP8 and having a high AT content (Le *et al.* 2002a). They share the same genetic code as the Echinodermata, apparently through convergent evolution, with ATG as the typical start codon and TAG and TAA acting as stop codons (Telford *et al*. 2000). Many of the protein coding genes are separated by short non-coding regions and genomes typically have two larger non-coding regions believed to be associated with genome replication (e.g. Le *et al.* 2002a). The majority of published full mt genomes are from economically or socially important species such as *Schistosoma* and *Echinoccocus* (see Le *et al.* 2001). There are currently 29 complete mt genomes available on GenBank for the Neodermata.

Full mitochondrial genomes have been used to examine relationships at the species level and also higher level relationships (Le *et al.* 2002b [parasites]; Simmons and Miya 2004 [fish]; Yamanoue *et al.* 2009 [fish]). It is not just the sequence data of a mt genome that can be used in phylogenetic analyses but the

arrangement of genes within the genome may also be informative. Gene order rearrangements in theory occur rarely and so when they are shared, it should indicate common ancestry (Littlewood *et al.* 2006). However some studies have shown that in parasitic lineages rearrangements may occur more frequently and should be viewed with caution as phylogenetic markers (Le *et al.* 2000; Dowton *et al.* 2009). Only four mt genomes have been sequenced for monogenean species: three *Gyrodactylus* spp. (Monopisthocotylea; see Huyse *et al.* 2007; Plaisance *et al.* 2007; Huyse *et al.* 2008) and *Microcotyle sebastis* (Polyopisthocotylea; see Park *et al.* 2007). Sequences of more monogenean mt genomes are required to assess the phylogenetic utility of rearrangements.

Coevolution and radiation

As a parasite spends much of its life in tight association with its host, it is thought that the evolution of the host will play a significant role in the radiation of the parasite (Banks *et al.* 2006). Coevolution between a parasite and host occurs when the parasite speciates following a host speciation event and is apparent when a parasite and host phylogeny appear congruent. This is known as Fahrenholz"s Rule: where parasite phylogeny should mirror host phylogeny (Fahrenholz 1913). This strict congruence has been demonstrated in some parasite-host associations such as pocket gophers and their lice (Light and Hafner 2008) but the majority of studies show that coevolution may be the exception rather than the rule (Paterson and Poulin 1999; Weckstein 2004). This cornerstone of coevolutionary studies is fast becoming Fahrenholz"s fallacy (Page and Charleston 1998). Demonstrating coevolution is difficult for many reasons. Coevolution analyses can only be as robust as the parasite and host phylogenies on which they are performed. Often it is not only the parasite phylogeny that needs estimating but confusion about the host relationships can lead to the need to generate phylogenetic hypotheses for the hosts as well. Coevolution not only requires topological congruence but also temporal congruence (Page 1996). In order to demonstrate temporal congruence, some kind of dating method is required for the host and parasite phylogenies. Molecular dating has been developed and there are now programs available that can implement strict and relaxed clock models (Drummond and Rambaut 2007). Critical to accurate molecular dating are multiple fossil calibration points (Hedges and Kumar 2004). For vertebrate hosts like fish, an extensive fossil record exists allowing robust dating for molecular phylogenetic

hypotheses (Azuma *et al.* 2008). However for parasitic flatworms a fossil record is exceptionally rare. The few known fossil parasitic flatworms can not be viewed as either a maximum or minimum age for these groups but only indicate the presence of these groups at that time (Combes 2001). There have been some molecular clock analyses of early metazoans and calibration points do exist for some of these groups (Peterson *et al.* 2004, 2008). Such data can be combined with phylogenetic data of parasites to infer dating for the parasitic groups.

In the absence of coevolution, a parasite phylogeny can be a result of a variety of events such as extinction, „missing the boat", duplication, failure to speciate and host switching (de Vienne *et al.* 2007). Distinguishing between these events is difficult with host switching the most commonly assumed cause. There can be many different drivers of host switching such as shared ecology, biology, behaviour and plasticity in morphological adaptations. To assess correlation between ecological factors and a parasite phylogeny, ancestral state reconstructions can be used to reconstruct the evolutionary history of an ecological trait across a parasite lineage (Pagel 1994). The combination of these analyses techniques allows an assessment of the timing and drivers behind diversification (Pagel 1997).

PhyloCode

A new classification system, PhyloCode, has been in development for the past few years, prompted by recognition that the current Linnaean rank-based system of nomenclature is not well suited to govern the naming of clades and species (Cantino and de Queiroz 2007). PhyloCode will provide rules for the direct purpose of naming clades and species with specific reference to phylogeny. It is designed to be used concurrently with the current rank-based system or as the only code governing the names of taxa if the scientific community so decides. Its intention is not to replace existing names but to provide a system governing the application of existing and new names. Names that apply to clades will be redefined in terms of phylogenetic relationships instead of taxonomic rank. This will prevent names being subject to the same changes that occur under the rank-based system when changes in rank occur (Cantino and de Queiroz 2007). The PhyloCode has been proposed as a means for governing nomenclature in a phylogenetic context (Cantino and de Queiroz 2007). A major criticism of PhyloCode has been a failure to develop a means to deal with species ranks. However, Dayrat *et al.* (2008) proposed a system where Linnaean

binomials can be used in a form that is consistent with phylogenetic nomenclature. A system that can accommodate the legacy of the use of Linnaean ranks and the principles of phylogenetic nomenclature based on molecular phylogenies is perhaps the way forward. Such a system may allow a classification that conveys biological data and the phylogenetic history of the organisms.

Aims

My study aims to provide insights into the phylogenetic relationships and evolutionary history of capsalid parasites using molecular phylogenetic approaches.

- I will use multiple nuclear loci to reveal relationships amongst the Capsalidae and examine its position within the Monogenea (Chapter II)
- Compare phylogenetic hypotheses to the current morphological classification of the family to assess homoplasy of key morphological characters (Chapter II)
- Use full mitochondrial genomes to assess monophyly of Monogenea and the evolution of diet across the Neodermata (Chapter III)
- Combine nuclear and mitochondrial genes across a broader representation of taxa to reassess relationships within the Capsalidae and its position within the Monogenea (Chapter IV)
- Use molecular dating techniques to provide dates for the radiation of the parasitic platyhelminths, Monogenea and the Capsalidae (Chapter IV)
- Use multiple nuclear and mitochondrial genes to generate a phylogeny for the fish hosts of the Capsalidae (Chapter V)
- Use molecular dating techniques to provide dates for the radiation of the major fish groups (Chapter V)
- Compare host and parasite phylogenies and chronograms to assess coevolution (Chapter V)

Looks can deceive: Molecular phylogeny of a family of flatworm ectoparasites (Monogenea: Capsalidae) does not reflect current morphological classification

Elizabeth M. Perkins [a,*], Steve C. Donnellan [b,c], Terry Bertozzi [b], Leslie A. Chisholm [a], Ian D. Whittington [a,d]

[a]*Marine Parasitology Laboratory, School of Earth and Environmental Sciences (DX 650 418), The University of Adelaide, North Terrace, Adelaide, SA 5005, Australia*

[b]*Evolutionary Biology Unit, The South Australian Museum, North Terrace, Adelaide, SA 5000, Australia*

[c]*Australian Centre for Evolutionary Biology and Biodiversity, The University of Adelaide, North Terrace, Adelaide, SA 5005, Australia*

[d]*Monogenean Research Laboratory, Parasitology Section, The South Australian Museum, North Terrace, Adelaide, SA 5000, Australia*

Molecular Phylogenetics and Evolution (2009) **52,** 705–714.

Statement of Authorship

This chapter is a published research article and is reproduced with kind permission from Elsevier Inc. (see Appendix I).

Perkins EM, Donnellan SC, Bertozzi T, Chisholm LA, Whittington ID (2009) Looks can deceive: Molecular phylogeny of a family of flatworm ectoparasites (Monogenea: Capsalidae) does not reflect current morphological classification. *Molecular Phylogenetics and Evolution* **52**, 705–714. (doi:10.1016/j.ympev.2009.05.008)

E.M. Perkins (Candidate)

Corresponding author: Responsible for laboratory work, drafted manuscript, conducted all analyses, produced all figures and oversaw manuscript revisions.

Signed …  …… Date…….. 28/10/2009

S.C. Donnellan

Sought and won funding, co-supervised the direction of study, assisted with analyses and contributed to the manuscript.

I give consent for E.M. Perkins to include this paper for examination towards the degree of Doctor of Philosophy.

Signed .  Date…….. 28/10/2009

T. Bertozzi

Provided technical laboratory assistance, advised and assisted with analyses and evaluated the manuscript.

I give consent for E.M. Perkins to include this paper for examination towards the degree of Doctor of Philosophy.

Signed ..  ….. Date …….. 28/10/2009

L.A. Chisholm

Contributed to the manuscript and assisted with manuscript revision.

I give consent for E.M. Perkins to include this paper for examination towards the degree of Doctor of Philosophy.

Signed .. .. Date..... 28/10/2009

I.D. Whittington

Sought and won funding, co-supervised the direction of study, acquired specimens, contributed to the manuscript and assisted with manuscript revision.

I give consent for E.M. Perkins to include this paper for examination towards the degree o

Signed .. Date..... 28/10/2009

ABSTRACT
The morphological based taxonomy of highly derived parasite groups is likely to poorly reflect their evolutionary relationships. The taxonomy of the monogenean family Capsalidae, which comprises approximately 180 species of flatworm parasites that predominantly attach to external surfaces of chondrichthyan and teleost fishes, is based mainly on six morphological characters. The phylogenetic history of the family is largely unknown. We reconstructed the phylogenetic relationships of 47 species in 20 genera from eight of the nine subfamilies, from nucleotide sequences of three unlinked nuclear genes, *28S ribosomal RNA*, *Histone 3* and *Elongation Factor 1 α*. Our phylogeny was well corroborated, with 75% of branches receiving strong support from both Bayesian posterior probabilities and maximum likelihood bootstrap proportions and all nodes showed positive partitioned likelihood support for each of the three genes. We found that the family was monophyletic, with the Gyrodactylidae and Udonellidae forming the sister group. The Capsalinae was monophyletic, however, our data do not support monophyly for the Benedeniinae, Entobdellinae and Trochopodinae. Monophyly was supported for *Capsala*, *Entobdella*, *Listrocephalos*, *Neobenedenia* and *Tristoma,* but *Benedenia* and *Neoentobdella* were polyphyletic. Comparisons of the distribution of character states for the small number of morphological characters on the molecular phylogeny show a high frequency of apparent homoplasy. Consequently the current morphological classification shows little correspondence with the phylogenetic relationships within the family.

1. Introduction

The Platyhelminthes is a diverse phylum of aquatic and terrestrial organisms that are classified into mostly free-living „turbellarians" and the wholly parasitic Neodermata (see Kearn, 1998). The Neodermata comprises three classes, the Cestoda (tapeworms), Trematoda (internal flukes) and Monogenea (principally ectoparasitic flukes of teleosts and chondricthyans). Monogenea have a direct life cycle and tend to be highly host specific, i.e. species commonly infect a single host species. The Monogenea is divided into two subclasses, the Monopisthocotylea that feed on epithelial cells and the Polyopisthocotylea that are exclusively blood feeders.

The Capsalidae (Monopisthocotylea) include parasitic flatworms that attach predominantly to external surfaces of marine fish. Capsalids are distributed

worldwide and some are among the largest monogenean species known (up to 3 cm long) (Whittington, 2004). Some can be site specific and different species parasitise different sites including the: epithelium covered lamina of teleost scales; smooth external ventral epithelium of batoids; gill lamellae, arches and rakers; fins; branchiostegal membranes; lip folds and pharyngeal tooth pads (Whittington, 2004). While capsalids generally parasitise „modern" marine teleosts, some parasitise „primitive" anadromous and freshwater teleosts, like acipenserids and also marine elasmobranchs (sharks and rays) (Whittington, 2004). Some capsalids are important pathogens in aquaculture and public aquaria e.g. *Benedenia seriolae*, *Neobenedenia „melleni"* and have been responsible for significant losses of fish stocks (Deveney et al., 2001). The current taxonomic classification, which comprises nine subfamilies, 45 genera and approximately 180 species (Whittington, 2004, Table 1), is based on very few morphological characters (e.g. attachment organ characteristics, testis number and arrangement). Within the Capsalidae, some subfamilies and genera are considered ill-defined and require taxonomic revision (Whittington et al., 2004). Four subfamilies contain only a single genus and many capsalid genera are monotypic.

Whittington et al. (2004) conducted a preliminary phylogenetic study of the Capsalidae which used partial *28S ribosomal DNA* (*28S rDNA*) nucleotide sequences, and included only 17 species, representing seven genera and five of the nine subfamilies. Monophyly for the Capsalidae was supported as was monophyly for the Encotyllabinae and Entobdellinae. Benedeniinae was paraphyletic with *Neobenedenia* species failing to fall within the subfamily. *Capsala* was not monophyletic due to the inclusion of *Tristoma integrum*. While this is the only phylogenetic analysis of the family to date, it emphasises the need to establish phylogenetic relationships to assess the substance of the current systematic classification. Far greater taxon sampling and use of multiple genes will be required to infer and resolve relationships within the Capsalidae robustly (Whittington, 2004).

Other than the preliminary phylogenetic hypothesis by Whittington et al. (2004), phylogenetic relationships among capsalids remain unexplored. Currently there are too few morphological characters adequate to establish evolutionary relationships for the entire group. The paucity of phylogenetically useful morphological characters is due largely to the fact that parasites tend to have simplified and conserved body plans compared to free-living relatives (Brooks and McLennan, 1993). Homology is another critical consideration when establishing a

morphological dataset for phylogenetic analyses. If relationships between taxa are unknown, homology is usually inferred by developmental, structural and positional similarity (Brooks and McLennan, 1993). Such an approach can be problematic in relation to parasites and may lead to inaccurate assumptions about homology, an issue of concern for capsalid morphological characters (Whittington, 2004). A molecular phylogenetic hypothesis will allow an examination of the issue of homology in these key morphological characters and an assessment of the frequency and the potential impacts of homoplasy.

Our study extends the preliminary work of Whittington et al. (2004) by increasing taxon and gene sampling. We base our analyses on 47 capsalid species in 20 genera representing eight of the nine subfamilies and also include 15 outgroup taxa (in nine families) from the Monopisthocotylea and Polyopisthocotylea. Presently the sister taxon of the Capsalidae is unknown. Our analyses combine partial sequence data for *28S rDNA*, *Histone 3* (*H3*) and *Elongation Factor 1* α (*EF1*α) and is the first molecular phylogeny of a monogenean family to include multiple unlinked nuclear markers. Six morphological characters commonly used in higher level capsalid classifications were assessed relative to the molecular phylogenetic hypothesis for their utility as phylogenetically informative characters.

2. Materials and methods

2.1. Sample collection

Specimens (preserved in 95% AR grade ethanol) were collected or obtained from various sources between 1993 and 2007 from 47 capsalid and 15 outgroup taxa (see Appendix III). Table 1 shows the current taxonomic classification of the capsalids. Trees were rooted with *Microcotyloides incisa* (Polyopisthocotylea: Microcotylidae), the most distant outgroup included in the analyses. The other 14 outgroup taxa belong to the subclass Monopisthocotylea and represent eight families (Acanthocotylidae, Amphibdellatidae, Calceostomatidae, Dactylogyridae, Gyrodactylidae, Microbothriidae, Monocotylidae and Udonellidae).

Table 1

Current capsalid subfamilies and included genera, listed alphabetically.

Subfamilies*	Included genera**
Benedeniinae (**13**)	*Allometabenedeniella* (1), *Ancyrocotyle* (2), [b]*Benedenia* (21), ***Benedeniella*** (2), *Calicobenedenia* (1), *Dioncopseudobenedenia* (1), *Lagenivaginopseudobenedenia* (2), *Menziesia* (5), *Metabenedeniella* (2), *Neobenedenia* (6), *Oligoncobenedenia* (1), *Pseudallobenedenia* (2), ***Trimusculotrema*** (5)
[a]Capsalinae (**4**)	[b]*Capsala* (22), *Capsaloides* (7), *Nasicola* (3), *Tristoma* (4)
Dioncinae (**1**)	[b]*Dioncus*[c] (11)
Encotyllabinae (**2**)	*Alloencotyllabe* (1), [b]*Encotyllabe* (17)
Entobdellinae (**5**)	*Branchobdella* (1), [b]*Entobdella* (7), ***Listrocephalos*** (4), ***Neoentobdella*** (10), ***Pseudoentobdella*** (1)
Interniloculinae (**1**)	[b]*Interniloculus* (2)
Nitzschiinae (**1**)	[b]*Nitzschia* (2)
Pseudonitzschiinae (**1**)	[b]*Pseudonitzschia* (1)
Trochopodinae (**17**)	*Allobenedenia* (8), *Allomegalocotyla* (2), ***Macrophyllida*** (1), *Mediavagina* (2), *Megalobenedenia* (2), *Megalocotyle* (6), *Pseudobenedenia* (3), *Pseudobenedeniella* (1), *Pseudobenedenoides* (2), *Pseudomegalocotyla* (1), *Sessilorbis* (1), ***Sprostonia*** (2?)[d], *Sprostoniella* (3), *Tetrasepta* (1), *Trilobiodiscus* (1), *Trochopella* (1), [b]*Trochopus* (15)

*Number of genera in bold; **Approximate number of species in parentheses; genera in bold denotes those with species that parasitise elasmobranchs. [a]Subfamily contains type species (*Capsala martinierei*) for the Capsalidae; [b]Type genus for each subfamily; [c]*Dioncus* postoncomiracidia are reported from skin of blacktip sharks (*Carcharhinus limbatus*) (Carcharhinidae), adult specimens of *Dioncus* occur on teleosts of the families Carangidae, Echeneidae and Rachycentridae (see Bullard et al., 2000); [d] host associations in *Sprostonia* require re-evaluation because according to Egorova (1994), the host of the type species, *S. squatinae*, is the angel shark *Squatina squatina* (Squatinidae) but the host of *S. longiphallus* is the teleost, *Epinephelus tauvina* (Serranidae). Table based on Whittington (2004) and updated from Tingbao et al. (2004), Chisholm and Whittington (2007), Kearn et al. (2007) and Whittington and Kearn (2009)

2.2. DNA preparation, PCR amplification and sequencing

DNA was extracted according to the Gentra Kit (Gentra Systems) protocol for animal tissues preserved in ethanol. Extracted DNA was stored in hydration solution at 4 °C. PCR amplification of partial *28S rDNA*, *H3* and *EF1α* sequence was carried out with published primers and additional primers designed using OLIGO 4.0 (Rychlik, 1992) listed in Table 2. For amplification of the *28S rDNA* dataset, primer combinations used were C1/D2 (approx. 800 bp), LSU5/EC-D2 (approx. 800 bp) and G904/G905 (approx. 400 bp). For amplification of the *H3* dataset, primer combinations used were H3aF/H3R2 (approx. 350 bp) and G926/G927 (approx. 300 bp). For amplification of the *EF1α* dataset, primer combinations used were G959/G960 (approx. 800 bp) and G1050/G1051 (approx. 800 bp). Primers used for PCR were also used for sequencing. PCR amplifications were performed in 25 μL reactions using the following cycle conditions: denaturation at 94 °C for 45 s, annealing at a minimum 50 °C and maximum 65 °C (dependent on primers being used) for 45 s and extension at 72 °C for 1 min; this was repeated for 34 cycles and increased to 38–40 cycles when PCR product yield was low. Each 25 μL PCR contained a final concentration of: 0.5 U AmpliTaq Gold® (5 U/μl), 0.2 μM of each primer, 200 μM of each dNTPs, 2–4 μM $MgCl_2$, 1 x AmpliTaq Gold® buffer. Annealing temperature and $MgCl_2$ concentration were varied to produce optimal amplification.

PCR products were cleaned using Agencourt® AMPure® PCR purification kit and were cycle sequenced using the BigDye Terminator v3.1 cycle-sequencing kit (Applied Biosystems). The cycling protocol consisted of 25 cycles of denaturation at 96 °C for 30 s, annealing at 50 °C for 15 s, and extension at 60 °C for 4 min. All samples were sequenced on an Applied Biosystems 3730 DNA sequencer.

Table 2

Primers used for PCR amplification

Gene	Primer ID	Sequence (5‟-3‟)	Forward/ Reverse	Source
28S rRNA	C1	ACCCGCTGAATTTAAGCAT	F	a
	D2	TGGTCCGTGTTTCAAGAC	R	a
	LSU5	TAGGTCGACCCGCTGAAYTTAAGCA	F	b
	EC-D2	CCTTGGTCCGTGTTTCAAGACGGG	R	b
	G904	GATTCTCYTAGTAACKGCGAGTG	F	c
	G905	GTTTAACCTYCAWGTRGTTTCA	R	c
H3	H3aF	ATGGCTCGTACCAAGCAGACVGC	F	d
	H3R2	ATRTCCTTGGGCATGATTGTTAC	R	d
	G926	GACCGCYCGYAAAAGYAC	F	c
	G927	AGCRTGRATDGCRCACAA	R	c
EF1α	G959	GATTTYATTAARAAYATGATYACTGG	F	c
	G960	CRGGATGRTTCATAAYRATAAC	R	c
	G1050	CTGGWACYAGYCARGCTGA	F	c
	G1051	CATACCATACCACGYTTKA	R	c

[a]Chisholm et al. (2001).

[b]Littlewood et al. (1997).

[c]This study.

[d]Colgan et al. (1998).

2.3. Phylogenetic analyses and hypothesis testing

Sequence chromatograms were edited using SeqEd version 1.0.3 and aligned initially using Clustal X (Thompson et al., 1997). Adjustments to alignments were made manually in SeAl version 2.0a11 (Rambaut, 1996) using inferred amino acid sequences where applicable (*H3* and *EF1α*). For the *28S rDNA* sequence data, we tried to align our sequences to the predicted RNA structure for *Gyrodactylus salaris* (see Matejusová and Cunningham, 2004). All sequences have been deposited on GenBank (Accession Nos. FJ971962–FJ972138). Voucher specimens (most mounted on slides but some are specimens or part specimens stored in 95% AR grade ethanol) of each monogenean species are deposited in the Australian Helminthological Collection (AHC) of the South Australian Museum (SAMA), Parasitology Section, North Terrace, Adelaide, South Australia 5000, Australia or in the Muséum National d‟ Histoire Naturelle (MNHN), Paris, France.

Monte Carlo Markov Chain (MCMC) Bayesian phylogenetic analyses were run using MrBayes 3.1.1 (Huelsenbeck and Ronquist, 2001). This analysis method allowed the data to be partitioned and optimal models of nucleotide substitution

applied to each partition. The model of nucleotide substitution for each partition was assessed using the Akaike Information Criteria (AIC – Akaike, 1985) in ModelTest version 3.7 (Posada and Crandall, 1998). The General Time Reversible (GTR) model with a proportion of invariable sites and a gamma distribution for rates across sites was selected. To determine an optimal partitioning strategy, preliminary Bayesian analyses (1 million generations) using each possible partitioning strategy were run and then the AIC for each partitioning strategy calculated. The final MCMC analyses were run for 10,000,000 generations with a sample frequency of every 100 generations. Tracer v1.4 (Rambaut and Drummond, 2007) was used (to plot the generation number against the log likelihood value) to identify the point at which log likelihood values became stable and all trees generated before this point were discarded. A 50% majority rule consensus tree of the remaining trees was computed.

Maximum likelihood (ML) analyses were run in RAxML (Stamatakis, 2006; Stamatakis et al., 2008) using the default rapid hill climbing algorithm. Adjusting the values of distinct rate categories and rearrangement settings did not improve the likelihood scores so the defaults were used in each case. The model of nucleotide substitution chosen was GTRMIX. These analyses were run for 200 replicates and the best tree chosen from those runs. Bootstrap proportions were estimated under the same conditions for 100 pseudoreplicates. Two constraint analyses (with monophyly enforced for all subfamilies and genera in ingroup and outgroup taxa and Acanthocotylidae and Gyrodactylidae forced to be sister taxa following Boeger and Kritsky (2001) were also run under the same criteria for use in hypothesis testing.

The 50% majority rule consensus tree from the Bayesian analyses was used to view the distribution of six morphological characters in relation to the phylogenetic hypothesis produced. Description of these characters (haptoral septa, haptoral accessory sclerites, haptoral hamuli, vagina and number of testes) follows Whittington (2004) and elaboration of the anterior attachment organ morphology is shown in Fig. 1.

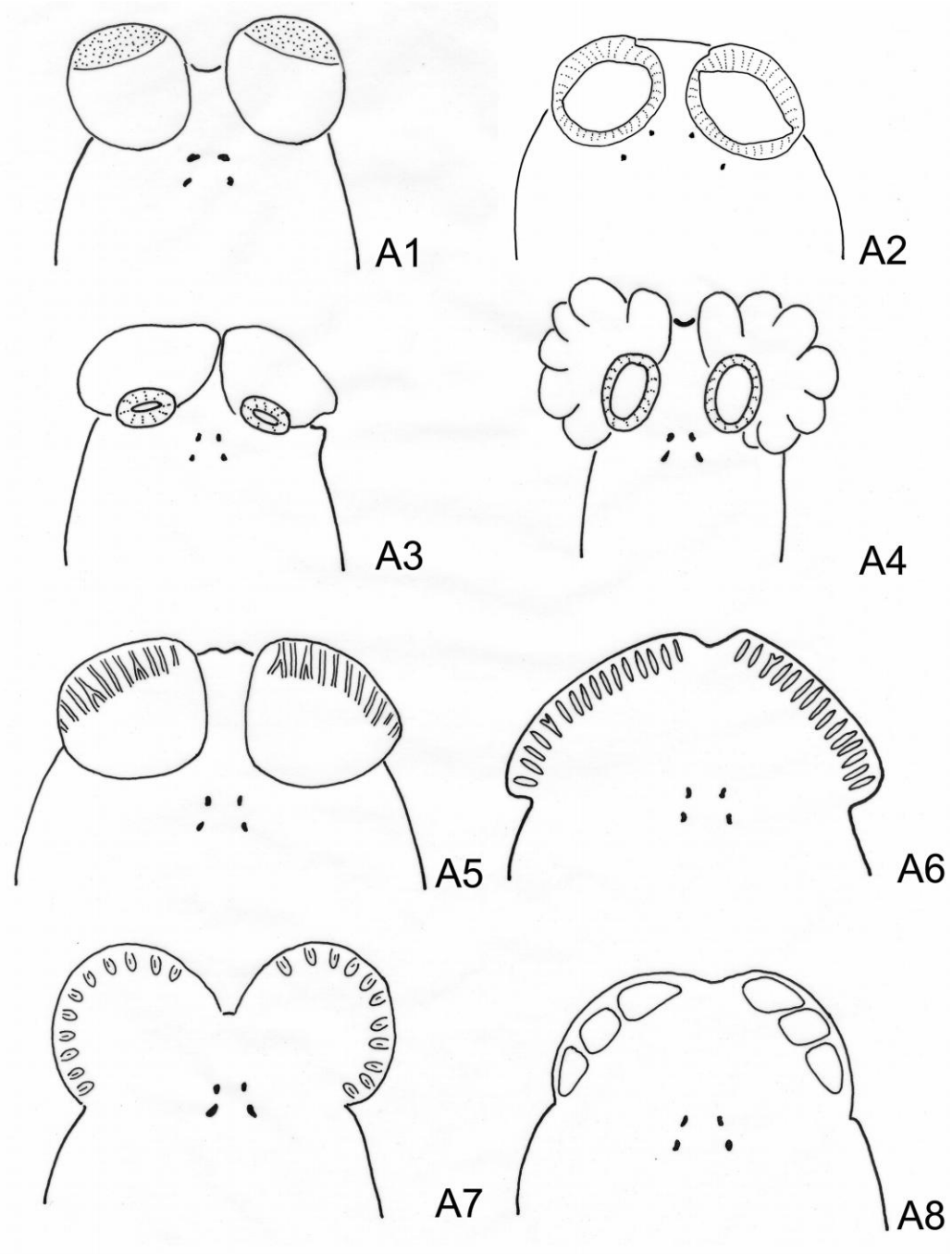


Fig. 1. Diagrammatic representations of the variation in anterior attachment organ morphology among the Capsalidae. A1 – paired circular discs, A2 – paired circular discs with anterior glandular and posterior muscular regions, A3 – paired circular discs with muscular suckers, A4 – paired structure with convoluted edges and muscular suckers, A5 – paired circular discs with anterolateral ridges, A6 – paired diadems, A7 – paired anterolateral adhesive areas with ventral columns of multiple raised ovoid structures, A8 – paired anterolateral adhesive pads each with three separate areas.

Partitioned Likelihood Support (PLS – Lee and Hugall, 2003) determines whether the different data partitions are in support or disagreement with each node of

the tree derived from the combined data matrix. PLS was assessed for all nodes found in the best ML tree produced in RAxML. PLS was analysed for the three different genes: *28S rDNA*, *H3* and *EF1α*. The log likelihood values for the three different genes for this tree were calculated in PAUP* using the site log likelihood function. The constraint trees necessary for PLS were constructed in MacClade v 4.0 (Maddison and Maddison, 1995). As reverse constraint analyses could not be run in RAxML, all analyses for the different nodes were run in GARLI v0.95 (Zwickl, 2006). The GTR model with a proportion of invariable sites and a gamma distribution for rates across sites was used. Termination conditions were set at 10,000 (genthreshfortopoterm) and 0.01 (significanttopochange). The remaining default settings were used as it has been shown that altering these generally has little effect on the likelihood scores (Zwickl, 2006). Bootstrap analyses in GARLI were run using 100 pseudoreplicates.

The approximately unbiased (AU) test is a multi-scale bootstrap technique developed for general hypothesis testing and provides a procedure to assess the confidence of tree selection. In the AU test, several sets of bootstrap replicates are generated by changing sequence length, with the number of times the hypothesis is supported by replicates counted for each set to obtain bootstrap probability values for different sequence lengths. The log likelihood values for each site (generated in PAUP*) for the ML tree without constraints, the monophyly constraint ML tree (monophyly constrained for all families, subfamilies and genera) and the ML tree with the Acanthocotylidae/Gyrodactylidae constraint (Acanthocotylidae and Gyrodactylidae were constrained to be sister taxa) were used in CONSEL version 0.1i (Shimodaira and Hasegawa, 2001) to run the AU test to determine in which trees to have confidence. Monophyly constrained for all families, subfamilies and genera was used to test the current hypothesis of capsalid classification. Acanthocotylidae and Gyrodactylidae were constrained to be sister taxa to test the hypothesis of Boeger and Kritsky (2001) who suggested that the Acanthocotylidae and Gyrodactylidae may be sister groups.

3. Results

3.1. DNA sequence characteristics

There were no premature stop codons within the coding regions of the protein coding nuclear genes. The secondary structure of the *28S rDNA* sequence for *Gyrodactylus salaris* could not be used to align our sequence data. Parts of the *28S rDNA* sequence data span a highly variable section of *28S rDNA* so areas where the model predicted stems did not correspond to conserved regions in the sequence data and so the model was not used to infer an alignment. The three loci for 47 ingroup taxa and 15 outgroup taxa were concatenated for a total alignment of 1528 characters of sequence including: 430 characters *28S rDNA*, 292 characters *H3* and 806 characters *EF1α*. This included 104 parsimony informative sites for *28S rDNA*, 141 parsimony informative sites for *H3* and 348 parsimony informative sites for *EF1α.* We were unable to obtain sequence for *H3* for *Udonella* sp. and *EF1α* for the following taxa: *Benedenia anticavaginata*, *Capsala* sp. 1, *Encotyllabe caranxi*, *Interniloculus chilensis*, *Neoentobdella diadema*, *Tristoma integrum*, *Tristoma* sp., and Trochopodinae sp. 3 (Appendix III). These taxa were included in analyses as missing data for this gene. The *EF1α* sequence spanned an intron of variable length (approx. 50–100 bp), which we excluded from our analyses because it could not be aligned unambiguously due to high variability. Some primer pairs for *28S rDNA* generated larger sequence fragments (approx. 800 bp) but because alignment at the 3‟ end of this sequence was ambiguous, only approximately 400 bp were included in analyses. Other areas of *28S rDNA* and *EF1α* sequence, where alignment was also ambiguous, were excluded from analyses reducing the final number of characters used in the analyses to 1280. Indels occurred at 29 sites in the *28S rDNA* sequence data (20 of which occurred only in *Udonella* sp.) and 14 sites in the *EF1α* sequence data. Sequencing of some *28S rDNA*, *H3*, and *EF1α* sequences revealed heterozygotes, indicated by overlapping signals for two kinds of bases in the sequence chromatograms data. These sites were scored with the IUPAC ambiguity codes for dimorphic sites.

3.2. Phylogenetic analyses

The preliminary Bayesian analyses and AIC showed that seven partitions (*28S rDNA*, *H3* 1st codon position, *H3* 2nd codon position, *H3* 3rd codon position, *EF1α* 1st codon position, *EF1α* 2nd codon position and *EF1α* 3rd codon position) were optimal for the data (Fig. 2).

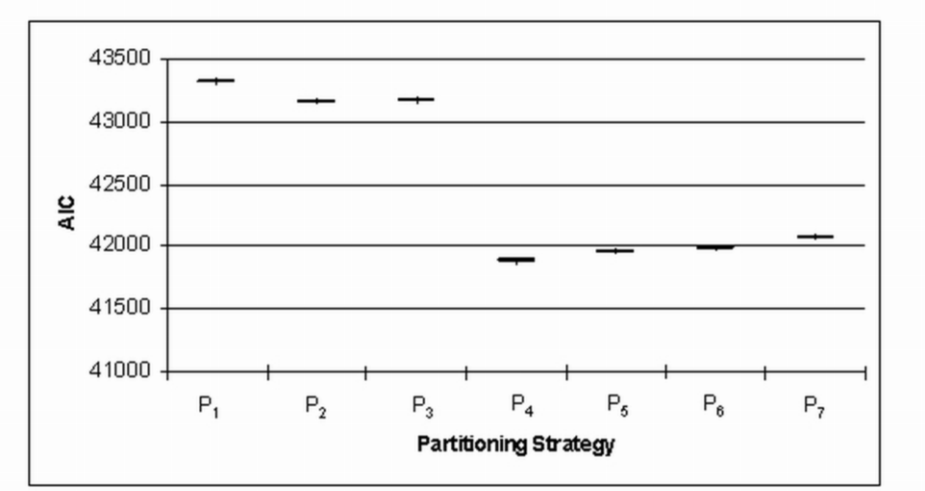


Fig. 2. AIC values for the different partitioning strategies. P_1 – All data combined (1 partition), P_2 – *28S*; *H3*; *EF1α* (3 partitions), P_3 – 28S; *H3* and *EF1α* combined (2 partitions), P_4 – *28S*; *H3* codon positions; *EF1α* codon positions (7 partitions), P_5 – 28S; *H3* and *EF1α* codon positions combined (4 partitions), P_6 – *28S; H3* codon position 1 and 2; *H3* 3rd codon position; *EF1α* codon position 1 and 2; *EF1α* 3rd codon position (5 partitions), P_7 – *28S*; *H3* and *EF1α* codon positions 1 and 2; *H3* and *EF1α* 3rd codon positions (3 partitions).

We present the Bayesian 50% majority rule consensus tree in Fig. 3 along with posterior probabilities and because the ML tree was so similar in topology, the ML bootstrap proportions (BS). For comparison, we present the ML tree in Appendix IV. Bayesian and ML analyses of the combined data (Fig. 3) yielded some interesting relationships that were recovered consistently and some were strongly supported as indicated by Bayesian posterior probabilities (PP) and non-parametric bootstrap proportions (BS). Monophyly of the Capsalidae was supported strongly (PP 100%, BS 99%) and consistently in all analyses. A clade comprising three *Gyrodactylus* species (Gyrodactylidae) and a *Udonella* sp. (Udonellidae) (Fig. 3, Clade 3) formed the sister group to the family (PP 97%, BS 63%). Of the three outgroup families

where two or more taxa were represented, two formed well supported clades: Gyrodactylidae (*Gyrodactylus* spp.; PP 100%, BS 100%) and the Microbothriidae (*Asthenocotyle*, *Dermophthirius* spp. and *Pseudoleptobothrium*; PP 100%, BS 93%). The Monocotylidae represented by a *Calicotyle* sp. and *Dendromonocotyle bradsmithi* were not monophyletic.

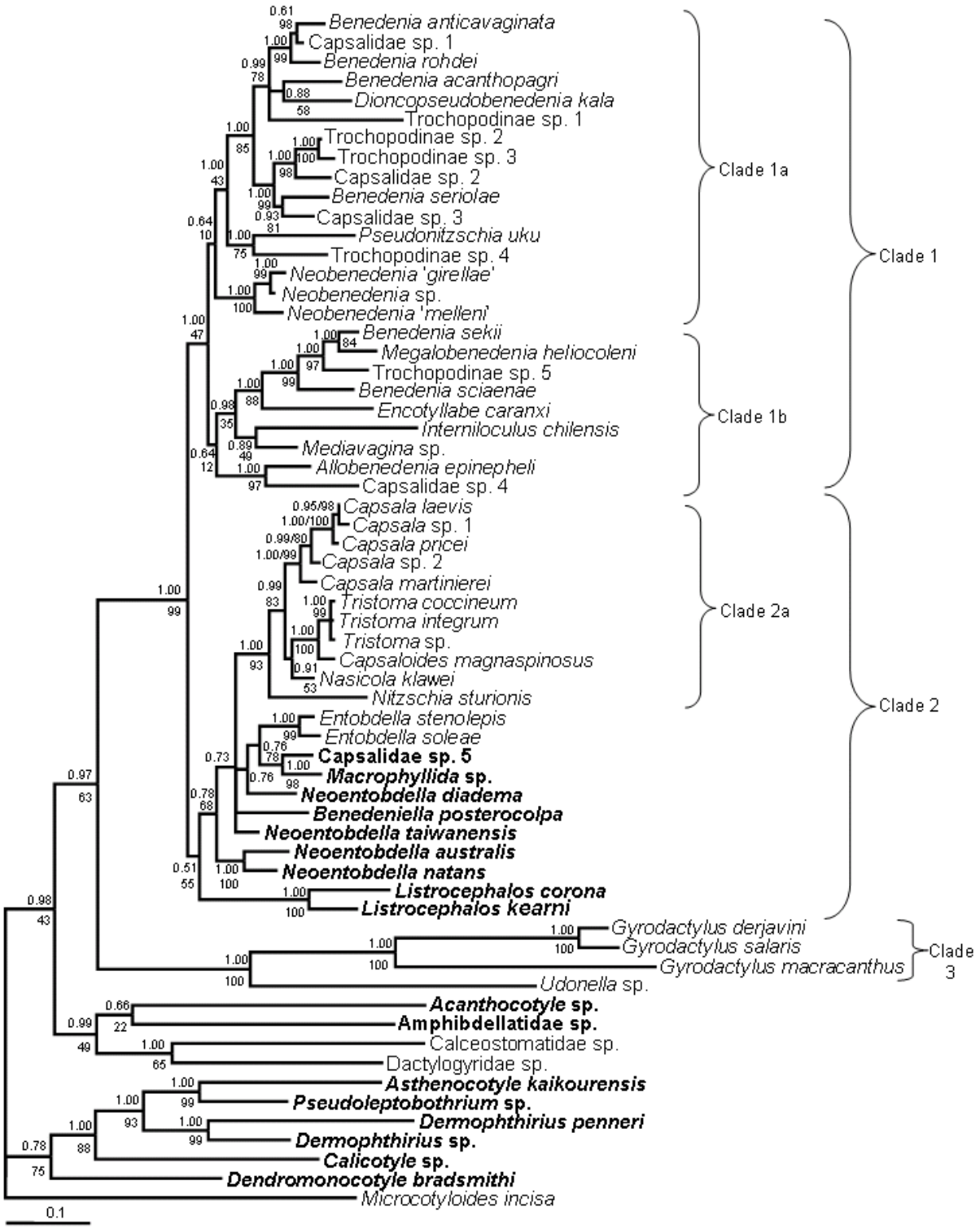


Fig. 3. A 50% majority rule consensus tree produced from Bayesian inference analyses of the combined nuclear sequence data for the Capsalidae and 15 outgroup taxa representing 9 families and 2 subclasses. Posterior probabilities and maximum likelihood bootstrap proportions are indicated above and below each node, respectively, or, in some cases in Clade 2a before and after a /, respectively. Taxa in bold parasitise elasmobranch hosts. See Table 1 for current capsalid classification, Fig. 4 for subfamily status of capsalid taxa studied and Appendix III for outgroup families.

Capsalids were split into two major clades (Fig. 3). Clade 1 comprised species currently in five subfamilies (Benedeniinae, Encotyllabinae, Interniloculinae, Pseudonitzschiinae and Trochopodinae) and nine genera. Clade 1 is further divided into two subclades (Clade 1a and Clade 1b) but while consistently recovered, these clades were not strongly supported (PP 64% for both, BS 10% and 12%, respectively). Clade 1a comprises species currently in *Neobenedenia*, Pseudonitzschiinae and other representatives of the Benedeniinae, Trochopodinae and seven undescribed capsalid species not yet assigned to a genus. Clade 1b consists of species currently in Benedeniinae, Encotyllabinae, Interniloculinae, Trochopodinae and one undescribed capsalid species unassigned to a genus. Clade 2 comprised species currently in five subfamilies: Benedeniinae (*Benedeniella posterocolpa*), Capsalinae, Entobdellinae, Nitzschiinae and Trochopodinae (*Macrophyllida* sp.) and ten genera. Clade 2 has a strongly supported subclade (PP 100%, BS 93%) within it (Clade 2a) containing all included species of Capsalinae that are the strongly supported sister group to *Nitzschia sturionis* (Nitzschiinae). The remainder of Clade 2 comprises species currently in Benedeniinae, Entobdellinae and Trochopodinae and one species unassigned to either subfamily or genus. Eight of the nine capsalid subfamilies were represented in our analyses but monophyly was only tested for four of those (Benedeniinae, Capsalinae, Entobdellinae and Trochopodinae) as three of the remaining subfamilies (Encotyllabinae, Interniloculinae, Nitzschiinae) were each represented by a single taxon and Pseudonitzschiinae is monotypic. The only capsalid subfamily not represented was the Dioncinae. Of the subfamilies tested, only the Capsalinae was found to be monophyletic (PP 99%, BS 83%). Of the 20 genera included, only seven (*Benedenia*, *Capsala*, *Entobdella*, *Listrocephalos*, *Neobenedenia*, *Neoentobdella* and *Tristoma*) were represented by multiple species to test generic monophyly. Of these, only five genera (*Capsala*, *Entobdella*, *Listrocephalos*, *Neobenedenia*, and *Tristoma*) were monophyletic and all with strong support (Fig. 3).

Despite poor support at some nodes, these phylogenetic hypotheses are strongly supported. Both Bayesian inference and ML produce concordant topologies and there is strong PP support and BS support for 75% of nodes. Positive PLS for each gene at every node (data not shown) indicates that all genes are contributing to the phylogenetic signal at all nodes, including those with poor PP and BS support, therefore supporting their usefulness as markers in analyses of phylogenetic

relationships of capsalid parasites. The PLS values did not vary significantly with the depth in the tree indicating they are contributing to all levels of the phylogeny. The large number of outgroup taxa included also allows for a better estimation of the root position.

We carried out AU tests of whether our data can reject a number of alternate hypotheses proposed in previous studies. The ML analysis produced a tree with a log likelihood of –31045.52. The ML analysis with monophyly constrained for subfamilies and genera of both ingroup and outgroup taxa produced a tree with a log likelihood of –32281.56. The results of the AU test are as follows: the ML tree without any topological constraints had a *p*-value ($\alpha = 0.05$) of 0.87, the ML tree with monophyly enforced had a *p*-value ($\alpha = 0.05$) of 0.00, indicating confidence in the ML tree produced without monophyly constraints. In the ML tree in which Acanthocotylidae and Gyrodactylidae were constrained to be sister taxa following Boeger and Kritsky (2001), the *p*-value ($\alpha = 0.05$) was 0.131 indicating confidence in both this tree and the ML tree where no topological constraints were enforced.

The distribution of six key morphological characters that are used commonly in combination to distinguish capsalid subfamilies and genera (e.g. Whittington, 2004) were assessed relative to the Bayesian hypothesis generated (Fig. 3) to examine the instance and frequency of homoplasy (Fig. 4). Haptoral septa are found in the Capsalinae, Encotyllabinae, Interniloculinae and Trochopodinae. In our study, septa were identified also in *Pseudonitzschia uku* (Pseudonitzschiinae) but were neither described nor illustrated by Yamaguti (1965, 1968). Accessory sclerites were absent in only one species, *P. uku* (Fig. 4). Hamuli are absent in the Capsalinae (represented by ten species), *Dioncopseudobenedenia kala* (Benedeniinae), Interniloculinae (represented in our study by one species) and Pseudonitzschiinae (monotypic) (Fig. 4). The vagina is absent only in *Neobenedenia* species (Benedeniinae). Anterior attachment organ morphology, not previously considered in detail, was the most complex morphological character included here with eight states present in the family (plus one uncharacterised state (A?)). Character state A1 (see Fig. 1) was predominant in both Clade 1 and Clade 2 (Fig. 4). Character states A2, A3 and A4 (see Fig. 1) were only found in Clade 1 (Fig. 4) and character states A5, A6, A7 and A8 (see Fig. 1) were only found in Clade 2 (Fig. 4). Indeed the most diverse anterior attachment organ variation is displayed in capsalid taxa infecting elasmobranchs (Fig. 4, taxa in bold) with three separate character states identified among the nine

included taxa (Clade 2). Multiple testes occur only in the Capsalinae and Pseudonitzschiinae but some Trochopodinae species not available for our analyses apparently also have multiple testes (Egorova, 1994). The only species included in the analyses with four testes was *Interniloculus chilensis* but some described Trochopodinae species also have four testes (Egorova, 1994; Whittington, 2004). Benedeniinae (11 species), Entobdellinae (eight species) and all remaining Trochopodinae species included (nine species) had two testes. Most species in the analyses have two juxtaposed testes with the exception of *Macrophyllida* sp. and *Mediavagina* sp. where they are in tandem.

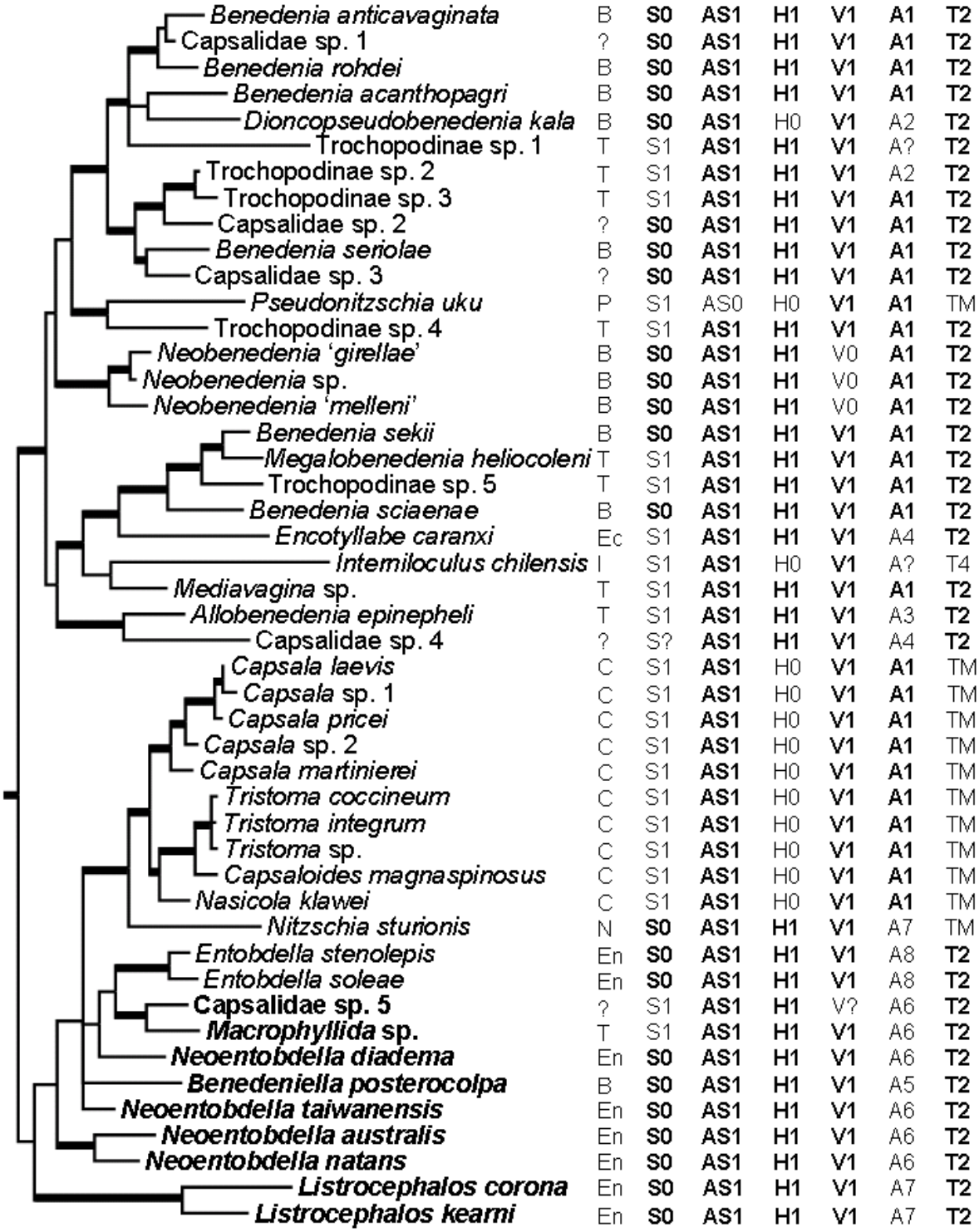


Fig. 4. A 50% majority rule consensus tree produced from Bayesian inference analyses (from Fig. 3) of the combined nuclear sequence data with current subfamily designations and distributions of key morphological characters displayed beside it. Thicker internal branches indicate those with strong support (PP > 90%). Column 1 – subfamilies: Benedeniinae (B), Capsalinae (C), Encotyllabinae (Ec), Entobdellinae (En), Interniloculinae (I), Nitzschiinae (N), Pseudonitzschiinae (P) and Trochopodinae (T); column 2 – haptoral septa (S): absent (S0), present (S1), unknown (S?); column 3 – haptoral accessory sclerites (AS): absent (AS0), present (AS1); column 4 – haptoral hamuli (H): absent (H0), present (H1); column 5 – vagina: absent (V0), present (V1), unknown (V?); column 6 – anterior attachment organ morphology (A; see Fig. 1): paired circular discs (A1), paired circular discs with anterior glandular and posterior muscular regions (A2), paired circular discs with muscular suckers (A3), paired structures with convoluted edges and muscular suckers (A4), paired circular discs with anterolateral ridges (A5), paired diadems (A6), paired anterolateral adhesive areas with ventral columns of multiple raised ovoid structures (A7), paired anterolateral adhesive pads each with three separate areas (A8), morphology unknown (A?), column 7 – number of testes: two (T2), four (T4) or multiple (TM). Characters in bold denote the most frequently occurring state. Taxa in bold parasitise elasmobranch hosts.

4. Discussion

4.1. Monophyly of the Capsalidae

Our study is the first molecular phylogeny of the Capsalidae with comprehensive taxon sampling (30 described species, seven species assigned to genus, five species assigned to subfamily and five species assigned to family) and multiple loci. Monophyly of the Capsalidae has been questioned and its composition has been changed multiple times and continues to be unstable (Yamaguti, 1963; Timofeeva, 1990; Egorova, 1999, 2000). The Dioncinae was considered previously to have familial status and to be the sister group to the Capsalidae (Bychowsky, 1957). *Dioncus* has since been incorporated into the family, based on haptoral characteristics and reproductive morphology (Timofeeva, 1990). Inclusion of the Dioncinae provides a unique morphological synapomorphy for the family (Whittington, 2004): the presence of accessory sclerites on the haptor (Kearn, 1963). Accessory sclerites are absent only in two capsalid species (*Pseudonitzschia uku*; Fig. 3, Clade 1a) and *Calicobenedenia polyprioni* (not represented in our study) which presumably represent secondary losses (Whittington, 2004). The perforated bead shape of the spermatid mitochondrion and the progressive disappearance of the microtubules of the zone of differentiation have also been suggested as synapomorphies with the inclusion of *Dioncus* into the Capsalidae (see Justine and Mattei, 1987). The Capsalidae was shown to be monophyletic by Mollaret et al. (1997) and by Whittington et al. (2004). However, as the Dioncinae was not included in their or in our analyses, a rigorous test of capsalid monophyly in future studies should include a representative taxon. Boeger and Kritsky (2001) suggested that those microbothriids which as adults lack haptoral sclerites and have two testes (e.g. *Dermophthirius penneri*, see Fig. 3) may actually be capsalids but this is not supported by our analyses because the four investigated microbothriids were monophyletic, forming a strongly supported clade (PP 100%, BS 93%) distantly related to capsalids.

4.2. Sister group to the Capsalidae

Phylogenetic hypotheses based on morphology have suggested that sister groups to the Capsalidae are the Loimoidae and Monocotylidae (see Boeger and Kritsky, 2001) while previous molecular analyses based on RNA only showed that

the Gyrodactylidae and Udonellidae are closest (Olson and Littlewood, 2002). The latter is a scenario strongly supported (PP 97%, BS 63%) in our analyses (see Fig. 3, Clade 3). It has also been hypothesised that the Acanthocotylidae is closely related to Gyrodactylidae based on multiple morphological synapomorphies (Boeger and Kritsky, 1997). While this relationship was not found in our analyses (Fig. 3), an AU test showed that our data could not reject it. More monopisthocotylean outgroups could be included to examine this relationship further.

4.3. The subfamily classification

Within the Capsalidae, the revision of some genera and species has required an ongoing reassessment of subfamilial classifications (Whittington and Horton, 1996; Egorova, 1999; Whittington, 2004). Many of these revisionary works have been done by Egorova particularly with subfamilial and generic classifications in the Capsalinae, Trochopodinae, Benedeniinae, Entobdellinae and Dioncinae (Egorova, 1989, 1994, 1997, 1999, 2000). Of the four subfamilies for which we tested monophyly (Benedeniinae, Capsalinae, Entobdellinae and Trochopodinae), only the Capsalinae is monophyletic. This subfamily has recently undergone significant revision by rigorous evaluation of original descriptions and type material. Chisholm and Whittington (2007) identified many synonymous species and reduced the seven genera and 60 species to four genera and 36 species. Interestingly, Nitzschiinae, species of which parasitise acipenserids, is sister to the Capsalinae in our analyses (Fig. 3). Capsaline species generally parasitise highly mobile pelagic species like tuna and marlin so this infers a host switching event between euryhaline sturgeons and cosmopolitan oceanic pelagic fish.

The Benedeniinae and Trochopodinae are both large subfamilies comprising 13 and 17 genera, respectively, and approximately 51 and 52 species each (Table 1; Whittington, 2004). Together they contain >50% of capsalid diversity but based on traditional morphological characters, differ principally by possession of an aseptate (Benedeniinae) or septate (Trochopodinae) haptor (Whittington, 2004). Our study demonstrates that polyphyly in the Benedeniinae is extensive indicating that relationships are widely misunderstood in this subfamily. Whittington et al. (2004) suggested that *Neobenedenia* could be placed in a separate subfamily and this is strongly supported (PP 100%, BS 100%) in our analyses since the three *Neobenedenia* species form a monophyletic group (Fig. 3). Monophyly is also

supported by the unique character, absence of a vagina (Fig. 4). The loss of the vagina may be an evolutionary innovation related to a specific mating behaviour or strategy among the species of *Neobenedenia* and this deserves further investigation. Insemination is likely achieved by sperm being introduced via the common genital pore (Whittington and Horton, 1996). A single specimen of *Neobenedenia* has been observed with its penis directed into its own uterus indicating they may self-inseminate (Whittington and Horton, 1996). With the confused composition of the Benedeniinae, it is currently unreasonable to erect a new subfamily without first re-examining the subfamily to which *Neobenedenia* presently belongs.

The Trochopodinae has been considered previously a "dumping ground" for capsalid species that are not assignable to other subfamilies and shows most morphological variation in testes number (Whittington, 2004). Its unsatisfactory definition is only further highlighted in our analyses. Whittington (2004) predicted that members of the Interniloculinae and Pseudonitzschiinae could be moved to the Trochopodinae on further study. While they do appear to be closely related to some so-called species of Trochopodinae, the extreme polyphyletic state of species currently assigned to this subfamily as shown in our analyses precludes inclusion of *Interniloculus* and *Pseudonitzschia* at this stage.

The Entobdellinae has undergone recent revision (Kearn and Whittington, 2005; Kearn et al., 2007) and is considered currently to comprise 23 species in five genera (see Table 1; *Entobdella*, *Branchobdella*, *Listrocephalos*, *Neoentobdella* and *Pseudoentobdella*). Our analyses, however, show paraphyly among this group of capsalids that parasitise both elasmobranchs and teleosts. In our hypothesis, a *Macrophyllida* sp. (currently considered to be a Trochopodinae) and *Benedeniella posterocolpa* (currently in the Benedeniinae) group with entobdellines and two *Listrocephalos* species group together in a separate but closely related clade. The positions of *Benedeniella postercolpa* (Benedeniinae) and *Macrophyllida* (Trochopodinae) within the Entobdellinae (Fig. 3) are consistent with the host range and these species share some morphological characteristics with other entobdellines (e.g. anterior attachment organ morphology, see Figs. 1 and 4). Species of *Trimusculotrema* (Benedeniinae) and *Sprostonia* (Trochopodinae), which were not included in our study, also infect elasmobranch hosts and will be valuable additions to future analyses.

No representative from Dioncinae was available. Dioncinae infect remoras of the Echeneidae such as *Echeneis* and *Remora* but species are also recorded from carangids and rachycentrids (Table 1). Remoras can be „carried" on larger organisms such as sharks, rays, teleosts, turtles and cetaceans. *Dioncus* attach their eggs to the gills of remoras and therefore these teleosts may provide a vector for host switching from chondrichthyans to teleost fish groups or perhaps in the other direction (Whittington, 2004). With capsalid parasites from sharks and rays grouping together, it is possible that remoras have been the means of transmission for ancestral capsalids on elasmobranchs to a diversity of teleost hosts.

4.4. Generic classifications

Of the 46 capsalid genera recognised, some remain poorly defined (Whittington, 2004). Five (*Capsala*, *Entobdella*, *Listrocephalos*, *Neobenedenia* and *Tristoma*) of the seven genera for which we had more than one representative were monophyletic. Genera represented by large numbers of species in our analyses such as *Benedenia* (six of 21 species included) were not monophyletic and were spread throughout Clade 1 of the tree. Unexpectedly, *Neoentobdella* was also not monophyletic in the Bayesian analyses (Fig. 3). It was monophyletic in the ML analyses but with very weak support (BS 11%, see Appendix IV). The genus was erected recently based on morphological characters and host association and comprises 10 species infecting rays (Whittington and Kearn, 2009). Our analyses included four described *Neoentobdella* species (Whittington and Kearn, 2009). Our analyses indicate that further revision of *Benedenia* is needed but monophyly for *Neoentobdella* cannot be rejected and further work incorporating faster evolving genes is required. The confused state of capsalid subfamilial classification is further complicated by poor generic definitions. Adding a mitochondrial dataset may also help to tease out some of the shallower relationships in the tree and further test support in these areas. This, along with broader taxon representation, will further elucidate relationships within the Capsalidae.

4.5. Systematic utility of morphological characters

In parasites, molecular genetic data have been viewed more favourably than morphological data for phylogenetic analyses due to the apparent lack of stability of morphological based hypotheses and the lack of available morphological characters

(Littlewood et al., 1999b). However, morphological analyses are important as they allow the identification of synapomorphies and lead to the development of a robust set of characters with which to describe taxa. Examination of the distribution of defining morphological characters relative to our molecular phylogenetic hypotheses generated shows that some of the character states (haptoral septa, haptoral hamuli, anterior attachment organ morphology and testis number) show apparent homoplasious evolution in the Capsalidae. These morphological characters may be homoplastic due to convergent evolution which is considered highly likely in parasites given the similar life history challenges they face (Poulin and Morand, 2000). A parasite must find its host, attach to it and then derive nutrition from it. Similarities in the type of host and specific microhabitat parasitised may elicit morphological adaptations by parasites that impose phylogenetic constraints on character evolution (Whittington, 2004). Homoplasy may also be an artefact of poor or insufficient character state definitions. While capsalid morphology is considered conserved, there is variation within some of these characters. The usefulness of morphological characters is thought to increase with the complexity with which they are described (Littlewood et al., 1999b). Currently five of the subfamilies have haptoral septa but the haptors are divided in very different ways. The Capsalinae haptor is divided into a series of peripheral compartments surrounding a central loculus. This arrangement is not seen in the haptoral septa of the other subfamilies (Whittington, 2004). Many of these characters, at the detail to which they are described, are also not unique to the Capsalidae. Septate haptors occur in other monogenean families (e.g. Monocotylidae) but there has been no assessment about whether these structures are homologous (Whittington, 2004).

Individual characters will only contribute to relationships at certain levels of a tree. Many of the anterior attachment organ morphologies only apply to species in a single genus and so provide no information on relationships at higher levels. Some combinations of these characters appear to define some subfamilies and genera relative to the molecular phylogenetic analysis. The Capsalinae are defined morphologically as having haptoral septa, presence of accessory sclerites, absence of haptoral hamuli, presence of a vagina, paired anterior circular discs and multiple testes (Fig. 4). There are no other taxa in these analyses that have this combination. Similarly, *Neobenedenia* can be defined as lacking haptoral septa and a vagina, but possessing accessory sclerites, haptoral hamuli, paired anterior circular discs and two

juxtaposed testes, a combination unique to this genus (Fig. 4). However, with only six morphological characters commonly used to distinguish capsalid subfamilies and genera, it is inconceivable that these characters can comprehensively define the relationships at all levels between the approximately 180 described capsalid taxa. More morphological characters are needed to provide phylogenetic information throughout all levels of the tree.

These simplistic definitions and paucity of morphological characters provide little information on relationships at any level and this is reflected in the disparity between the molecular phylogenetic hypothesis and morphological taxonomy. Perhaps these characters need examining at an ultrastructural level to identify informative structural differences. New characters need exploring, such as larval characters, as they are believed to be less modified by parasitism and better reflect ancestry (Whittington, 2004). Care must be taken when examining and inferring further characters and states. Biological and environmental variables such as parasite and host age, host species and water temperature can also induce changes in morphology making characters problematic when used in phylogenetic analyses due to phenotypic plasticity and low heritability (Brooks and McLennan, 1993). The phylogenetic framework presented here provides a basis to explore further morphological characters.

The Linnaean ranks used for classification of taxa within the Capsalidae are subjective because they are not based on phylogenetic hypotheses. Furthermore our analyses show they are also poor estimates of relationships within the family likely due to homoplasy. As a consequence of the small number of informative adult morphological characters in these parasites and the logistical problems associated with documenting variation in larval or gamete characters, it is unlikely that a morphological dataset robust enough to establish a comprehensive phylogenetic hypothesis will be compiled any time soon. While molecular data are providing new and valuable insights into the relationships of these parasites, by themselves they are no more useful in defining Linnaean ranks. This is not an uncommon dilemma and there has been much debate in the literature about how to combine traditional taxonomy with phylogenetic relationships (Moore, 1998; Brummitt, 2002; Schuh, 2003; Horandl, 2006). The PhyloCode has been proposed as a means for governing nomenclature in a phylogenetic context (Cantino and de Queiroz, 2007). Since its inception, one of the biggest criticisms has been a failure to develop a means to deal

with species ranks. However, there is now a system proposed whereby Linnaean binomials can be used in a way that is consistent with phylogenetic nomenclature (Dayrat et al., 2008). Such a system, that can bridge the legacy of the extensive use of Linnaean ranks with the principles of phylogenetic nomenclature based on molecular phylogenies, is perhaps where the answer lies for producing a classification that both conveys biological information and the phylogenetic history of these organisms.

Acknowledgments

We are exceptionally grateful to those colleagues and collaborators who have collected and provided us with specimens, without which this study would not have been possible (Appendix III). We also thank Gaynor Dolman for her continual support and advice throughout this project. This study was funded by an Australian Research Council Discovery Grant (DP0556780) awarded to I.D.W. and S.C.D. E.M.P. was supported by an Australian Postgraduate Award during PhD candidature in the School of Earth and Environmental Sciences at the University of Adelaide.

Supplementary data

Supplementary data associated with this chapter is shown in Appendices III and IV.

Closing the mitochondrial circle on paraphyly of the Monogenea (Platyhelminthes) infers evolution of diet in parasitic flatworms★

Elizabeth M. Perkins [a,*], Steve C. Donnellan [b,c], Terry Bertozzi [b], Ian D. Whittington [a,c,d]

[a]*Marine Parasitology Laboratory, School of Earth and Environmental Sciences (DX 650 418), The University of Adelaide, North Terrace, Adelaide, SA 5005, Australia*
[b]*Evolutionary Biology Unit, The South Australian Museum, North Terrace, Adelaide, SA 5000, Australia*
[c]*Australian Centre for Evolutionary Biology and Biodiversity, The University of Adelaide, North Terrace, Adelaide, SA 5005, Australia*
[d]*Monogenean Research Laboratory, Parasitology Section, The South Australian Museum, North Terrace, Adelaide, SA 5000, Australia*

★ *Note*: Nucleotide sequence data reported in this paper are available in GenBank™ under the accession numbers: XXXXXXXXX-XX.

*Corresponding author. Address: *Marine Parasitology Laboratory, School of Earth and Environmental Sciences (DX 650 418), The University of Adelaide, North Terrace, Adelaide, SA 5005,* Australia. Tel.: +61 8 8303 8245; fax: +61 8 8303 4364.

Email address: elizabeth.perkins@adelaide.edu.au or lizzie.perkins@gmail.com (E. Perkins).

International Journal for Parasitology (2010) **In Press**.

ABSTRACT

Relationships between the three classes of Neodermata (parasitic Platyhelminthes) are much debated and restrict our understanding of the evolution of parasitism and contingent adaptations. The historic view of a sister relationship between Cestoda and Monogenea (Cercomeromorphae; larvae bearing posterior hooks) has been dismissed and the weight of evidence against monogenean monophyly has mounted. We present the nucleotide sequence of the complete mitochondrial (mt) genome of *Benedenia seriolae* (Monogenea: Monopisthocotylea: Capsalidae), the first complete non-gyrodactylid monopisthocotylean mt genome to be reported. We also include nucleotide sequence data for some mt protein coding genes for a second capsalid, *Neobenedenia* sp. Analyses of the new mt genomes with all available platyhelminth mt genomes provides new phylogenetic hypotheses, which strongly influence perspectives on the evolution of diet in the Neodermata. Our analyses do not support monogenean monophyly but confirm Digenea and Cestoda are each monophyletic and sister groups. Epithelial feeding Monopisthocotylea of fish hosts are basal in the Neodermata and represent the first shift to parasitism from free-living ancestors. The next evolutionary step in parasitism was a dietary change from epithelium to blood. The common ancestor of Digenea + Cestoda was monogenean-like and most likely sanguinivorous. From this ancestral condition, adult Digenea and Cestoda independently evolved dietary specialisations to suit their diverse microhabitats in their final vertebrate hosts. These improved perspectives on relationships fundamentally enhance our understanding of the evolution of parasitism in the Neodermata and in particular, the evolution of diet.

1. Introduction

The Neodermata (Platyhelminthes) comprises three classes of parasitic flatworms (Cestoda, Trematoda and Monogenea) that are diverse in morphology, host association, microhabitat, lifecycle, behaviour and diet (Rohde, 1994). The replacement of larval epidermis by a syncytial neodermis with subepithelial perikarya unites the Neodermata phylogenetically (Ehlers, 1985). Cestoda are endoparasitic tapeworms with complex lifecycles (multiple hosts) that infect a diversity of final vertebrate (all classes) and intermediate invertebrate (typically crustacean, insect, mollusc and annelid) hosts. The cestodes lack a gut and feed by actively absorbing nutrients across their neodermis (Dalton et al., 2004). Trematoda are endoparasitic flukes with complex lifecycles, the adults of which infect all vertebrate classes, their intermediate hosts are principally molluscs and adults exhibit a diversity of diets including blood and epithelia digested in their gut (Rohde, 1994). Monogenea differ in that they have no intermediate host and predominantly parasitise the external surfaces of fish and display two distinctive diets that traditionally divide them into two subclasses; the blood feeding Polyopisthocotylea and the epithelial feeding Monopisthocotylea (see Littlewood et al., 1999). These are sometimes named Heteronchoinea and Polyonchoinea, respectively (Boeger and Krisky, 2001). These subclasses are united by various morphological synapomorphies: larvae with three ciliated zones, adults and larvae with two pairs of pigmented eyes, one pair of ventral anchors (hamuli) and one egg filament (Lockyer et al., 2003). Conjecture about whether the Monogenea is monophyletic has been rife for decades (e.g. Bychowsky, 1957; Llewellyn, 1970; Justine, 1991, 1998; Mollaret et al., 2000; Lockyer et al., 2003).

Phylogenetic analyses based on morphological characters typically support monogenean monophyly (e.g. Boeger and Kritsky, 1993, 2001) but an exception is work by Justine (1991) who could not define a synapomorphy for Monogenea based on character analysis of spermatozoon and spermiogenesis ultrastructure. Independent synapomorphies for each of the Monopisthocotylea and Polyopisthocotylea, however, were defined from spermatological characters (Justine, 1991). Pariselle and Euzet (1995) argued that differences in the origin, structure and postlarval development of attachment sclerites suggested that they are not homologous in Monopisthocotylea and Polyopisthocotylea. Studies of the larval osmoregulatory system and changes during postlarval development by Euzet et al.

(1995) supported the hypothesis of polyphyletic origins for Monopisthocotylea and a monophyletic origin for Polyopisthocotylea. Molecular genetic analyses of parasitic platyhelminths have reported paraphyly (Baverstock et al., 1991 – partial *18S rRNA* including two monogenean species; Rohde et al., 1993 – partial *18S rRNA* including three monogenean species) or monophyly (Lockyer et al., 2003 – near complete *18S rRNA*, partial *28S rRNA* including four monogenean species) for the Monogenea. Morphological and molecular genetic analyses of monogenean relationships to date are subject to the same set of problems, i.e. variation in taxon and character selection between studies (Littlewood et al., 1998) and the molecular genetic studies are limited further by the use of only two linked rRNA loci (Duvall and Ervin, 2004; Xie et al., 2008).

Relationships between the three classes of Neodermata have been explored but a strong consensus has not been reached. Historically Cestoda and Monogenea have been considered sister groups (i.e. the Cercomeromorphae; united by larvae bearing posterior hooks) with Trematoda (=Aspidogastrea + Digenea) sister to the Cercomeromorphae (see Rohde, 1994; Littlewood et al., 1999). Recent molecular genetic studies based on *rDNA* genes (Lockyer et al., 2003) and mitochondrial (mt) genomes (Park et al., 2007) contradict this hypothesis with varying degrees of confidence by placing monogeneans outside of clades comprising representatives of Cestoda and Digenea. Assessing monophyly of the Monogenea and the consequent relationships of its subclasses is vital to our understanding of the evolution of parasitism in the Neodermata. The branching order at the base of the neodermatan clade is critical for reconstruction of the ancestral character states of the biological adaptations promoting the transition from free-living to ectoparasitic to obligate internal parasitic life histories. Some of the next steps to advance our understanding of neodermatan evolutionary relationships should involve a diverse range of molecular genetic markers and an increase in the diversity of taxa sampled (Littlewood et al., 1998). Here we address these two propositions and apply the phylogenetic power of mt genomics to neodermatan relationships in an analysis that includes a wider taxonomic range of monogeneans.

Complete mt genome sequences are capable of resolving some persistent phylogenetic problems in metazoan evolution (e.g. Ruiz-Trillo et al., 2004). Currently 29 complete mt genomes of parasitic flatworms are available, dominated by medically and economically important digeneans and cestodes. Only four

complete mt genomes for Monogenea are available: three species of *Gyrodactylus* represent one family (Gyrodactylidae) of the skin feeding Monopisthocotylea (see Huyse et al., 2007, 2008; Plaisance et al., 2007) and a fourth, *Microcotyle sebastis* (Microcotylidae), represents the blood feeding Polyopisthocotylea (see Park et al., 2007), but these data have not been analysed phylogenetically together. Indeed Park et al. (2007) included only a single monogenean (*M. sebastis*) in their analyses and so could not test monophyly of the Monogenea. We present the first complete mt genome for a non-gyrodactylid monopisthocotylean, *Benedenia seriolae* (Capsalidae), which parasitises jacks and trevallies (Carangidae) and is a pathogen in marine aquaculture (Whittington et al., 2001). We also include nucleotide sequence data for eight mt protein coding genes for a second capsalid, *Neobenedenia* sp. We phylogenetically analyse these and all other available platyhelminth mt genomes to test monogenean monophyly and infer the evolution of diet in the Neodermata, a group of global evolutionary and biomedical significance.

2. Materials and methods

2.1. Specimen collection and DNA extraction

Specimens of *B. seriolae* were collected by Dr K. Hutson from skin of wild *Seriola hippos* (Carangidae) in South Australia and specimens of *Neobenedenia* sp. were collected by E.M.P., I.D.W., J.C. Perez Urbiola and R. Inohuye Rivera from fins of wild *Sphoeroides annulatus* (Tetraodontidae) off La Paz, Mexico. DNA was extracted according to the Gentra Kit (Gentra Systems) protocol for animal tissues preserved in ethanol. Extracted DNA was stored in hydration solution at 4 °C.

2.2. PCR and sequencing

Initial short fragments of the *B. seriolae* mt genome were obtained using primers listed in Table 1. From these sequence data, long range primers were designed to amplify the genome in two sections. Sequence for these large overlapping fragments was obtained by primer walking and sequencing (Table 1). Long range PCR amplifications were performed in 25 μL reactions using the Expand Long Template PCR System (Roche) following kit protocols. Two large fragments of the *Neobenedenia* sp. mt genome were also amplified using the same system. Protein coding regions of the *Neobenedenia* sp. mt genome were sequenced with primers

listed in Table 1. All PCR products were cleaned using Agencourt® AMPure® PCR purification kit and were cycle sequenced using the BigDye Terminator v3.1 cycle-sequencing kit (Applied Biosystems). The cycling protocol consisted of 25 cycles of denaturation at 96 °C for 30 s, annealing at 50 °C for 15 s and extension at 60 °C for 4 min. All samples were sequenced on an Applied Biosystems 3730 DNA sequencer. Sequence chromatograms were edited using SeqEd version 1.0.3.

Table 1

Primers used for PCR amplification and sequencing.

Gene	Primer ID	Sequence (5‟-3‟)	Source
CytB	CytbF	GGWTAYGTWYTWCCWTGRGGWCARAT	a
CytB	CytbR	GCRTAWGCRAAWARRAARTAYCAYTCWGG	a
COX1	LCO1490	GGTCAACAAATCATAAAGATATTGG	a
COX1	HCO2198	TAAACTTCAGGGTGACCAAAAAATCA	a
COX1	LM702	CACATTGAAACCTTTAATTGCCAC	b
CytB	HM703	GAAAATAAACAGCATCAGTGTAAC	b
CytB	LM710	TACTTTAATGCGGGTTTTTGCTGC	b
COX1	HM711	ATCAAGCACTCAAAGCATTTAAAC	b
12S	HM833	GTATGGAATTTYTGGRCC	b
COX2	LM730	ATCATTTGGGGGTATTTATAGG	c
ND5	HM778	CTGATTAAATAAACCACTAC	c
ND5	HM767	GCTCTCTTTGATATAACTAC	c
ND5	HM733	ATATTAAAAAATCCTGTCCC	c
COX3	HM722	GAAAAGTGAAGCCCTACTGTAC	c
16S	LM702	CACATTGAAACCTTTAATTGCCAC	c
COX1	LM708	CGTAACTATGATTATTGGTGTTCC	c
12S	LM721	TTAGGGYCCARAAATTCCATAC	c
COX2	LM730	ATCATTTGGGGGTATTTATAGG	c
ND6	LM884	ACTAAGATTTATTTTTGTCG	c
ND2	LM749	GTTGGATTGTTATTTGGTTGAT	c
ND1	HM830	AAATTCTTAGACCCTCACC	c
ND2	HM829	TAATCAACCAAATAACAATCC	c
CytB	LM742	GCTGATCCTTTAGTTACTCC	c
16S	HM832	TAAGAAAATAAATGTTGTG	c
ND4	HM883	CGATAAATACCAATAATACC	c
CytB	BSND4L	GTTTTTATTAGTCAGTTAGG	c
ND4	BSND4H	GAAATTCTTACTAACAACCG	c
16S	LM925	GACGGAAAGACCCTGAAATC	d
CYTB	HM855	CATGAGCAGCAAACACACGC	d
ND5	LM942	TTGRTTARAGGCTATGCGKGC	d
16S	HM926	GATTTCAGGGTCTTTCCGTC	d
COX3	HM927	TCACAACATGAGTGAAATGTAGGC	e
COX2	HM939	ASSWTCARTACCACTGMCGRC	e
ND5	LM959	ATGTGGTATTATATCTTTGG	e
ND5	HM960	TACCAAAGATATAATACCAC	e
ND4	LM945	CMCGCATAGCCTYTAAYCAAC	e
COX1	LM930	CTTTAATAGGTGTTTGAGCTGG	e
ND1	LM982	ATTATATTAAATCGCAAAGCTG	e
ND3	HM1026	TTTTCTTTTAAGAATTTTTTCG	e
ND3	LM1027	CTTGTTCAAGAAAACTTACACC	e
COX1	HM1023	GCTTATTTAGTATGGTTATGCG	e
ND6	LM967	TTTAYTTAATTTATGKGGG	e

[a]Boore and Brown (2000); [b]*Benedenia seriolae* long range PCR primers (this study); [c]*B. seriolae* sequencing primers (this study); [d]*Neobenedenia* sp. long range PCR primers (this study); [e] *Neobenedenia* sp. sequencing primers (this study).

2.3. Annotation of Benedenia seriolae *mt genome*

Protein coding and ribosomal RNA genes were identified through BLAST searches and by alignment to available platyhelminth complete mt genomes. These alignments were used to determine likely start and stop codons of the protein coding genes. Protein coding regions were translated with the echinoderm and flatworm mitochondrial code (code 9). The tRNAs were identified by eye and structures confirmed using the secondary structure folding program RNA structure (Mathews, 2006). Final annotation of the genome was carried out in MacVector® 10.5 (Accelrys).

2.4. Vouchering of material

A voucher specimen of *B. seriolae* and *Neobenedenia* sp. are deposited in the Australian Helminthological Collection (AHC) of the South Australian Museum, Parasitology Section, North Terrace, Adelaide, South Australia 5000 with the registration numbers AHC 29181 (slide) and AHC 45392 (spirit), respectively.

2.5. Phylogenetic analyses and hypothesis testing

The 29 complete mt genomes for parasitic platyhelminths, the partial mt genome of the „turbellarian" *Microstomum lineare* and the complete mt genomes of three lophotrochozoan outgroups *Terebratulina retusa* (Brachiopoda: Rhynchonellata), *Platynereis dumerilii* (Annelida: Polychaeta) and *Loligo bleekeri* (Mollusca: Cephalopoda) were downloaded from GenBank™ using BACA (Antao et al., 2007) (Table 2).

Table 2

Accession numbers for mitochondrial genomes of taxa used.

GenBank™ No.	Species	Higher taxon
NC_008945	*Diphyllobothrium latum*	Cestoda
NC_009463	*Diphyllobothrium nihonkaiense*	Cestoda
NC_011121	*Echinococcus canadensis*	Cestoda
NC_008075	*Echinococcus granulosus*	Cestoda
NC_000928	*Echinococcus multilocularis*	Cestoda
NC_009461	*Echinococcus oligarthrus*	Cestoda
NC_011122	*Echinococcus ortleppi*	Cestoda
NC_009460	*Echinococcus shiquicus*	Cestoda
NC_009462	*Echinococcus vogeli*	Cestoda
NC_002767	*Hymenolepis diminuta*	Cestoda
NC_011037	*Spirometra erinaceieuropaei*	Cestoda
NC_004826	*Taenia asiatica*	Cestoda
NC_002547	*Taenia crassiceps*	Cestoda
NC_009938	*Taenia saginata*	Cestoda
NC_004022	*Taenia solium*	Cestoda
NC_010976	*Gyrodactylus derjavinoides*	Monogenea
NC_008815	*Gyrodactylus salaris*	Monogenea
NC_009682	*Gyrodactylus thymalli*	Monogenea
NC_009055	*Microcotyle sebastis*	Monogenea
NC_012147	*Clonorchis sinensis*	Digenea
NC_002546	*Fasciola hepatica*	Digenea
NC_011127	*Opisthorchis felineus*	Digenea
NC_002354	*Paragonimus westermani*	Digenea
NC_008074	*Schistosoma haematobium*	Digenea
NC_002544	*Schistosoma japonicum*	Digenea
NC_002545	*Schistosoma mansoni*	Digenea
NC_002529	*Schistosoma mekongi*	Digenea
NC_008067	*Schistosoma spindale*	Digenea
NC_009680	*Trichobilharzia regenti*	Digenea
AY228756	*Microstomum lineare*	„Turbellaria“
NC_000931	*Platynereis dumerilii*	Polychaeta
NC_000941	*Terebratulina retusa*	Brachiopoda
NC_002507	*Loligo bleekeri*	Cephalopoda

The 12 protein coding regions were aligned in SeAl version 2.0a11 (Rambaut, 1996) using inferred amino acid sequences against each other and to our genome. The protein coding regions of eight mt genes (*ND5*, *ND4*, *ND4L*, *ND3*, *CYTB*, *COX3*, *COX2* and *COX1*) from *Neobenedenia* sp. were also included. Monte Carlo Markov Chain (MCMC) Bayesian phylogenetic analyses were run using Mr Bayes version 3.1.1 (Huelsenbeck and Ronquist, 2001). To determine an optimal partitioning strategy, preliminary Bayesian analyses (10^6 generations) using each possible

partitioning strategy were run and then the AIC for each partitioning strategy calculated. The model of nucleotide substitution for each partition was assessed using the Akaike Information Criteria (AIC – Akaike, 1985) in ModelTest version 3.7 (Posada and Crandall, 1998). The General Time Reversible (GTR) model with a proportion of invariable sites and a gamma distribution for rates across sites was selected. The final MCMC analyses were run for 10^7 generations with a sample frequency of every 200 generations. Tracer version 1.4 (Rambaut and Drummond, 2007) was used (to plot the generation number against the log likelihood value) to identify the point at which log likelihood values became stable and all trees generated before this point were discarded. A 50% majority rule consensus tree of the remaining trees was computed.

Maximum Likelihood (ML) analyses were run in RAxML (Stamatakis, 2006; Stamatakis et al., 2008). Three constraint analyses were also run: 1) monophyly was enforced for Monogenea, 2) Cestoda and Monogenea were forced to be sister groups as proposed by the Cercomeromorphae theory and 3) monophyly was enforced for the Monopisthocotylea. The Approximately Unbiased (AU) test is a multi-scale bootstrap technique developed for general hypothesis testing and provides a procedure to assess the confidence of tree selection. In the AU test, several sets of bootstrap replicates are generated by changing sequence length, with the number of times the hypothesis is supported by replicates counted for each set to obtain bootstrap probability values for different sequence lengths. The log likelihood values for each site (generated in PAUP*) for the ML tree without constraints, the ML constraint tree where monophyly was enforced for the Monogenea, the ML tree with the Cercomeromorphae constraint and the ML tree enforcing monophyly for the Monopisthocotylea were used in CONSEL version 0.1i (Shimodaira and Hasegawa, 2001) to run the AU test to determine in which trees to have confidence.

3. Results

3.1. Genome organisation and gene order

The complete mt sequence of *Benedenia seriolae* is 13498 bp (Fig. 1). The entire mt genome of *B. seriolae* is transcribed on the light strand as seen in other monogenean species. The genome comprises 12 protein coding genes, two ribosomal RNAs and 22 tRNAs. Like other platyhelminth mt genomes, it lacks the gene encoding *ATP8*. While the gene order is the same as the published *Gyrodactylus* spp.

there are some tRNA rearrangements. *Gyrodactylus* spp. have $tRNA^{Phe}$ between ND4 and non-coding region 1 while *B. seriolae* has $tRNA^{Thr}$, $tRNA^{Phe}$, $tRNA^{Gln}$ and $tRNA^{Met}$ in this location. $tRNA^{Thr}$ occurs between *COX1* and *16S rRNA* in *Gyrodactylus* spp. In *Gyrodactylus* spp., seven tRNAs ($tRNA^{Tyr}$, $tRNA^{Leu1}$, $tRNA^{Gln}$, $tRNA^{Met}$, $tRNA^{Ser2}$, $tRNA^{Leu2}$ and $tRNA^{Arg}$) are located between *ND6* and *ND5* with a non-coding region in the middle. By contrast in *B. seriolae* five tRNAs ($tRNA^{Tyr}$, $tRNA^{Leu1}$, $tRNA^{Ser2}$, $tRNA^{Leu2}$ and $tRNA^{Arg}$) were found between *ND6* and *ND5*. Interestingly this gene order follows the exact tRNA arrangement seen between *ND6* and *ND5* in the polyopisthocotylean *Microcotyle sebastis*.

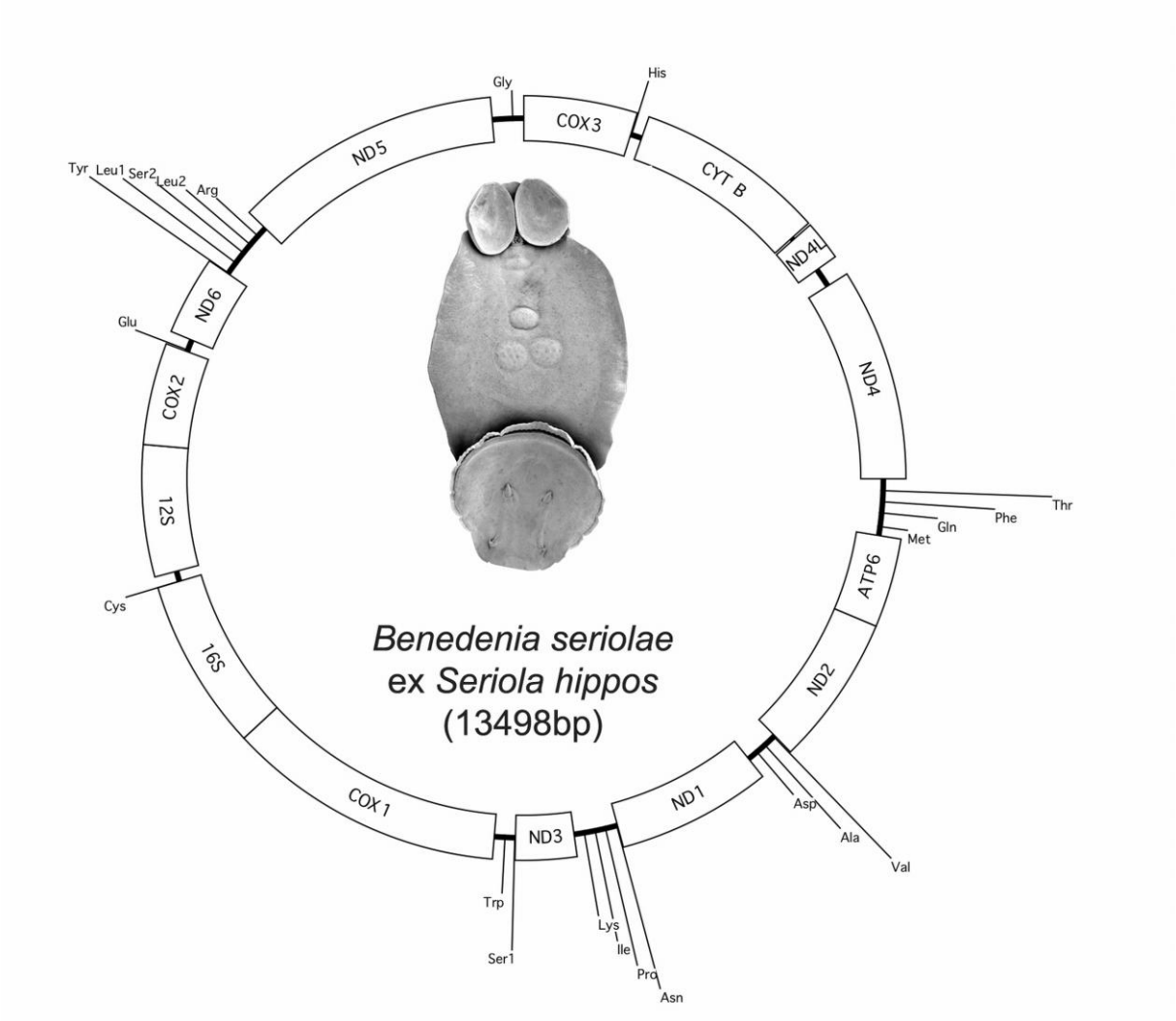


Fig. 1. Order and relative size of the genes in the complete mitochondrial genome of the capsalid *Benedenia seriolae* (Monogenea: Monopisthocotylea).

3.2. Non-coding regions

Intergenic spacers are usually limited in number and size in invertebrates making their mt genomes generally smaller in comparison to vertebrate mt genomes. *Gyrodactylus* spp. have two large, conserved non-coding regions, one located between $tRNA^{Met}$ and $tRNA^{Ser2}$ and the other between *ND4* and *ATP6*. These long non-coding regions are absent in *B. seriolae* making the latter genome more compact. Repeat regions have been found in many different mt genomes but are considered uncommon and hotspots for replication slippage errors and translocation (Le et al., 2002). Only the one tandem repeat region (TA_{12}) was found in the *B. seriolae* genome and is located between *ND4L* and *ND4*. Repeat regions are often associated with the origin of replication (also known as the control region; Le et al., 2002). The control region has proved difficult to locate in parasitic platyhelminths but the TA_{12} repeat region is most likely embedded within the control region in *B. seriolae*. Mononucleotide cytosine repeats have also been associated with the control region in other animals (e.g. fish; Zardoya et al., 1995) and one is also present between *ND4L* and *ND4* in the *B. seriolae* mt genome. Many mt genes in *B. seriolae* are separated by short non-coding regions (<30 bp).

3.3. Protein-coding genes and ribosomal RNAs

The protein coding genes and ribosomal RNAs in *B. seriolae* are similar in length to those in *Gyrodactylus* spp. ATG is the typical start codon with the exception of *ND2* in *B. seriolae* which has a GTG start codon (Table 3). Stop codons varied between TAG and TAA and none was truncated. No gene overlaps were detected. Codon usage is shown in Table 4. There is a considerable A+T richness consistent with other invertebrate mt genomes with *Benedenia* showing the highest value amongst monogeneans by nearly 6% (Table 5). There is T and G skew potentially caused when strands remain single-stranded for unequal amounts of time during replication (Masta et al., 2008).

Table 3

List of annotated mitochondrial genes of *Benedenia seriolae* (see also Fig. 1).

Genes (protein)	Lengths bp	AmAcid	Codon Start	Stop	Position 5‟-3‟	Genes (RNAs)	Length (bp)	Position 5‟-3‟
COX3	648	216	ATG	TAA	1 – 648	tHis	64	649 – 653
CYTB	1089	363	ATG	TAG	720 – 1806	tThr	65	3450 – 3514
ND4L	228	76	ATG	TAA	1821 – 2048	tPhe	67	3516 – 3582
ND4	1215	405	ATG	TAA	2165 – 3379	tGln	63	3583 – 3645
ATP6	510	170	ATG	TAG	3720 – 4229	tMet	55	3665 – 3719
ND2	861	287	GTG	TAA	4230 – 5090	tVal	70	5098 – 5167
ND1	885	295	ATG	TAG	5299 – 6183	tAla	63	5170 – 5232
ND3	351	117	ATG	TAA	6449 – 6799	tAsp	67	5232 – 5298
COX1	1590	530	ATG	TAA	6927 – 8516	tAsn	64	6187 – 6250
COX2	582	194	ATG	TAG	10314 – 10895	tPro	63	6257 – 6319
ND6	450	150	ATG	TAA	10970 – 11419	tIle	65	6320 – 6384
ND5	1536	512	ATG	TAA	11774 – 13309	tLys	63	6386 – 6448
						tSer(AGN)1	52	6803 – 6854
						tTrp	65	6862 – 6926
						rrnL (16S)	982	8517 – 9498
						tCys	63	9499 – 9561
						rrnS (12S)	752	9562 – 10313
						tGlu	65	10905 – 10969
						tTyr	66	11426 – 11491
						tLeu(CUA)1	67	11501 – 11567
						tSer(UCA)2	64	11576 – 11639
						tLeu(UUA)2	69	11640 – 11708
						tArg	65	11709 – 11773
						tGly	68	13431 – 13498

Table 4

Nucleotide codon usage for 12 protein-coding genes of the mitochondrial genome of *Benedenia seriolae*.

Am Acid	Codon	No.	%	Am Acid	Codon	No.	%	Am Acid	Codon	No.	%	Am Acid	Codon	No.	%
Phe	TTT	405	12.22	Leu	CTT	61	1.84	Ile	ATT	197	5.94	Val	GTT	116	3.5
Phe	TTC	16	0.48	Leu	CTC	10	0.30	Ile	ATC	18	0.54	Val	GTC	6	0.18
Leu	TTA	363	10.95	Leu	CTA	43	1.30	Ile	ATA	157	4.74	Val	GTA	66	1.99
Leu	TTG	46	1.39	Leu	CTG	10	0.30	Met	ATG	60	1.82	Val	GTG	24	0.72
Ser	TCT	117	3.53	Pro	CCT	65	1.96	Thr	ACT	80	2.42	Ala	GCT	71	2.14
Ser	TCC	19	0.57	Pro	CCC	10	0.30	Thr	ACC	15	0.45	Ala	GCC	4	0.12
Ser	TCA	23	0.69	Pro	CCA	8	0.24	Thr	ACA	25	0.75	Ala	GCA	12	0.36
Ser	TCG	2	0.06	Pro	CCG	1	0.03	Thr	ACG	2	0.06	Ala	GCG	2	0.06
Tyr	TAT	168	5.07	His	CAT	48	1.45	Asn	AAT	113	3.41	Asp	GAT	48	1.45
Tyr	TAC	34	1.03	His	CAC	11	0.33	Asn	AAC	10	0.30	Asp	GAC	8	0.24
*	TAA	8	0.24	Gln	CAA	18	0.54	Asn	AAA	104	3.14	Glu	GAA	39	1.18
*	TAG	4	0.12	Gln	CAG	18	0.54	Lys	AAG	48	1.45	Glu	GAG	20	0.60
Cys	TGT	81	2.44	Arg	CGT	32	0.97	Ser	AGT	103	3.11	Gly	GGT	119	3.59
Cys	TGC	8	0.24	Arg	CGC	4	0.12	Ser	AGC	8	0.24	Gly	GGC	7	0.22
Trp	TGA	49	1.48	Arg	CGA	4	0.12	Ser	AGA	52	1.57	Gly	GGA	20	0.60
Trp	TGG	26	0.78	Arg	CGG	7	0.21	Ser	AGG	16	0.48	Gly	GGG	26	0.79

Table 5

Nucleotide content of protein-coding genes from complete mitochondrial genomes of monogeneans.

Species	Base Composition (%)					Total bp usage	Total No. codons
	T	C	A	G	A+T		
Gyrodactylus salaris	33.2	17.2	29.1	20.5	62.3	10050	3350
G. thymalli	35.6	15.6	23.3	25.5	58.9	9944	3314
G. derjavinoides	35.7	14.3	32.2	17.8	67.9	9945	3315
Microcotyle sebastis	41.8	10.1	27.8	20.3	69.6	10254	3418
Benedenia seriolae	48.18	9.99	27.13	14.7	75.31	9945	3315

3.4. Phylogenetic analyses

There were no premature stop codons within the coding regions of the protein coding genes. The 12 protein coding regions for each of the 33 taxa were concatenated for a total alignment of 6573 characters of sequence including 222 bp of *ATP6* (173 parsimony informative sites), 1392 bp of *COX1* (922 parsimony informative sites), 393 bp of *COX2* (287 parsimony informative sites), 285 bp of *COX3* (236 parsimony informative sites), 1032 bp of *CYTB* (792 parsimony informative sites), 696 bp of *ND1* (547 parsimony informative sites), 363 bp of *ND2* (327 parsimony informative sites), 243 bp of *ND3* (197 parsimony informative sites), 783 bp of *ND4* (665 parsimony informative sites), 192 bp of *ND4L* (169 parsimony informative sites), 684 bp of *ND5* (524 parsimony informative sites) and 288 bp of *ND6* (252 parsimony informative sites). As the mt genome of the „turbellarian" *Microstomum lineare* has been sequenced only partially, it was included with data for only five genes: *ATP6*, *COX1*, *COX3*, *CYTB* and *ND5*. *Neobenedenia* sp. has also only been partially sequenced (present study) and data for *ND5*, *ND4*, *ND4L*, *ND3*, *CYTB*, *COX3*, *COX2* and *COX1* were included. Some highly variable regions, particularly where some taxa had large insertions, could not be aligned unambiguously and so were removed (3657 bp in total).

We present the ML tree (Fig. 2) with non-parametric bootstrap proportions (BS) and posterior probabilities (BPP) as the Bayesian tree was identical in topology. Digenea and Cestoda are each monophyletic and sister groups (Fig. 2, node 4) with

all three nodes receiving strong support from both analyses. In contrast, the Monogenea does not form a single clade. Instead the single polyopisthocotylean, *Microcotyle sebastis* (Fig. 2, node 3) is the sister to the Digenea and Cestoda. Futhermore the Monopisthocotylea is not monophyletic with the Capsalidae (*Benedenia* and *Neobenedenia*) sister to the (Polyopisthocotylea (Digenea and Cestoda)) clade (Fig. 2, node 2) and the Gyrodactylidae is the sister to all of the other neodermatans (Fig. 2, node 1). Hypothesis testing of the two previously proposed alternative neodermatan relationships using AU tests (Table 6), show that our data do not support either the monophyly of Monogenea or a sister group relationship for the Cestoda and Monogenea (i.e. the Cercomeromorphae). The AU test of monopisthocotylean monophyly was not significant showing our data can not reject monophyly of the Monopisthocylea.

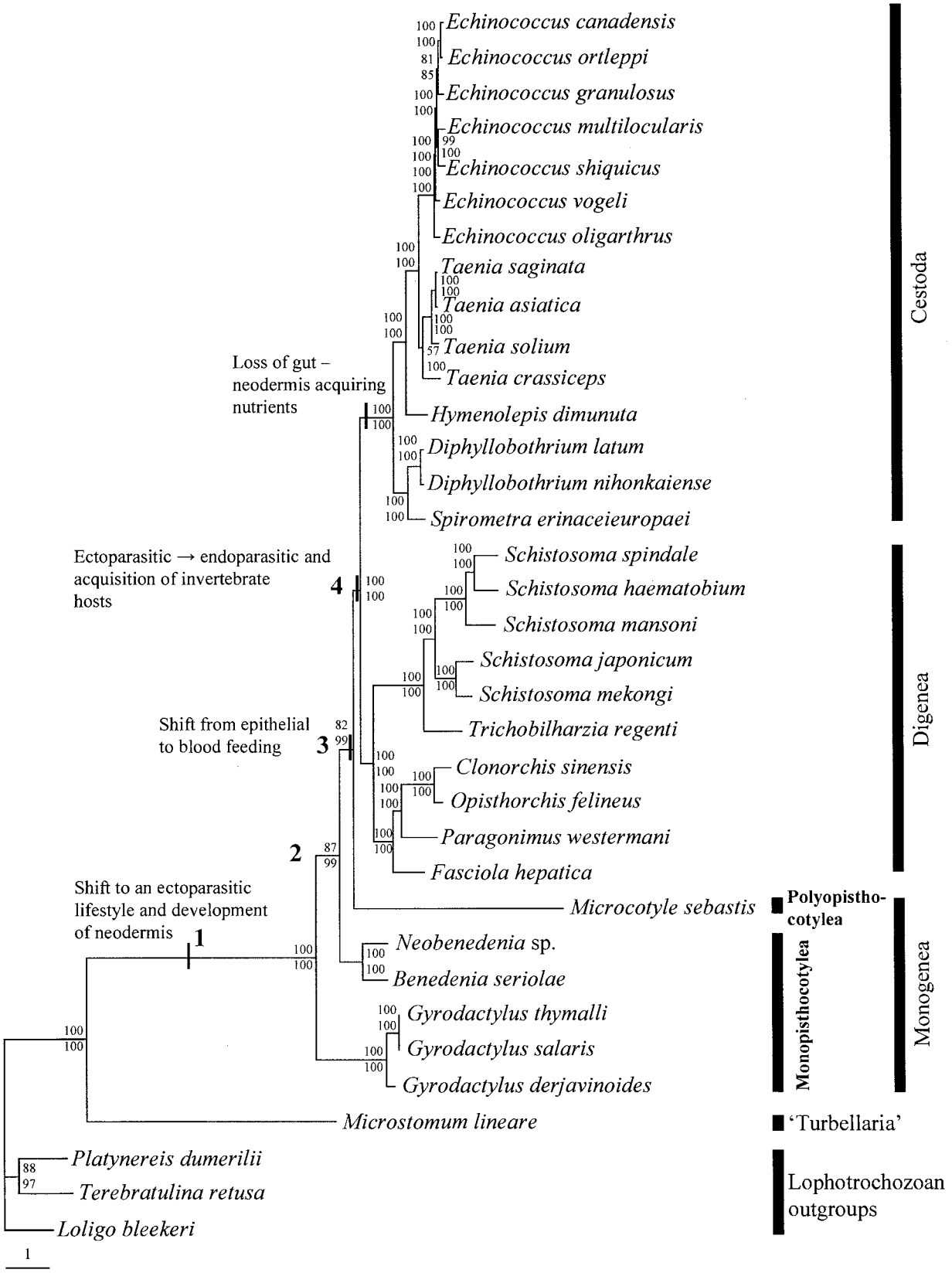


Fig. 2. A Maximum Likelihood (ML) tree from analyses of the 12 protein coding genes of the mitochondrial genomes of 32 Platyhelminthes. Posterior probabilities from Bayesian inference (above) and ML bootstrap proportions (below) the branches are indicated.

Table 6

Log Likelihood values and results from the Approximately Unbiased test for the alternative tree topologies.

Topology	-Ln	*P*-value ($\alpha = 0.05$)
Unconstrained	-139343.45	0.994
Monophyly of Monogenea	-139356.47	0.007
Cercomeromorphae (Monogenea and Cestoda as sisters)	-139362.94	0.001
Monophyly of Monopisthocotylea	-139350.05	0.098

4. Discussion

4.1. Systematic implications

Our phylogenetic analyses of neodermatan mitochondrial genomes produced three findings of note. First, our data reject the Cercomeromorphae (i.e. Cestoda+Monogenea) in favour of a Digenea+Cestoda clade. We acknowledge, however, that representatives of the Aspidogastrea are absent from our analyses but this is likely to have just a „local" affect on the nature of relationships among digeneans or between digeneans and cestodes. Previous molecular genetic analyses have provided equivocal resolution of the relationships among these groups, either supporting the Cercomeromorphae based on combined morphology and partial rDNA sequence coverage (Littlewood et al., 1999) or alternatively being consistent with our findings based on near complete rDNA nucleotide and complete mitochondrial genome sequences (Lockyer et al., 2003; Park et al., 2007). Thus the Digenea+Cestoda clade is supported by both of the largest molecular datasets and intriguingly has important consequences for understanding the evolution of parasitism, as it provides a far simpler explanation for the shift from ectoparasitism and the acquisition of multiple host life histories (see Littlewood et al. 1999), issues that we explore below.

Second, we found that Monogenea is paraphyletic, contradicting recent studies (e.g. Littlewood et al., 1999; Lockyer et al., 2003). The relatively few similarities between Monopisthocotylea and Polyopisthocotylea have been emphasised as evidence for the sister group relationship for decades. However, a poor understanding of character homology and the many morphological and ultrastructural differences emphasise their long independent evolutionary histories. This led Euzet and Combes (2003) to suggest that the lack of shared characters

between Monopisthocotylea and Polyopisthocotylea (e.g. Justine, 1991) implied paraphyly rather than simply lack of evidence for monophyly of the Monogenea (Justine, 1998). While the findings of Lockyer et al. (2003) based on near complete rDNA sequences emphasise monogenean monophyly, their taxon sampling was limited in phylogenetic scope (see below), which could have affected the arrangement of long branches towards the root of the tree.

Third, we found that Monopisthocotylea, a diverse group (Olson and Littlewood, 2002; Whittington, 2005), may not be monophyletic. This is in stark contrast to some previous tests of monophyly of the Monopisthocotylea and the morphological synapomorphies that unite the group (Justine, 1991) but is perhaps not surprising given the huge diversity of the Monopisthocotylea (see Whittington, 2005). However, clearly far greater sampling is required to determine if monopisthocotylean paraphyly is the case and if so, the extent. In contrast, the Polyopisthocotylea is regarded as a more cohesive group based on morphological, microhabitat, dietary, molecular genetic and spermatological characters (Mollaret et al., 2000) but monophyly of this subclass still requires testing.

The more phylogenetic hypotheses that are proposed from molecular genetic data, the more homoplasies are apparent among the current suite of morphological characters used in phylogenetic hypotheses for the Neodermata (e.g. Tkach et al., 2003; Palm et al., 2009; Perkins et al., 2009; Olson et al., 2010). Different parasites face similar selection pressures which has led to abundant convergent evolution thereby reducing the useful phylogenetic application of morphological data (Poulin and Morand, 2000). The debate about morphology versus molecules in phylogenetic analyses has all but become redundant with rapid advances in sequencing techniques expanding available sequence data exponentially. Faster than species are described, progress in sequencing techniques makes molecular data a far more rapid, efficient and cost-effective way to produce robust phylogenetic hypotheses (Rusk, 2009). While morphology may currently have limited applications in phylogenetics, its value cannot be underestimated. For example, morphological descriptions are vital to distinguish the plethora of monogenean species already described and to provide detailed descriptions of undescribed taxa. However, it is molecular genetic studies that have the power to provide the phylogenetic framework to investigate the evolutionary origins of parasitism, host associations and lifecycles and the

physiological and ecological adaptations for parasitism among the vast diversity of the Neodermata (e.g. Littlewood et al., 1999; Park et al., 2007).

4.2. The evolution of parasitism in the Platyhelminthes

Despite decades of study on the evolutionary relationships of Platyhelminthes, uncertainty remains about the origin of parasitism and evolution in the Neodermata (see Llewellyn, 1970; Littlewood et al., 1999). Different and inconsistent hypotheses about sister group relationships among the Neodermata have precluded informed inference about the evolution and radiation of the parasitic flatworms. Our data demonstrate paraphyly of the Monogenea and hint at paraphyly of the subclass Monopisthocotylea. Our establishment of phylogenetic relationships using complete mt genomes among neodermatans allows us to examine the evolution of parasitism in the Neodermata.

The common ancestor to the Neodermata is thought to be a free-living „turbellarian"-like omnivore, perhaps a scavenger or micropredator (Rohde, 1994). Transition from this free-living ancestor to a parasitic lifestyle required numerous adaptive modifications to the basic „turbellarian" plan resulting in specialisations of physiology, body form and behaviour (Littlewood et al., 1999). Adaptations that relate to food acquisition by neodermatans are among the most significant changes (Halton, 1997). Ectoparasitic, epithelial feeding Monopisthocotylea which parasitise fishes are the basal neodermatans in our analyses and transition from a „turbellarian"-like omnivorous ancestor requires no significant change in diet and digestive physiology. The most essential morphological innovation from a free-living ancestor was development of attachment organs to maintain permanent host contact (Fig. 2, node 1). The next step in the evolution of parasitism was a diet shift from epithelium (Monopisthocotylea) to blood (Polyopisthocotylea) (Fig. 2, node 3) and was associated with morphological differences that may be congruent with this food change. There are differences in the cellular structure of the gastrodermis in Monopisthocotylea and Polyopisthocotylea (see Smyth and Halton, 1983; Dalton et al., 2004). Furthermore, Polyopisthocotylea have a genito-intestinal canal that links the oviduct and right intestinal branch, which is absent in all Monopisthocotylea. While its function is not well understood, one hypothesis is a link to a blood diet (Euzet and Combes, 2003). Polyopisthocotyleans also produce anticoagulants, which

are not known in the Monopisthocotylea, to prevent clotting during blood feeding (Dalton et al., 2004).

The position of Polyopisthocotylea in our analyses implies a polyopisthocotylean-like ancestor for Digenea and Cestoda (Fig. 2, node 3). Phylogenetic analyses by Olson et al. (2003) show the basal group of Digenea contains many blood feeders with adults from several groups living inside the circulatory system (e.g. sanguinicolids = aporocotylids, spirorchids and schistosomes) further supporting a sanguinivorous ancestor. Schistosomes, amongst the basal Digenea, also produce anticoagulant proteins but homology with polyopisthocotylean anticoagulants has not been assessed. Other adult Digenea parasitise a broad range of microhabitats including the gut, organs associated with the gut or the lungs of all vertebrate classes. These microhabitat changes are associated with dietary diversification, an evolutionary possibility because Digenea possess a gut. Cestoda, considered the most derived of the Neodermata, live as adults almost exclusively in the gut or associated organs of all vertebrate classes, display the most dramatic shift in morphology and diet through loss of the gut and use the neodermis to acquire nutrients from host gut contents (Halton, 1997).

The neodermis is a key evolutionary adaptation in the Neodermata. The cestode neodermis has active transport systems that facilitate the selective uptake of nutrients, allowing them to compete actively for food with the host mucosa (Dalton et al., 2004). The dietary role of the neodermis is not well understood in Digenea and Monogenea but studies have shown they are capable of acquiring some nutrients across the neodermis to supplement food digested by the gut (Dalton et al., 2004). In an in vitro study on *Diclidophora merlangi* (Polyopisthocotylea) in which the foregut of worms was ligated to prevent nutrient acquisition via the mouth, unligated worms had similar uptake of trace-labelled substrates compared to ligated worms supporting hypotheses of facilitated diffusion across the neodermis (Smyth and Halton, 1983). Diffusion across the neodermis has not been assessed in Monopisthocotylea and the relative importance of nutrient acquisition across the tegument in Monogenea is unknown. Studies on the digeneans *Fasciola hepatica* and *Schistosoma mansoni* show the neodermis may be the principal route of absorption for low molecular weight solutes such as glucose and that *S. mansoni* can survive for prolonged periods without a functional gut (Halton, 1997). However, active transport of nutrients across the neodermis has not been demonstrated in either Digenea or Monogenea.

The neodermis is a defining structure of the Neodermata, a specialisation important in the evolution of diet, but research is needed to fully understand its role and the potential evolutionary advantage it could confer for nutrient absorption for more basal groups of Neodermata. The complete nuclear genomes of *S. mansoni* and *S. japonicum* were recently sequenced (Berriman et al., 2009; Zhou et al., 2009). Comparisons of these schistosome genomes to that of the sea anemone *Nematostella vectensis* revealed an expansion of proteases in schistosomes which may be directly related to parasitism because gene families that were more numerous included those that related to blood feeding (cathepsins; Berriman et al., 2009). Whether these genes related to blood feeding occur in all neodermatan groups is unknown but their wide presence and a shared function may provide further insights into the evolution of parasitism across the Neodermata.

4.3. Concluding comments

Historically, the debate about monogenean monophyly is perhaps a reflection of our desire to focus on perceived similarities between taxa and therefore we wrongly overlook the differences. It is human nature to try to categorise organisms neatly but it seems inevitable that more molecular evolutionary and phylogenetic studies will further blur the lines between categories that were once perceived as distinct. We must accept that many traditional morphological classifications of parasitic groups may fail to define discrete, monophyletic groups. The arrangement of haptoral sclerites in larvae and adults has been used extensively to propose hypotheses about monogenean evolution (e.g. Bychowsky, 1957; Llewellyn, 1970). Arguments against the homology of haptoral sclerites in Monopisthocotylea and Polyopisthocotylea have been advanced (Pariselle and Euzet, 1995) together with other differences such as the larval osmoregulatory system (Euzet et al., 1995) and in the diet and gut structure of adults (Euzet and Combes, 2003). Until character homology among Monogenea is better understood through intensive embryology, ontogeny, larval and adult studies, elucidation of their phylogenetic relationships likely remains the domain of molecular genetics. There is tremendous species diversity and great evolutionary depth among platyhelminths in general and the parasitic Neodermata in particular (e.g. Littlewood et al., 2004). The prospect of overcoming issues highlighted by Littlewood et al. (1998) concerning lack of correspondence of taxon representation among not only morphological but also

molecular studies seems unlikely to be resolved without a focus provided by well co-ordinated approaches. Examples include the Second International Workshop for Tapeworm Systematic (Hoberg et al., 1999), funded by large scale global research support schemes such as Assembling the Tree of Life (Cracraft and Donoghue, 2004) and Planetary Biodiversity Inventories programs. We should apply some balance here as the three most robust molecular genetic analyses of neodermatan relationships conducted to date each rely, effectively, on a single gene history, i.e. Lockyer et al. (2003) used two closely linked rDNA genes, and Park et al. (2007) and the present study on the non-recombining haploid mtDNA molecule. Despite very large numbers of informative sites (5091) for the mtDNA dataset analysed here, some caution is required (e.g. Castoe et al., 2009) until our hypothesis can be tested with nuclear datasets with at least equivalent numbers of informative characters, a situation only likely to be addressed readily through the developing field of phylogenomics (e.g. Dunn et al., 2008).

Acknowledgements

This study was funded by an Australian Research Council Discovery grant (DP0556780) awarded to IDW and SCD. EMP was supported by an Australian Postgraduate Award.

When the worms turned: molecular dating for diversification of the parasitic Platyhelminthes, Monogenea and Capsalidae (Monopisthocotylea)

Elizabeth M. Perkins [a,*], Steve C. Donnellan [b,c], Terry Bertozzi [b] and Ian D. Whittington [a,c,d]

[a]*Marine Parasitology Laboratory, School of Earth and Environmental Sciences (DX 650 418), The University of Adelaide, North Terrace, Adelaide, South Australia 5005, Australia*

[b] *The South Australian Museum, North Terrace, Adelaide, South Australia 5000, Australia*

[c]*Australian Centre for Evolutionary Biology and Biodiversity, The University of Adelaide, North Terrace, Adelaide, South Australia 5005, Australia*

[d]*Monogenean Research Laboratory, Parasitology Section, The South Australian Museum, North Terrace, Adelaide, South Australia 5000, Australia*

*Author for correspondence (elizabeth.perkins@adelaide.edu.au)

This chapter is formatted in a style appropriate for submission to the *Proceedings of the Royal Society B – Biological Sciences*.

There are currently no independent estimates of dates for the origin and radiations of the parasitic Platyhelminthes (Neodermata). Using *Elongation Factor 1 α* (*EF1α*) data from 51 Capsalidae species, 10 species in five other monogenean families, seven other platyhelminth species in combination with sequence data and fossil calibrations from 23 metazoan groups we estimate molecular clock dates for the radiation of the Neodermata, Monogenea and Capsalidae. The Neodermata diverged from the free-living Platyhelminthes approximately 513 million years ago (mya) (95% HPD (highest posterior density): 473–605) predating the appearance of vertebrates in the fossil record. This suggests the origins of the parasitic platyhelminths may be on invertebrate hosts. We estimate the Monogenea diverged from the other neodermatan classes 441 mya (95% HPD: 420–547) a similar time to the appearance of fish in the fossil record. Our molecular dating analyses indicate the most recent common ancestor of the Capsalidae arose approximately 235 mya (95% HPD: 200–274) following the massive Permian/Triassic extinction event. Their origin coincides with the diversification of marine bony fishes. These molecular dates are the first independent estimates of neodermatan, monogenean and capsalid diversification and provide valuable insights into the radiation of the parasitic Platyhelminthes.

1. INTRODUCTION

Estimating divergence dates is an important component of understanding the mode and tempo of the diversification of biological radiations. Many organisms have extensive fossil records that allow accurate dating of their origin and subsequent radiations. Their diversity can be examined in relation to significant geological events and known extinctions. Soft bodied organisms that lack hard structures are recorded rarely, if ever, in fossil deposits and so molecular clock estimates of divergences provide the only way to date their origins and radiations. Parasitic Platyhelminthes (Neodermata) are generally small, soft bodied organisms that lack a useful fossil record. Molecular clock studies, using multiple nuclear loci, which have focused on the origin of the early metazoans, suggest an origin for the Platyhelminthes 500–600 million years ago (mya) (Peterson *et al*. 2004, 2008). However the timing of the origins and subsequent radiation of the parasitic flatworms was not explored in these studies.

The Capsalidae (Neodermata: Monogenea), a large (approx. 180 described species) and diverse family of ectoparasites of marine fish has a global distribution

with species recorded across marine fish lineages (Chondrichthyes, Acipenseriformes and a broad array of teleosts). This is exceptional among Monogenea with most families restricted to either Chondrichthyes or teleosts. The broad host associations of the Capsalidae make them a fascinating focal group for molecular dating techniques to examine their origin and radiation. Molecular phylogenetic analyses of the Capsalidae by Perkins *et al*. (2009) used three unlinked nuclear genes to assess phylogenetic relationships within the family. Here we extend our work with the addition of 31 capsalid taxa and mitochondrial (mt) gene sequence data. Our analyses are based on 78 capsalid taxa in 28 genera representing all nine subfamilies and also include 30 outgroup taxa in 13 families from the Monopisthocotylea and Polyopisthocotylea. A large outgroup representation is required as the sister taxon to the Capsalidae is unresolved. Analyses combine partial sequence data for *28S rDNA*, *Histone 3* (*H3*) and *Elongation Factor 1* α (*EF1*α), *Cytochrome Oxidase 3* (*COX3*) and *Cytochrome B* (*CYTB*). We use our *EF1*α dataset for 51 capsalids and 10 other monogeneans with that of other basal metazoans which have fossil records to estimate divergence times for the parasitic playhelminthes, Monogenea and the Capsalidae using molecular clock dating techniques. We increase the taxon sampling and array of genes sequenced for the Capsalidae to improve our ability to resolve parasite relationships and to be more certain that we sampled the major capsalid lineages in order to better estimate the age of the root of the family. Using this improved phylogenetic hypothesis, we constrain topologically a tree of *EF1*α sequences to develop a molecular clock based chronology of diversification for the Neodermata, Monogenea and in particular the Capsalidae.

2. MATERIALS AND METHODS

(a) *Taxon selection*

Specimen and voucher deposition details are presented in Appendix V. For the current taxonomic classification of the Capsalidae see Table 1 in Perkins *et al*. (2009). Trees were rooted with 22 monopisthocotylean species from nine families and eight more distant polyopisthocotylean species in six families (Appendix V).

(b) *Molecular genetic methods*

DNA was extracted according to the Gentra Kit (Gentra Systems) protocol for animal tissues preserved in ethanol. PCR amplification of the nuclear loci *28S rDNA*, *H3* and *EF1α* sequence was carried out following Perkins *et al*. (2009). The mt genes, *COX3* and *CYTB*, were amplified using primers listed in Appendix VI, Table 1. Amplifications, in 25 μL reactions, and sequencing reactions using the PCR primers followed Perkins *et al*. (2009). Samples were sequenced on an Applied Biosystems 3730 DNA sequencer. Additional materials and methods are in Appendix VI.

(c) *Phylogenetic analyses and hypothesis testing*

Edited sequences were aligned initially using Clustal X (Thompson *et al*. 1997) and manual adjustments were made using inferred amino acid sequences where applicable (*H3*, *EF1α*, *COX3* and *CYTB*). Two alignments were used for analyses, the first comprising only nuclear genes, and a second using all five genes, in order to examine the effect of the mt genes on the phylogenetic hypotheses generated. The taxon composition of the two alignments is specified in Appendix V. All sequences have been deposited in GenBank (see Appendix V for accession numbers).

ModelTest, preliminary and final Bayesian phylogenetic analyses methods followed Perkins *et al*. (2009). Maximum likelihood (ML) analyses of datasets were run in RAxML (Stamatakis 2006; Stamatakis *et al*. 2008) with data partitioning using the default rapid hill climbing algorithm. Bootstrap proportions were estimated under the same conditions for 100 pseudoreplicates.

(d) *Molecular Clock Dating*

EF1α sequence data for monogeneans were combined with *EF1α* data from 28 selected metazoan groups including seven other platyhelminth species and two outgroups from the Choanoflagellida and Viridiplantae (Appendix VII, Table 1). A fully resolved constraint tree was constructed based on phylogenetic hypotheses of the Capsalidae from our study and major metazoan groups of Dunn *et al*. (2008) and Peterson *et al*. (2004, 2008). ML analyses of the *EF1α* sequence data were performed to confirm these data retrieved the nodes we were interested in estimating dates for. Different alignments and fossil calibrations were used to infer dates of divergence for the Capsalidae. As the *EF1α* sequence data produced for 51 capsalid taxa in these analyses is shorter than that of the available metazoans sequences (Appendix VII, Table 1) two alignments were used: a longer alignment (960 bp) with the sequences

from this study included as missing for some parts of the alignment and a shorter alignment (573 bp) in which available downloaded sequences of metazoans were shortened to the length of sequences produced in this study. BEAST analyses (Drummond & Rambaut 2007) implemented an uncorrelated relaxed clock method, which assumes an overall distribution of rates across branches but does not assume that the rates on adjacent branches are autocorrelated. We implemented the log normal rate distribution and calibrations were treated as exponential priors including soft maxima with 95% of their density lying between the uniform maximum and minimum. There are three fossil parasitic platyhelminths that are relevant calibrations (Combes 2001) but given the inadequacy of the parasitic platyhelminth fossil record, these calibration points were treated with much uncertainty. A fossil trematode from 10 mya was not used for calibration due to the high likelihood that this class is considerably older (Combes 2001). The monogenean fossil from 400 mya (Upeniece 2001) was used as a maximum age for the Capsalidae and as a minimum age for the Neodermata. The tapeworm fossil (Combes 2001) was used as a minimum age for the Cestoda. The parasite fossils were used to calibrate two nodes in some analyses and were compared to analyses that did not use these calibrations points. Due to the high likelihood that these fossil records represent neither maxima nor minima for these groups, calibrations had large exponential means. Fossil calibrations for another 15 metazoan nodes were taken from Peterson *et al*. (2004, 2008). Burnin and 95 % highest posterior densities (HPD) for results were calculated using the program Tracer v1.4 (Rambaut & Drummond 2007).

3. RESULTS

(a) *Phylogenetic relationships of monogeneans*

Patterns of DNA sequence characteristics, phylogenetic information content, size of data partitions and optimal data partitioning are presented in Appendix VI. The Bayesian 50% majority rule consensus tree of the complete dataset is shown in Appendix VIII along with posterior probabilities (PP) and as the ML analyses produced very similar results, the ML bootstrap proportions (BS). The Bayesian 50% majority rule consensus tree of the nuclear dataset is in Appendix IX for comparison. Analyses of the two datasets produced largely similar results (compare Appendices VIII & IX) with some differences in outgroup arrangement. Analyses of the nuclear

data place the Gyrodactylidae and Udonellidae as sister group to the Capsalidae whereas analyses of the complete dataset leave the sister group unresolved as indicated by a polytomy of the Capsalidae, Microbothriidae+Monocotylidae (clade 3, Appendix VIII) and Gyrodactylidae+Udonellidae (clade 4; Appendix VIII). The arrangement of *Anoplodiscus* sp., Amphibdellatidae sp. Dactylogyridae sp. and Calceostomatidae sp. (clade 5) also varies between analyses (Appendix VIII and IX). The polyopisthocotylean outgroups form a strongly supported group in both analyses (Appendix VIII and IX, clade 6, PP 100%, BS 99%). Monophyly for the Capsalidae is strongly supported by the nuclear data (PP 100%) and the ML combined data analyses (Appendix X, BS 85%) but weaker support in the combined data Bayesian analysis (Appendix VIII, PP 86%). The Capsalidae is characterised by short internal branches, often poorly supported internal nodes and longer branches at the tips with strongly supported nodes. In Bayesian analyses of the combined dataset, capsalids split into two major groups from the base of the family (Appendix VIII, clade 1 and clade 2) but these were not strongly supported (PP 73% and 85%, BS 61% and 64%). This split is not seen in analyses of the nuclear data alone (Appendix IX). Clade 1 (Appendix VIII) comprises taxa from four subfamilies (Benedeniinae, Encotyllabinae, Pseudonitzschiinae and Trochopodinae), 12 genera and 17 unattributed taxa. Clade 1 further splits into multiple, mostly strongly supported clades – A to E (Appendix VIII). The Encotyllabinae forms a strongly supported clade (PP 100%, BS 100%) as does *Neobenedenia* (clade 1E, PP100%, BS 100%). The Benedeniinae and Trochopodinae are polyphyletic.

Clade 2 (Appendix VIII) comprises taxa from seven subfamilies (Benedeniinae, Capsalinae, Dioncinae, Entobdellinae, Interniloculinae, Nitzschiinae and Trochopodinae), 16 genera and two unattributed taxa. Clade 2 divides into four major clades – A to D (Appendix VIII). The Capsalinae is monophyletic (clade 2A, PP 100%, BS 100%) with the Nitzschiinae as sister group (PP 87%, BS 70%). The Entobdellinae (clades 2B and 2C) was polyphyletic, forming multiple, distinct clades that also included members of the Benedeniinae and Trochopodinae. Clade 2C in not strongly supported (PP 68%, BS 21%) and comprises benedeniines and entobdellines. With the exception of *Calicobenedenia polyprioni* all species in clade 2C infect elasmobranch hosts. Most basal in Clade 2 is clade 2D comprising members from Dioncinae, Interniloculinae and Trochopodinae with included taxa infecting teleosts.

(b) ***Molecular clock dating***

The shorter *EF1α* alignment consistently gave slightly older dates on average but confidence intervals overlapped with results from the longer *EF1α* alignment (figures 1A, 1B). Analyses that included two neodermatan fossils as calibrations gave similar results to those without the neodermatan calibration points but with narrower confidence intervals. Results from the longer alignment with the parasite fossil calibrations included are presented in figures 1A and 1B.

Our estimates of dates of metazoan nodes are consistent with results from Peterson *et al*. (2004, 2008) indicating confidence in methodology and results. Molecular dating estimates place the split of the Platyhelminthes and polychaetes+molluscs at 521 mya (95% HPD: 534–656) with the Neodermata versus „Turbellaria" split at approximately 513 mya (95% HPD: 473–605). The Monogenea diverged from the other neodermatan classes approximately 441 mya (95% HPD: 420-547). The split between the major monogenean lineages, Monopisthocotylea and Polyopisthocotylea (*Microcotyloides incisa*), dates to 393 mya (95% HPD: 379–506). The estimated divergence of the Gyrodactylidae + Undonellidae clade versus Capsalidae is approximately 312 mya (95% HPD: 274–396). Estimated age of the extant capsalid diversity is approximately 235 mya (95% HPD: 200–272), younger than the split between the Gyrodactylidae and Udonellidae at 251 mya (95% HPD: 153–315). The two major radiations of capsalids are similar in age: 227 mya (95% HPD: 190–270) and 232 mya (95% HPD: 185–264), respectively. There are recent radiations in some genera (e.g. *Encotyllabe* >10 mya)

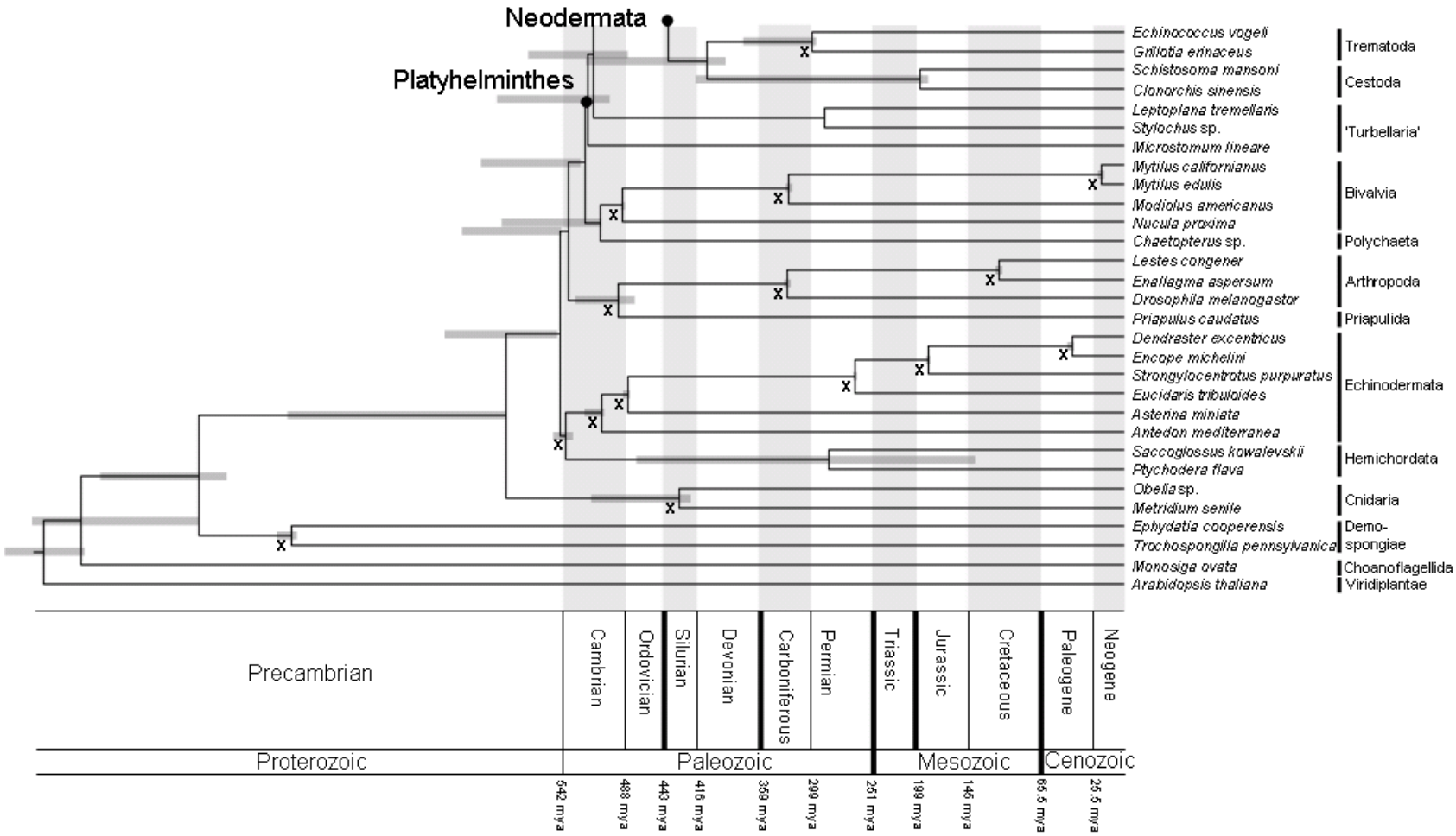


Figure 1A. Basal part of BEAST maximum credibility ultrametric tree for 61 monogenean taxa (51 capsalid species, 2 gyrodactylids, 1 udonellid, 4 microbothriids, 2 monocotylids and 1 microcotylid species) representing 6 families and both subclasses, seven platyhelminth species, 21 early metazoan groups and two outgroups (Viridiplantae and Choanoflagellida). Node bars indicate highest posterior densities. Timescale = millions of years ago (mya). X indicates nodes that were fossil calibrated from Peterson et al. (2004, 2008). Bold lines on timescale indicate major extinction events.

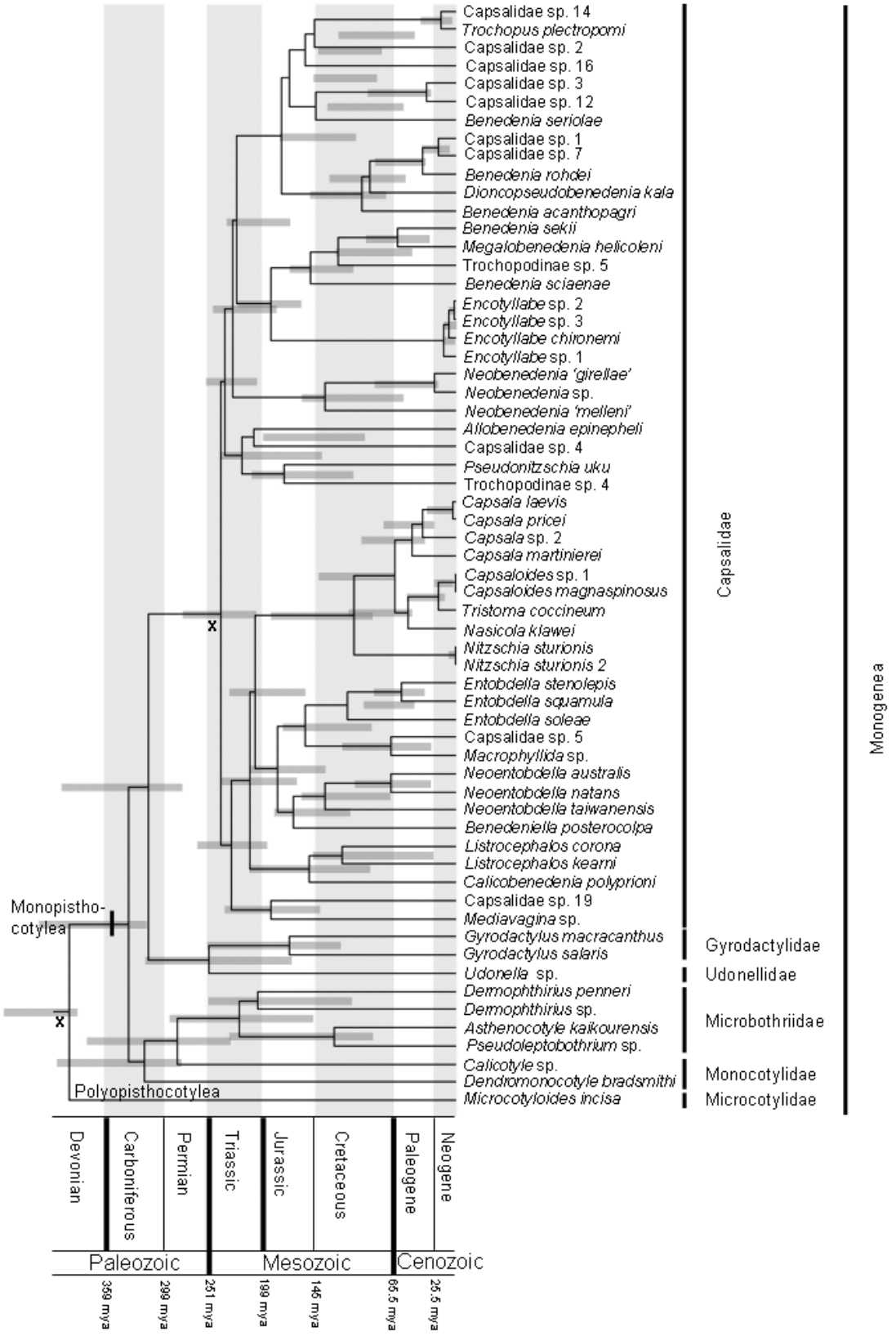


Figure 1B. Monogenena part of BEAST maximum credibility ultrametric tree for 61 monogenean taxa (51 capsalid species, 2 gyrodactylids, 1 udonellid, 4 microbothriids, 2 monocotylids and 1 microcotylid species) representing 6 families and both subclasses, seven platyhelminth species, 21 early metazoan groups and two outgroups (Viridiplantae and Choanoflagellida). Node bars indicate highest posterior densities. Timescale = millions of years ago (mya). X indicates nodes that were fossil calibrated from Upeniece (2001) and Combes (2001). Bold lines on timescale indicate major extinction events.

4. DISCUSSION

(a) *Radiation of the Neodermata*

Timing of the radiations of the different groups of parasitic Platyhelminthes have been inferred largely from the age of host lineages and assumptions about when the parasites acquired their vertebrate and invertebrate hosts (Boeger & Kritsky 1997; Littlewood *et al*. 1999a). Dates often show congruence between parasite and host divergences, as this has been a prior assumption placed on previous estimates. No independent test of divergences using molecular dating techniques has been attempted for Neodermata. No adequate fossil record for parasitic Platyhelminthes exists and is highly unlikely to ever be more than an occasional rare find (e.g. Combes 2001). Peterson *et al*. (2004, 2008) estimated the origin of the Platyhelminthes between 500-600 mya and our estimates are in agreement. Peterson *et al*. (2004, 2008) did not include any parasitic platyhelminths and so provided no estimate of the origin dates for Neodermata. We estimate the Neodermata diverged from the free-living platyhelminths approximately 513 mya (95% HPD: 473-605) and that the Monogenea diverged from the other neodermatan classes approximately 441 mya (95% HPD: 420-547). This places the origins of neodermatans in the Cambrian.

Most major types of invertebrate lineages appear in the Cambrian fossil record but vertebrates are absent from the record implying that vertebrate ancestors were either extremely rare or absent from the salt water in which most Cambrian deposits were laid down (Romer 1945; Janvier 1996a). This suggests the origins of the parasitic platyhelminths may have been on invertebrate hosts, perhaps not acquiring vertebrate hosts until they appeared later in the Paleozoic. This is in contrast with other hypotheses which suggest the ancestor to the Neodermata gained a vertebrate host first and later acquired invertebrate hosts in the Trematoda and Cestoda (see Littlewood *et al*. 1999a). The Monogenea diverged from the other neodermatan classes in the Silurian, coinciding with the appearance of vertebrates in the late Silurian (Romer 1945; Benton 1990). While there is limited evidence of jawed vertebrates in Silurian deposits the diversity present in the Devonian suggests their development during the Silurian (Romer 1945; Zhu *et al.* 2009). The Devonian, referred to as the age of fishes, saw lower fish groups radiate rapidly with frequent records in Devonian deposits. Much Devonian fish life is recorded from freshwater deposits. Ostracoderms and jawed fishes were abundant but no Chondrichthyes or

higher bony fishes have been identified (Romer 1945). Placoderms were abundant and had already diversified and flourished in lower Devonian seas. In the late Devonian, Osteichthyes appeared in fresh water deposits and coincided with a sharp decline among the lower fish groups. Crossopterygians, primitive coelacanths, dipnoans and primitive actinopterygians are also present in fresh water deposits (Romer 1945). The Devonian contained many important events in the evolutionary history of fish: decline and fall of ostracoderms; rise and decline of acanthodians; rise and disappearance of arthrodires; appearance and extinction of antiarchs; rise of shark like forms and the appearance of bony fishes (Romer 1945; Janvier 1996a). It is impossible to know if the ancestors of monogeneans arose on marine fishes but given the majority of early fish diversification appears to be in fresh water, it is possible that the ancestors of the monogeneans were hosted by freshwater fishes. Molecular dating analyses including all families of Monogenea will provide valuable insights to explore this.

(b) ***Radiation of the Capsalidae***

In the late Permian and its boundary with the Triassic (approximately 251 mya), the Earth‟s most severe extinction event occurred (Bambach *et al*. 2004). It is estimated that up to 96 % of marine species went extinct (Benton 2003). We estimate the age of the Capsalidae as slightly younger than this indicating that radiation of these monogeneans may have followed this major marine extinction event (Bambach *et al*. 2004). After an extinction event, extensive habitats and niches become available for the radiation of surviving lineages. Patterns and timing of radiations of the parasitic Platyhelminthes cannot be tracked through geological time due to the near absence of a fossil record. However, fish can be followed through the record and given the dependence of capsalid parasites on their marine fish hosts, it is reasonable to assume that their evolutionary radiations have been intertwined. Marine fish records in the Permian are not extensive with only a few chondrichthyan groups and little evidence of marine bony fishes (Romer 1945). During the Permian/Triassic extinction, marine vertebrate fauna was almost entirely wiped out. The fossil record for some chimaeras, hybodonts, primitive chondrosteans and one family of Actinopterygii (Semionotidae) crosses the Permian/Triassic boundary suggesting much of the extensive fish radiations following this mass extinction event arose from these groups (Sepkoski 1982). There was a strong trend among ray-finned fishes towards a marine

existence during the Triassic marking the beginning of the oceans as a major focus for actinopterygian diversity (Romer 1945). The origins of the Capsalidae coincide with this shift to the marine environment and extensive radiation of fishes and no extant capsalids are known from freshwater fishes (Whittington 2004).

Neoptergygian fishes appeared in the middle Triassic (245–228 mya) when the extant groups of sharks and rays also appeared (Nelson 2006). During the Jurassic (200–146 mya), the early representatives of bony fishes appeared and by the Cretaceous (145–65 mya), rays, sharks and teleosts are common in the fossil record (Nelson 2006). The diversification of the Capsalidae estimated here at ~235 mya (95% HPD: 200-272) is temporally congruent with the diversification of fishes. There have been two other major extinction events in the history of the Capsalidae (Triassic/Jurassic and end of Cretaceous). While these did not have as great an impact on the marine environment as the Permian/Triassic extinction (Raup and Sepkoski 1982), they may have resulted in the extinction of many parasite and host lineages. The long evolutionary history of the Capsalidae means there may have been many extinction events, various host switching events and long distance dispersals making it difficult to understand the past interrelationships of parasites and hosts based on extant taxa.

Our independent estimates of divergence dates for the Neodermata allow their evolutionary history to be explored without any bias from prior assumptions based on host origins and diversification. The parasitic Platyhelminthes have a long evolutionary history with their origins estimated to date back to the early Paleozoic. In contrast to previous hypotheses (Littlewood *et al*. 1999a), this suggests their diversification may have begun on invertebrate hosts only later acquiring vertebrate hosts when they appeared late in the Silurian. However, monogenean origins do coincide with the appearance of fish in the fossil record and given the diversity of freshwater fishes, Monogenea may have arisen in fresh water. The Capsalidae, known exclusively from marine fish hosts, are estimated to have arisen in the early Triassic when bony fishes began rapidly diversifying in the marine environment. Our analyses provide the first insights into the diversification of the Neodermata through time and while results are tentative, they are an independent assessment of neodermatan evolution.

We thank the people listed in Appendix V for supplying us with specimens. Our study was funded by an Australian Research Council Discovery Grant (DP0556780), a JD Smyth Travel Award from the Australian Society for Parasitology, Inc. and an ARC/NHMRC Research Network for Parasitology Travel Award. E.M.P. was supported by an Australian Postgraduate Award.

Supplementary data

Supplementary data associated with this chapter is shown in Appendices V, VI, VII, VIII, IX and X.

CHAPTER V

COLLECTING SOUVENIRS: FAMILY OF ECTOPARASITIC PLATYHELMINTHES (MONOGENEA: MONOPISTHOCOTYLEA) DISPLAYS EXTENSIVE HOST SWITCHING ACROSS MARINE FISHES

Elizabeth M. Perkins,[1,2] Terry Bertozzi,[3] Ian D. Whittington,[1,4,5] Steve C. Donnellan[3,4]

[1]*Marine Parasitology Laboratory, School of Earth and Environmental Sciences (DX 650 418), The University of Adelaide, North Terrace, Adelaide, South Australia 5005, Australia*

[2]*E-mail: elizabeth.perkins@adelaide.edu.au*

[3]*Evolutionary Biology Unit, The South Australian Museum, North Terrace, Adelaide, South Australia 5000, Australia*

[4]*Australian Centre for Evolutionary Biology and Biodiversity, The University of Adelaide, North Terrace, Adelaide, South Australia 5005, Australia*

[5]*Monogenean Research Laboratory, Parasitology Section, The South Australian Museum, North Terrace, Adelaide, South Australia 5000, Australia*

This chapter is formatted in a style appropriate for submission to *Evolution*.

Understanding the diversification and radiation of parasitic groups through time requires examining parasite relationships and the diversification and radiation of their hosts. The Capsalidae (Platyhelminthes: Monogenea: Monopisthoctylea) is a diverse family of ectoparasites of marine fish. Their host range includes all major lineages of fishes, from Chondrichthyes to a broad range of teleosts. The radiation of the Capsalidae across their diverse marine fish hosts has not been explored using robust parasite-host phylogenies. Phylogenetic relationships amongst and within the 62 orders of fishes remain unresolved due to difficulties in finding sufficient phylogenetically informative characters. We present multilocus phylogenetic analyses of fishes based on seven nuclear and three mitochondrial genes for 61 orders of fishes (including capsalid host taxa) and use fish fossil data to calibrate the phylogeny for molecular dating analyses. Our phylogenetic hypothesis for fishes indicates widespread polyphyly in the Beloniformes, Gasterosteiformes, Perciformes, Pleuronectiformes and Scorpaeniformes. Comparisons of parasite and host phylogenies and chronograms show no topological or temporal congruence indicating capsalid radiation may be independent of host radiation. Extensive host switching, within phylogenetic and ecological constraints, may explain the distribution of capsalids across their diverse marine fish hosts. Further biological studies of parasites and hosts are required to explore these interactions.

Parasites spend much of their lives in intimate associations with their hosts. Their host species form their primary environment and so play an important role in the evolution and diversification of parasite species. Coevolution is the most commonly explored phylogenetic relationship between a parasite and its host. Coevolution is where parasite speciation events closely follow those of their hosts and, consequently, their phylogenies show congruence (Charleston and Perkins 2006). Parasite phylogeny mirroring host phylogeny was first proposed by Fahrenholz and became known as Fahrenholz‟s rule (Fahrenholz 1913). The parasite phylogeny must mirror the host phylogeny not just in topology but also in temporal sequence (Page 1996). Parasites that follow this coevolutionary path can be thought of as biological „heirlooms‟ (Banks and Paterson 2005). Many parasites are host specific, infecting only a single host species. One of the best known cases of close parasite-host association is the North American pocket gophers and their chewing lice (Hafner and Page 1995). The ecology of pocket gophers means interspecific interactions are rare

and their chewing lice are incapable of surviving long periods separated from their hosts perhaps leading to the high instances of coevolution (Hafner and Page 1995). However many other studies of parasite and host associations have shown that coevolution may be the exception rather than the rule (Paterson and Poulin 1999; Weckstein 2004; Morand et al. 2008). As the number of studies of relationships of other parasite-host groups has increased, Fahrenholz"s rule has become known as Fahrenholz"s fallacy, with few comparisons of parasite-host relationships exhibiting high levels of coevolution (Page and Charleston 1998). There are multiple types of events that could explain the apparent absence of coevolution (Paterson and Banks 2001). First, a parasite might „miss the boat" by simply being absent from the host population that speciates, which may occur if parasites are patchily distributed and at low prevalence across their host range. Second, „drowning on arrival" is an extinction event where the parasite was present on the speciating host population but went extinct perhaps due to an extreme bottleneck that may take place in parasites that occur at low intensities (Paterson and Banks 2001). A third explanation is that the parasite may exist on the host population but has not yet been recorded from that host species despite being sampled for parasites. Insufficient sampling can occur if the parasite is present at low prevalence and intensities. Host switching events occur when a parasite successfully colonises a host species other than its current host and then speciates. These parasites can be thought of as „souvenirs" (Banks and Paterson 2005). Distinguishing between these events is difficult when presented with an extant group of parasites because different events could explain the observed distribution of parasites across their hosts. Extensive taxon sampling is required to identify different events confidently (Paterson and Banks 2001). One event is unlikely to explain the distribution of a parasite species because ecological, biological and morphological factors also play a role in shaping parasite diversity.

Parasite-host studies typically focus within a genus of parasites or hosts and are often limited to parasites that have restricted host and/or microhabitat ranges (e.g. Desdevises et al. 2002; Miller and Cribb 2007). There have been few studies that have examined the radiation of a family of parasites across its entire host range. The Capsalidae (Platyhelminthes: Monogenea: Monopisthocotylea) is a large family of ectoparasites of marine fish, diverse in host association and microhabitat range (Whittington 2004) and thus provide a model family to examine the radiation of a large lineage across its hosts through time. The family has a worldwide distribution

occurring in tropical, temperate and cold waters. As in other monogenean families, the majority of species exhibit strict host specificity and all have direct lifecycles (no intermediate host is required). Individual capsalid species also often display strict site specificity but different species parasitise a diversity of sites including gills, oral cavity, pharyngeal tooth pads, nares and fins. Capsalids have a free-swimming larval stage which must find or encounter its host and then attach to it. The larvae may then move to the specific site of adult attachment (Whittington 2004). Prevalence and intensity of infections on wild hosts can vary considerably between taxa.

Morphologically the group is characterised by a simple, flattened, leaf-like body plan with a reduction in morphological complexity, typical of parasites (Whittington 2004). Important morphological characters are considered to be the anterior attachment organs, posterior attachment organ (haptor) and the number and arrangement of testes, though homology of some characters is questionable (Perkins et al. 2009). There are approximately 180 described species of capsalids but this is definitely an underestimate of the true diversity given the currently documented range of hosts and the available host range still to be explored (Whittington 2004). Most described capsalid species parasitise marine bony fishes but there are representatives that parasitise sharks and rays, unlike other monogenean families that may parasitise only Chondrichthyes (e.g. Hexabothriidae; Monocotylidae) or Teleostei (e.g. Dactylogyridae; Diplectanidae). The evolution and radiation of the Capsalidae across the diverse array of hosts is yet to be explored. It is not known whether their radiation has been characterised by coevolution or other events like host switching. A robust phylogeny of parasites and hosts is required to investigate the radiation of capsalids across the lineage of fish and to reveal coevolution events.

Fishes are the most diverse group of vertebrates. There are in excess of 28,000 described species of bony fish along with more than 900 species of Chondrichthyes (see Nelson 2006). A recent taxonomic classification of fishes divides them into 62 orders, 515 families and approximately 4500 genera (Nelson 2006). Due to the extensive range of diversity of the fishes, deciphering relationships between and within the different orders, families and genera is challenging. Relationships amongst the basal fish groups are considered to be reasonably well understood but much controversy remains over higher-level teleost relationships (Miya et al. 2003). Classification of orders, families and genera is based largely on morphological comparisons (Nelson 2006). However, fishes have a long evolutionary

history and there is extensive variation in morphology, behaviour, ecology and physiology that has created difficulties in comparative approaches (Helfman et al. 1997). Problems in finding sufficient phylogenetically informative characters and accurate homology assessments have been on-going in unravelling fish relationships (Miya et al. 2003; Li et al. 2007). This has led to a recent focus on molecular phylogenetic approaches (Miya and Nishida 2000; Smith and Wheeler 2004, Li et al. 2007), however, phylogenetic hypotheses for the entire lineage of fishes using molecular data have been attempted rarely (Miya et al. 2003). Analyses using complete mitochondrial (mt) genomes have shown some orders to be polyphyletic, particularly the Gasterosteiformes, Perciformes and Scorpaeniformes (see Miya et al. 2003; Smith and Wheeler 2004). The extent of this polyphyly remains to be determined and will require extensive taxon sampling to resolve relationships fully. How other orders of fishes fit into this puzzle also needs resolving but does not require as extensive taxon sampling at the generic level. With extensive molecular data now available on GenBank for many fish taxa, multilocus phylogenetic analyses of fishes using nuclear and mt genes is possible.

To begin to unravel the complex evolutionary history of the Capsalidae, we have taken the hypothesis of capsalid phylogeny of Chapter IV and compared it with a new phylogeny of fish hosts constructed here. Using sequence data that we generated *de novo* in combination with the extensive data available on GenBank, we propose a phylogenetic hypothesis for not only the marine fish hosts of the Capsalidae but, for the first time, all of the major lineages of fishes. We base this phylogenetic hypothesis on combined nucleotide sequence data for three mt genes and seven nuclear genes from representatives of 61 orders and 139 families (based on the classification of Nelson 2006). Knowledge of the time scale for fish diversification is essential to assess temporal congruence with the timing of parasite diversification episodes. The extensive fossil record for fishes, recently developed relaxed molecular clock methods and the availability of nuclear gene sequence data suitable for molecular clock analyses (e.g. *Recombination Activating Gene 1* [*RAG1*]) has given valuable insights into the timing of the radiation of fishes (e.g. Santini et al. 2009). However to date, rigorous molecular clock analyses have focused only on the ray-finned fishes, i.e. Osteichthyes (see Underwood 2006, Hurley et al. 2007, Azuma et al. 2008, Alfaro et al. 2009, Santini et al. 2009). We extend molecular clock dating of the fishes by including the Chondrichthyes,

providing the first molecular clock analysis of all major fish lineages. Thus, we compared both the topologies of parasite and host phylogenies and the relative timing of cladogenesis in each from molecular clock analyses to examine coevolutionary relationships and potential drivers of capsalid diversification.

Materials and Methods

HOST SAMPLING

Host tissue was sourced from collections by the authors and is now deposited in the Australian Biological Tissue Collection (South Australian Museum). Details of taxa are listed in Appendix XI.

DNA PREPARATION, PCR AMPLIFICATION AND SEQUENCING

DNA was extracted according to the Gentra Kit (Gentra Systems) protocol for animal tissues preserved in ethanol. Extracted DNA was stored at 4°C. PCR amplification of partial *NADH dehydrogenase subunit 4* (*ND4*) and *RAG1* was carried out with primers listed in Table 1. Primers used for PCR were also used for sequencing. PCR amplifications were performed in 25 μL reactions using the following cycle conditions: denaturation at 94°C for 45 s, annealing at a minimum 50°C and maximum 60°C (dependent on primers being used) for 45 s and extension at 72°C for 1 min; this was repeated for 34 cycles and increased to 38–40 cycles when PCR product yield was low. Each 25 μL PCR contained a final concentration of: 0.5 U AmpliTaq Gold® (5 units/μl), 0.2 μM of each primer, 200 μM of each dNTPs, 2–4 μM $MgCl_2$, 1X AmpliTaq Gold® buffer. Annealing temperature and $MgCl_2$ concentration were varied to produce optimal amplification.

Table 1. Primers used for PCR amplification and sequencing. Sources: [1]Arevalo et al. (1994), [2]Inoue et al. (2001), [3]Iglésias et al. (2005), [4]Holcroft (2004), [5]This study.

Primer name	Gene	Primer sequence 5′ to 3′	Source
ND4	*ND4*	CACCTATGACTACCAAAAGCTCATGTAGAAGC	1
L11424-ND4	*ND4*	TGACTTCCWAAAGCCCATGTAGA	2
H12293-Leu	*tRNA-Leu*	TTGCACCAAGAGTTTTTGGTTCCTAAGACC	2
H11618-ND4	*ND4*	TGGCTGACKGAKGAGTAGGC	2
Chon-Rag1-R029	*RAG1*	AGTGTACAGCCARTGATGYTTCA	3
Of2fu	*RAG1*	CTGAGCTGCAGCCAGTATCATAAAATGT	4
Or2fu	*RAG1*	CCGAGTCTTTGTGCGCGTTCATAAAGTT	4
G1206F	*RAG1*	CACRGGGTATGATGARAAGCTGGT	5

PCR products were cleaned using Agencourt® AMPure® PCR purification kit and were cycle sequenced using the BigDye Terminator v3.1 cycle-sequencing kit (Applied Biosystems). The cycling protocol consisted of 25 cycles of denaturation at 96°C for 30 s, annealing at 50°C for 15 s and extension at 60°C for 4 min. All samples were sequenced on an Applied Biosystems 3730 DNA sequencer.

PHYLOGENETIC ANALYSES

Sequence chromatograms were edited using SeqEd version 1.0.3 and aligned manually in SeAl version 2.0a11 (Rambaut 1996) using inferred amino acid sequences. All new fish sequences have been deposited on GenBank (Accession Nos XXXXXXXX – XXXXXXXX; see Appendix XI). Additional sequence data for mt *Cytochrome oxidase 1* (*COX1*), *Cytochrome B* (*CYTB*), *ND4* and nuclear *Histone 3* (*H3*), *Interphotoreceptor retinoid-binding Protein* (*IRBP*), *Mixed lineage leukemia* (*MLL*), *RAG1*, *Recombination Activating Gene 2* (*RAG2*), *Rhodopsin* (*Rhod*) and *RNF213* (an anonymous nuclear protein coding gene) genes were obtained from GenBank June 1, 2009 (Appendix XI). As our aim was to include all major fish lineages where feasible, we constrained the mt gene selection to just three genes, despite the plethora of complete mt genome sequences for fishes, as some of the

major groups, e.g. Chondrichthyes, have scant sampling for complete mt genomes. Our selection of genes maximised taxon representation and minimised the number of "cells" with missing data. For some terminals that represented higher level taxa, e.g. orders or families, the concatenated sequence comprised individual gene sequences from different families (for orders) or different genera (for families). In these cases we were careful to assess the putative con-ordinal or con-familial status of the constituent taxa from an inspection of the fish phylogenetic literature. For diverse taxa, e.g. Perciformes and Scorpaeniformes, where published phylogenetic analyses indicate polyphyly (e.g. Chen et al. 2003; Dettai and Lecointre 2005; Li et al. 2009; Miya et al. 2003; Smith and Craig 2007; Smith and Wheeler 2004, 2006), we included representatives of these unrelated lineages to improve our ability to establish relationships.

Maximum likelihood (ML) analyses were run in RAxML (Stamatakis 2006; Stamatakis et al. 2008) using the default rapid hill climbing algorithm. This analysis method allowed the data to be partitioned and optimal models of nucleotide substitution applied to each partition. The model of nucleotide substitution for each partition was assessed using the Akaike Information Criteria (AIC – Akaike 1985) in ModelTest version 3.7 (Posada and Crandall 1998). The General Time Reversible (GTR) model with a gamma distribution for rates across sites was selected. To determine an optimal partitioning strategy, preliminary analyses were run using each possible partitioning strategy then the AIC for each partitioning strategy was calculated. Bootstrap proportions were estimated under the same conditions for 100 pseudoreplicates. Some nodes at the base of the phylogeny of fishes were constrained in analyses (explained in Appendix XII, Fig. 1) to improve performance. Constraining analyses to choose between alternate hypotheses has been shown to be effective when combining fast-evolving markers (e.g. mtDNA) with conservative traits (e.g. nuclear genes) to resolve ancient, closely spaced divergences (Lee 2009). When analysed in isolation, fast-evolving markers can converge on false trees whereas combining markers with constraints can provide important phylogenetic signal to resolve divergences (Lee 2009).

MOLECULAR CLOCK DATING

To qualitatively explore the temporal similarity between the parasite and host phylogenies, we performed molecular clock dating of the fish phylogeny using the

RAG1 sequence data in BEAST (Drummond and Rambaut 2007). A fully resolved constraint tree was used in analyses based on results from the present study. Taxa for which there were no *RAG1* data were removed from the tree. ML analyses of the *RAG1* data were performed to confirm these data retrieved the nodes we were interested in estimating dates for. Twenty fossil calibrations were taken from Underwood (2006) and Azuma et al. (2008) and were treated as exponential priors including soft maxima with 95% of their density lying between the uniform maximum and minimum. BEAST analyses were run for 10^7 generations. Burnin and 95 % highest posterior densities (HPD) for results were calculated using the program Tracer version 1.4 (Rambaut and Drummond 2007). Trees were viewed and annotated in TreeAnnotator version 1.4.8 and FigTree version 1.2.2 (Rambaut and Drummond 2008; Rambaut 2009). Date estimates from the fish analyses were compared with parasite dating information from Chapter IV.

COEVOLUTION TEST

For coevolution analyses, the host phylogeny was pruned to those host taxa for which we had sequence data. Similarly the parasite phylogeny was reduced by the removal of parasite species for which the host taxon was not represented and all non-capsalid taxa were removed. A distance based method was implemented as a test of cophylogeny. ParaFit was used to assess the null hypothesis of random association between parasites and hosts (Legendre et al. 2002). It accommodates uncertainty in tree topologies, multiple parasites per host lineage and multiple hosts per parasite lineage. ParaFit uses patristic distances of parasite and host phylogenies transformed into principle coordinates. Distance matrices were converted to principal coordinate matrices using the programs DistPCoA (Legendre and Legendre 1998). Tests were performed with 9999 permutations globally across both matrices and for each individual parasite-host association.

Results

FISH PHYLOGENY

Alignment of the 185 fish taxa for the seven nuclear genes and three mt genes was concatenated for a total alignment of 8903 characters of sequence including: 1533 characters *COX1*, 1122 characters *CYTB*, 331 characters *H3*, 819 characters *IRBP*,

551 characters *MLL*, 1002 characters *ND4*, 915 characters *RAG1*, 888 characters *RAG2*, 780 characters *Rhod* and 962 characters *RNF213*. This included the following number of parsimony informative sites for each gene: 756 for *COX1*, 725 for *CYTB*, 104 for *H3*, 492 for *IRBP*, 313 for *MLL*, 760 for *ND4*, 571 for *RAG1*, 618 for *RAG2*, 562 for *Rhod* and 258 for *RNF213*. Details of sequence data used in analyses are in Appendix XI. Length variable regions of sequence data were removed for analyses, particularly where large insertions were present.

Preliminary ML analyses and AIC showed that 30 partitions with each gene separated into codon positions was optimal for the data. We present the best scoring ML tree in Fig. 1A and 1B with bootstrap proportions (BS). Basal relationships in the fishes are believed to be well understood with relationships in the more derived groups being unknown or less certain (Miya et al. 2003). Some basal nodes were constrained in analyses (details in Appendix XII, Fig. 1) and are not discussed further. Data presented here is compared with a current classification of fish by Nelson (2006) (see Appendix XII, Fig. 1).

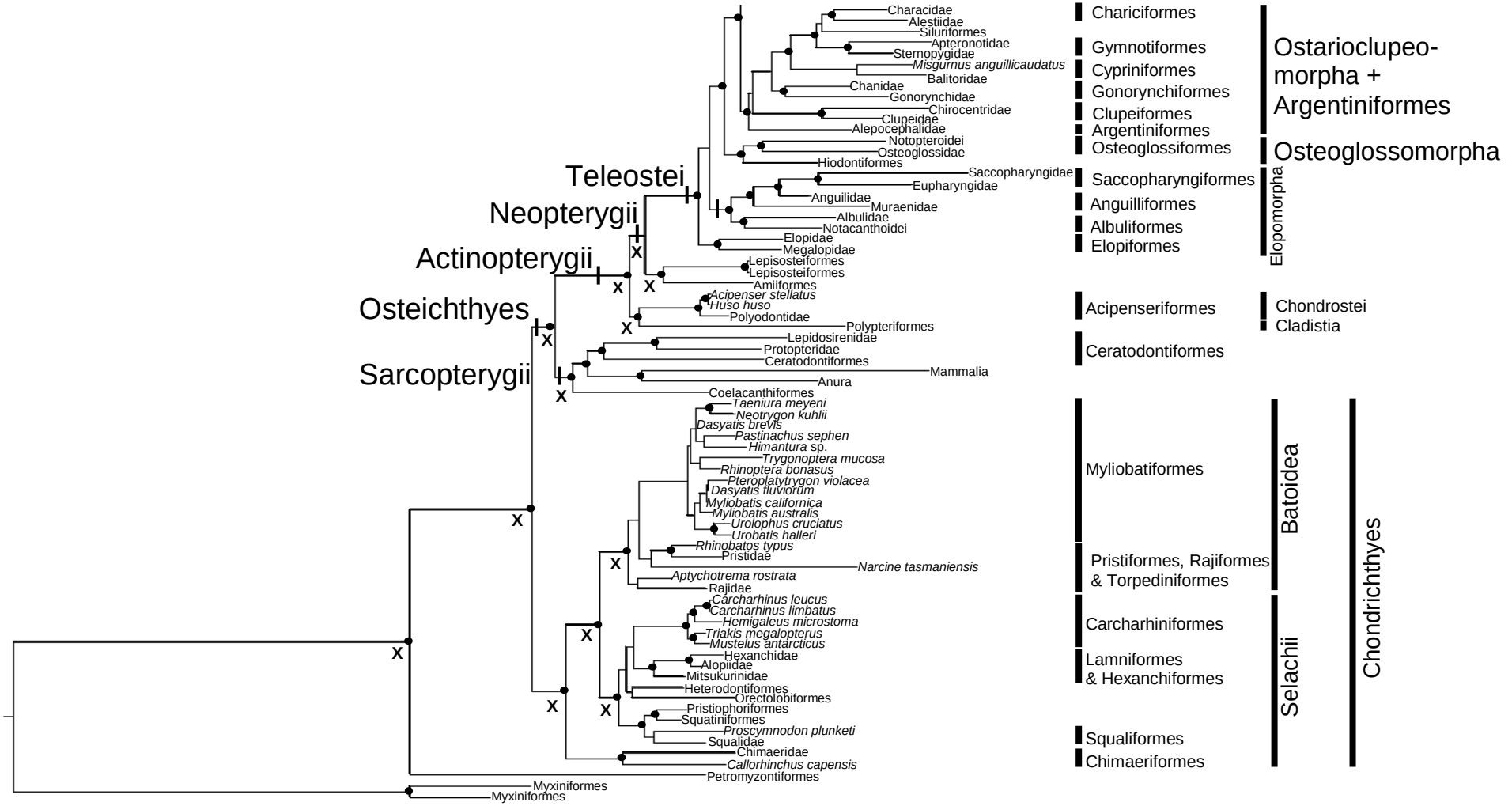


Figure 1A. Basal part of the Maximum Likelihood tree derived from analyses of seven nuclear and three mitochondrial genes from 61 orders of fishes. Non-parametric bootstrap support proportions above 70% are indicated by circles at the nodes. X indicates nodes that were constrained in analyses. Log likelihood score -349038.48.

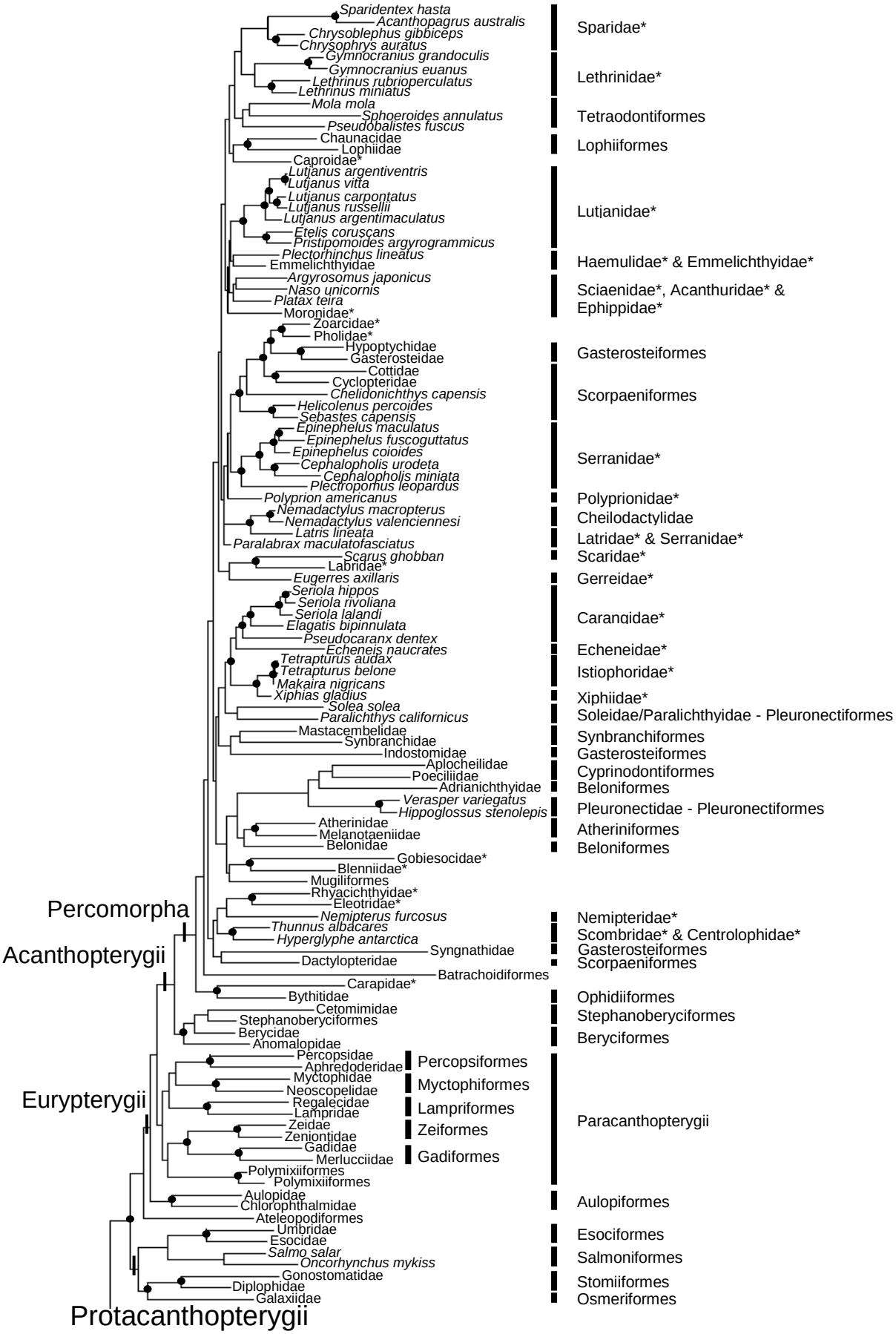


Figure 1B. The Euteleostei part of the Maximum Likelihood tree derived from analyses of seven nuclear and three mitochondrial genes from 61 orders of fishes. Non-parametric bootstrap support proportions above 70% are indicated by circles at the nodes. * indicate families that belong to the Perciformes. Log likelihood score -349038.48.

Chondrichthyes

Myliobatiformes was monophyletic (Fig. 1A, BS 58%) with Pristiformes, Rajiformes and Torpediniformes as sister (Fig. 1A, BS 58%). Rajiformes was not monophyletic (Fig. 1A, BS 47%) with *Rhinobatos typus* (Rhinobatidae) sister to Pristidae and the two other Rajidae taxa (*Aptychotrema rostrata* and Rajidae) sister to the other rays. Within the sharks, the Carcharhiniformes was monophyletic (Fig. 1A, BS 100%) and sister to a Lamniformes + Hexanchiformes clade. The Lamniformes (Alopiidae and Mitsukurinidae) was not monophyletic (Fig. 1A). Orectolobiformes grouped with the Heterodontiformes while the Pristiophoriformes and Squatiniformes grouped together and were sister to the Squaliformes.

Teleostei

The Elopiformes, represented by Elopidae and Megalopidae (Fig 1A, BS 98%), fall out at the base of the Teleostei. The next basal group consists of Albuliformes, Anguilliformes and Saccopharyngiformes and was strongly supported (Fig 1A, Elopomorpha BS 89%). Together these groups make up the Elopomorpha but this was not monophyletic. Anguilliformes was paraphyletic with the Saccopharyngiformes (Fig. 1A, Elopomorpha). Osteoglossiformes was monophyletic and sister to the Hiodontiformes (Fig. 1A, Osteoglossomorpha BS 94%). The Argentiniformes, Characiformes, Clupeiformes, Cypriniformes, Gonorynchiformes, Gymnotiformes and Siluriformes formed a strongly supported group (Fig. 1A, Ostarioclupeomorpha + Argentiniformes BS 90%). Esociformes, Osmeriformes, Salmoniformes, and Stomiiformes (each monophyletic) were the basal group of the Euteleostei (excluding Argentiniformes) (Fig. 1B, BS 50%). Aulopiformes was monophyletic. The next grouping could be considered equivalent to the Paracanthopterygii of Nelson (2006) (see Fig. 1 in Appendix XII). It comprised the Gadiformes, Lampriformes, Myctophiformes, Percopsiformes, Polymixiiformes and Zeiformes and is sister to the Acanthopterygii. Sister to the Percomorpha were the Beryciformes (represented by Berycidae and Anomalopidae and were paraphyletic) and Stephanoberyciformes.

Basal nodes in the Percomorpha are not strongly supported; indeed branches across the base of this entire radiation are characteristically short. There are 13 orders represented in the Percomorpha in these analyses (Atheriniformes, Batrachoidiformes, Beloniformes, Cyprinodontiformes, Gasterosteiformes,

Lophiiformes, Mugiliformes, Ophidiiformes, Perciformes, Pleuronectiformes, Scorpaeniformes, Synbranchiformes and Tetraodontiformes). Perciformes and Scorpaeniformes were extensively polyphyletic and the Beloniformes, Gasterosteiformes and Pleuronectiformes were also polyphyletic. Many families with multiple species represented were monophyletic (Carangidae, Istiophoridae, Lethrinidae, Lutjanidae, Pleuronectidae, Scorpaenidae and Sparidae).

MOLECULAR CLOCK DATING

Chondrichthyes

Results from Bayesian molecular dating analyses using the *RAG1* data are presented in Fig. 2A. Molecular dating analyses estimate the split between the Batoidea and Selachii occurred early in the Silurian, 435 mya (95% HPD: 421–448). The extant batoid lineages first appeared in the Jurassic, approximately 173 mya (95% HPD: 159–187). Much diversification of the Batoidea occurred in the last 120 million years. The first extant selachians also appeared in the Jurassic, 191 mya (95% HPD: 179–205).

Osteichthyes – Actinopterygii + Sarcopterygii

Molecular dating analyses estimate the split between the sarcropterygians and actinopterygians at 428 mya (95% HPD: 417–441) (Fig. 2A). The most recent common ancestor (MRCA) of the Sarcropterygii is estimated at 389 mya (95% HPD: 356–425). The MRCA of the Actinopterygii is estimated at 385 mya (95% HPD: 368–390) while the MRCA of the Neopterygii is estimated at 286 mya (95% HPD: 282–294). It is estimated that the Teleostei diverged from the Amiiformes 283 mya (95% HPD: 282–287) (Fig. 2A) with the MRCA of the Euteleostei estimated at 180 mya (95% HPD: 181–204) (Fig. 2B). Many of the extant lineages of the Euteleostei are present by 66 mya (Fig. 2B).

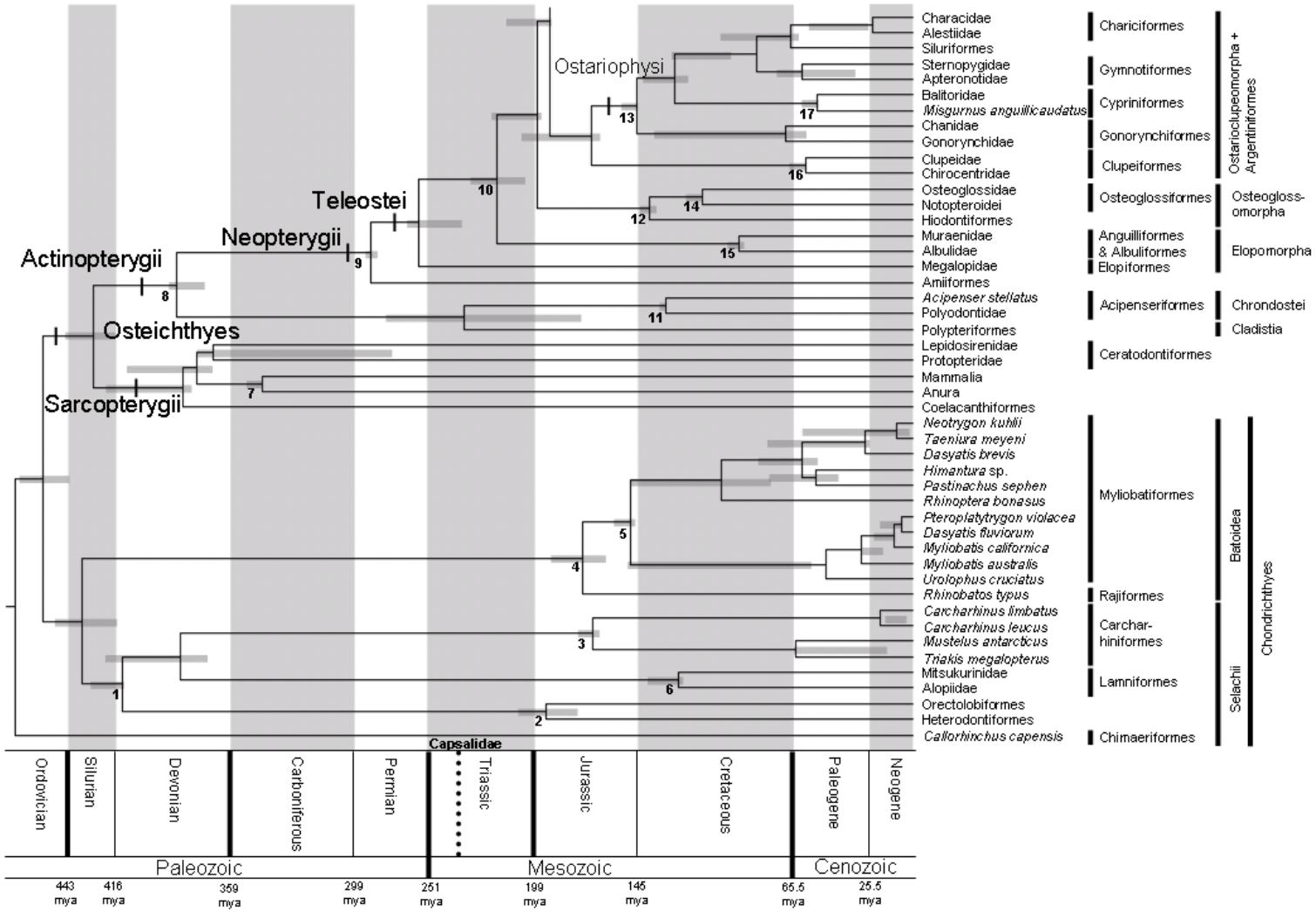


Figure 2A. BEAST maximum credibility ultrametric tree of basal fishes. Node bars indicate highest posterior densities. Timescale is in millions of years ago (mya). X indicates nodes that were fossil calibrated. Bold lines on timescale indicate major extinction events. Dotted line on timescale indicates estimated origin of the Capsalidae in the Triassic, 235 mya.

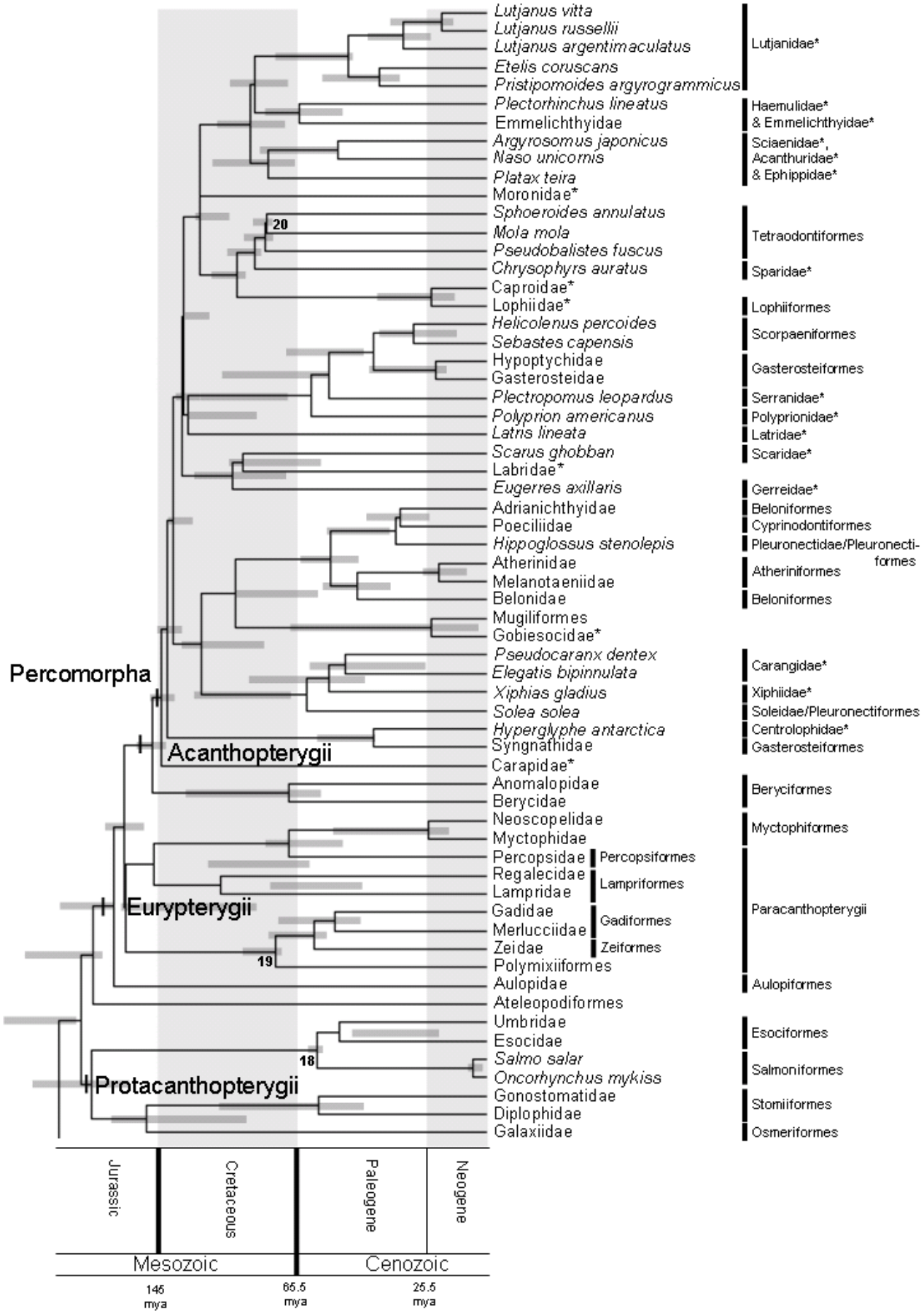


Figure 2B. BEAST maximum credibility ultrametric tree of the Euteleostei. Node bars indicate highest posterior densities. Timescale is in millions of years ago (mya). X indicates nodes that were fossil calibrated. Bold lines on timescale indicate major extinction events. * indicate families that belong to the Perciformes.

COEVOLUTION ANALYSES

The global test of parasite and host association using ParaFit showed no global relationship between capsalid and host phylogenies (Table 2, $P = 0.926$). The test computed by ParaFit for individual parasite-host links indicated a statistically significant structure for only five capsalid-host associations (Table 2). This suggests a mixed structure with potentially some coevolution. The trend across the Capsalidae sampled is parasite radiation has been independent of host radiation as indicated by results from the global test (Table 2).

Table 2. Results from ParaFit analysis. Probabilities are computed after 9999 random permutations. The null hypothesis (H_o) of the global test (bottom of table) that parasite and host evolution has been independent is not rejected. *Significant association ($P \leq 0.05$).

Parasite	Host	*P*
Allobenedenia epinepheli	*Epinephelus coioides*	0.692
Allomegalocotyla johnstoni	*Latris lineata*	0.965
Allometabenedeniella sp.	*Platax teira*	0.337
Benedenia acanthopagri	*Sparidentex hasta*	0.254
Benedenia anticavaginata	*Acanthopagrus australis*	0.230
Benedenia rohdei	*Lutjanus carponotatus*	0.705
Benedenia sciaenae	*Argyrosomus japonicus*	0.756
Benedenia sekii	*Chrysophrys auratus*	0.961
Benedenia seriolae	*Seriola hippos*	0.673
Benedeniella posterocolpa	*Rhinoptera bonasus*	0.125
Calicobenedenia polyprioni	*Polyprion americanus*	0.270
Capsala laevis	*Tetrapterus audax*	0.612
Capsala martinierei	*Mola mola*	0.959
Capsala pricei	*Tetrapterus audax*	0.623
Capsala sp. 1	*Tetrapterus belone*	0.615
Capsala sp. 2	*Makaira nigricans*	0.460
Capsalidae sp. 1	*Plectorhinchus chaetodonoides*	0.984
Capsalidae sp. 2	*Nemadactylus valenciennesi*	0.691
Capsalidae sp. 3	*Scarus ghobban*	0.457
Capsalidae sp. 4	*Paralabrax maculatofasciatus*	0.531
Capsalidae sp. 5	*Triakis megalopterus*	0.075
Capsalidae sp. 6	*Lutjanus russellii*	0.741
Capsalidae sp. 7	*Lutjanus vitta*	0.797
Capsalidae sp. 8	*Gymnocranius grandoculis*	0.267
Capsalidae sp. 9	*Naso unicornis*	0.979
Capsalidae sp. 10	*Lethrinus miniatus*	0.876
Capsalidae sp. 11	*Lethrinus rubrioperculatus*	0.920
Capsalidae sp. 12	*Cephalopholis urodeta*	0.888
Capsalidae sp. 13	*Nemipterus furcosus*	0.634
Capsalidae sp. 14	*Cephalopholis urodeta*	0.936
Capsalidae sp. 15	*Cephalopholis miniata*	0.825
Capsalidae sp. 16	*Lethrinus miniatus*	0.778
Capsalidae sp. 17	*Pristipomoides argyrogrammicus*	0.938
Capsalidae sp. 18	*Epinephelus maculatus*	0.619
Capsalidae sp. 19	*Latris lineata*	0.438
Capsaloides magnaspinosus	*Tetrapterus audax*	0.610
Capsaloides sp. 1	*Tetraptures belone*	0.615
Capsaloides sp. 2	*Tetrapterus belone*	0.625
Dioncopseudobenedenia kala	*Naso unicornis*	0.885
Dioncus remorae	*Echeneis naucrates*	0.181
Encotyllabe caballeroi	*Lethrinus miniatus*	0.943
Encotyllabe caranxi	*Pseudocaranx dentex*	0.986
Encotyllabe chironemi	*Nemadactylus macropterus*	0.838
Encotyllabe sp. 1	*Lethrinus rubrioperculatus*	0.925
Encotyllabe sp. 2	*Gymnocranius grandoculis*	0.722
Encotyllabe sp. 3	*Gymnocranius euanus*	0.703
Entobdella soleae	*Solea solea*	0.873
Entobdella squamula	*Paralichthys californicus*	0.905
Entobdella stenolepis	*Hippoglossus stenolepis*	0.196
Interniloculus chilensis	*Sebastes capensis*	0.780
Lagenivaginopseudobenedenia sp.	*Etelis coruscans*	0.941
Listrocephalos kearni	*Dasyatis brevis*	0.120
Macrophyllida sp.	*Hemigaleus microstoma*	0.063
Mediavagina sp.	*Nemadactylus valenciennesi*	0.419
Megalobenedenia helicoleni	*Helicolenus percoides*	0.813
Metabenedeniella sp.	*Lutjanus argentimaculatus*	0.778
Nasicola klawei	*Thunnus albacares*	0.997

Table 2. Continued

Parasite	Host	*P*
Neobenedenia 'girellae'	*Verasper variegatus*	0.202
Neobenedenia 'melleni'	*Sphoeroides annulatus*	0.261
Neobenedenia sp.	*Seriola rivoliana*	0.744
Neoentobdella australis	*Neotrygon kuhlii*	0.040*
Neoentobdella diadema	*Pteroplatytrygon violacea*	0.030*
Neoentobdella natans	*Pastinachus sephen*	0.012*
Neoentobdella taiwanensis	*Taeniura meyeni*	0.022*
Nitzschia sturionis	*Huso huso*	0.156
Nitzschia sturionis 2	*Acipenser stellatus*	0.156
Trilobiodiscus lutiani	*Lutjanus argentimaculatus*	0.196
Trimusculotrema sp.	*Trygonoptera mucosa*	0.012*
Tristoma coccineum	*Xiphias gladius*	0.859
Tristoma integrum	*Xiphias gladius*	0.860
Tristoma sp.	*Xiphias gladius*	0.856
Trochopodinae sp. 1	*Pseudobalistes fuscus*	0.129
Trochopodinae sp. 4	*Epinephelus fuscoguttatus*	0.234
Trochopodinae sp. 5	*Chelidonichthys capensis*	0.282
Trochopus plectropomi	*Plectropomus leopardus*	0.987
Global test		0.926

Discussion

PHYLOGENY AND MOLECULAR CLOCK DATING OF FISHES

Our analyses represent the first in depth examination of the interrelationships of the orders of fishes using multiple nuclear and mt genes. Of the 62 extant orders of fishes recognised by Nelson (2006), 61 are represented here with only sharks of the Echinorhiniformes not able to be included. Results are discussed in relation to the hypotheses presented in Miya et al. (2003; 2005) and Nelson (2006). The ray-finned fishes (Actinopterygii) are sister to the Sarcopterygii and each is believed to be monophyletic. The Sarcopterygii includes two subclasses (Coelacanthimorpha and Dipnotetrapodomorpha) along with the unranked Tetrapodomorpha. Actinopterygii comprises the majority of fish orders (44 of 62 orders) and approximately 28,000 species (Nelson 2006) and the MRCA of the Actinopterygii is approximately 385 mya old 95% HPD: 368–390). The Cladistia and Chondrostei are the most basal subclasses of the Actinopterygii (Fig. 1A). The Neopterygii is the more derived subclass of actinopterygian fishes and our molecular clock dating estimates the MRCA of the Neopterygii is approximately 286 mya (95% HPD: 282–294). It is generally agreed that the actinopterygians are monophyletic and their fossil record extends back to the late Permian (Nelson 2006). The Lepisosteiformes and Amiiformes are the basal group of the Neopterygii. The division Teleostei comprises 42 orders (Fig. 1A). Their rich fossil record suggests they arose in the middle or late

Triassic (Nelson 2006) and this is largely supported by our estimated MRCA of the Teleostei at 258 mya (95% HPD: 235–266) (Fig. 2A). The Osteoglossomorpha has been considered the most primitive group but others have challenged this hypothesising that the Elopomorpha is more primitive based on the caudal skeleton of *Elops* and other studies have even considered the Elopomorpha sister to all other extant teleosts (Nelson 2006). Elopomorpha are largely characterised by having leptocephalus (ribbonlike) larvae. The classification of this group has been controversial with some believing the leptocephalus larvae to be the primitive condition rather than the derived condition and so of limited systematic significance (Bohlke 1989). Monophyly has been debated because many of the morphological characters used to support monophyly are considered weak. Hypotheses based on ribosomal RNA sequence data have also reached different conclusions (Filleul and Lavoue 2001; Wang et al. 2003). Our analyses do not support monophyly for the Elopomorpha due to the exclusion of the Elopidae and Megalopidae, which form a separate basal group (Fig. 1A), but our analyses can also not exclude monophyly as the node excluding Elopidae and Megalopidae was not strongly supported. Traditionally the Ostarioclupeomorpha does not include the Argentiniformes, which is classified as part of the Euteleostei (Nelson 2006). The Osteoglossomorpha includes only two orders, the Osteoglossiformes and Hiodontiformes. Most representatives of this group exhibit some kind of parental care (e.g. mouth brooding or nest protection) and most possess only the left ovary (Britz 2004). Our analyses also support monophyly for the Osteoglossomorpha (Fig. 1A).

Ostarioclupeomorpha arises from a proposed sister group relationship between the Clupeomorpha and Ostariophysi, which together comprise six orders and have been recognised as the sister group to the Eutelostei (see Nelson 2006). The Clupeomorpha (the only extant order is Clupeiformes) have multiple morphological synapomorphies (Patterson and Rosen 1977) and monophyly is strongly supported in our analyses. Fishes of the Ostariophysi (comprising Characiformes, Cypriniformes, Gonorynchiformes, Gymnotiformes and Siluriformes) possess a fright reaction elicited by an alarm substance along with multiple morphological synapomorphies (Nelson 2006) and we estimate they diverged from the Clupeiformes approximately 168 mya (95% HPD: 163–207) (Fig. 2A). Our analyses support monophyly for this group and for each order but include the Argentiniformes as sister to the Ostarioclupeomorpha (Fig. 1A). There is not convincing evidence for monophyly of

the Euteleostei and this is highlighted in our analyses with the Argentiniformes not falling within the Euteleostei (Fig. 1B). In our study, the Protoacanthopterygii are the most basal Euteleostei but this classification has been notoriously unstable (Rosen 1973; Nelson 1984; 1994). Traditionally it contains the Argentiniformes, Esociformes, Osmeriformes and Salmoniformes (see Nelson 2006). Our analyses exclude the Argentiniformes from the Protoacanthopterygii but include the Stomiiformes (see Nelson 2006), which has not been included previously. Ateleopodiformes have been considered closely related to the Stomiiformes but in our analyses the Ateleopodiformes are sister to the eurypterygians but not with strong support (Fig. 1B). Aulopiformes are monophyletic and sister to the remaining eurypterygians. There is little evidence to suggest monophyly of the Paracanthopterygii and its composition has been redefined multiple times (Patterson and Rosen 1989; Wiley et al. 2000). Miya et al. (2003) considered the Paracanthopterygii included the Gadiformes, Lampriformes, Percopsiformes, Polymixiniformes and Zeiformes. Our analyses support this grouping but also include the Myctophiformes, which was not included in the analyses of Miya et al. (2003). The Stephanoberyciformes and Beryciformes (not monophyletic) are the basal group of the Acanthopterygii and sister to the Percomorpha (Fig. 1B). Miya et al. (2003) also suggested the Ophidiiformes are primitive Percomorpha and this is supported further by our analyses. In our analyses, the Percomorpha comprises 13 orders (Atheriniformes, Batrachoidiformes, Beloniformes, Cyprinodontiformes, Gasterosteiformes, Lophiiformes, Mugiliformes, Ophidiiformes, Perciformes, Pleuronectiformes, Scorpaeniformes, Synbranchiformes and Tetraodontiformes) and diverged from the Beryciformes approximately 148 mya (95% HPD: 143–153) (Fig 2B). We found that the Beloniformes, Gasterosteiformes, Perciformes, Pleuronectiformes and Scorpaeniformes are polyphyletic. Some families for which multiple genera are represented were monophyletic indicating that the classification at the family level may be robust in these groups (e.g. Carangidae, Istiophoridae, Lethrinidae and Lutjanidae). However, as five of the large percomorph orders were polyphyletic, our understanding of relationships at this level and what constitutes lineages is poor. Our analyses represent the most comprehensive higher level phylogeny of fishes and provide some of the first molecular genetic insights into the relationships between some orders. It is clear further study into the relationships of some orders of higher Teleostei is required.

COEVOLUTION AND RADIATION OF THE CAPSALIDAE

The host phylogeny, reduced to taxa for which we had parasite and corresponding host nucleotide sequence data, was compared with the capsalid phylogeny from Chapter IV using ParaFit to assess the null hypothesis that parasite and host evolution was independent (Table 2). Coevolutionary analyses showed no significant signal of coevolution between capsalids and their marine fish hosts. This result is perhaps expected as capsalid taxa infecting sharks and rays are not the basal group in the family (Chapter IV) and so is not a result of coevolution but a secondary host-switching event. There were five capsalid-host associations that appeared to not be independent and all were capsalids of Batoidea hosts. While these results support a coevolutionary relationship between the four *Neoentobdella* species and the *Trimusculotrema* sp. and their batoid hosts we can not conclude this is the case due to some incongruence between the estimated divergence dates of the parasites and their hosts. *Neoentobdella* species diverged from *Benedeniella posterocolpa* in the mid-Jurassic (Fig 1B, Chapter IV) predating the appearance of their hosts according to our estimates. Taxon sampling within the Chondrichthyes is also quite sparse and so relationships found in these hypotheses are unlikely to be accurate (T. Bertozzi, pers. comm.). It is possible that instead of coevolution, capsalids have host switched within a lineage of Batoidea giving the appearance of coevolution. Further taxon sampling is required to confirm or refute coevolution.

Our analyses do not necessarily conclude that coevolution has not occurred between the Capsalidae and marine fishes but the extant taxa sampled provide no evidence of it. Given the long evolutionary history of this family estimated at 235 mya (95% HPD: 200–272) (Fig. 1B, Chapter V) there have probably been many extinction events in the parasite and fish lineages. One of the proposed criteria for coevolution studies is extensive sampling across parasite and host clades (Paterson and Banks 2001). This is not always possible at the family scale especially among parasites with such an extensive host range in terms of species diversity, global range, microhabitat range and also significant undiscovered diversity of parasites. Addition of parasite and host taxa to phylogenetic analyses could reveal coevolution events not shown here but it seems improbable even with taxon additions that the extant taxa of this family have a strong history of coevolution. Sorting and duplication events are unable to be detected in coevolution analyses at this level, but

some inference about relationships can be made. Addition of taxa and more complete taxon sampling within certain parasite clades may allow these events to be revealed.

Given the absence of any global coevolutionary signal between the Capsalidae and their fish hosts, a comparison of their divergence dates is important to our understanding of the origins, radiation and diversification of the Capsalidae. The origin of the Capsalidae lies early in the Triassic approximately 235 mya (Chapter IV) and the basal capsalids based on current sampling infect derived fish hosts (e.g. Cheilodactylidae, Echeneidae, Latridae and Sebastidae). This date precedes the MRCA of the Euteleostei, which we estimate began diversifying only 180 mya. The lineages of Myxiniformes, Petromyzontiformes, Chondrichthyes, Sarcopterygii, Chondrostei, Cladistia and the most basal groups of the Actinopterygii (Amiiformes, Elopiformes and Lepisosteiformes) were all present at 235 mya (Fig. 2A). It is possible that one or more of these fish lineages or a lineage that is now extinct hosted the first capsalids and that they host switched from these lineages onto available niches when fish rapidly radiated. Whether these early capsalids went extinct cannot be determined. It is possible that some may be extant and have yet to be discovered. Capsalids like many other parasites can occur at very low prevalence and intensity and so even if some hosts are sampled, the parasites may not be detected (Whittington 2004). Given the diversity of fish hosts that capsalids infect, there is no doubt more species are yet to be discovered and a primitive capsalid on these ancient, potential hosts may be among them. However, it is also possible that early capsalids went extinct along with their hosts and the extant capsalid representation may reflect one, or several, host switches prior to the extinction event. Alfaro et al. (2009) showed that there have been multiple instances of increases in diversification rates amongst fishes particularly in the Euteleostei, Ostariophysi and Percomorpha. Such increases would have led to the rapid diversification of many fish lineages providing abundant unoccupied niches for capsalids to colonise.

Some systems display high levels of coevolution where parasite and host ecology and behaviour promote tight associations and where opportunities for host switching or other events are unlikely (Hafner and Page 1995). In associations where opportunities for host switching are abundant, coevolution has been found to explain very few parasite-host relationships (Morand et al. 2008). Parasite-host associations in Clade 1 in the phylogenetic analyses of the Capsalidae in Chapter IV can be explained by many host switching events. Parasite-host combinations in Clade 1

cover a wide geographic distribution including the Atlantic, Indian and Pacific oceans. Parasites do not group in relation either to their host phylogeny or on their geographic distribution, thus showing no evidence of either coevolution or any biogeographic structuring, respectively. Parasites in Clade 1 parasitise a diverse range of teleost hosts but are restricted to the most derived groups of bony fishes. No representatives in Clade 1 are found on basal teleosts or on Chondrichthyes. Perhaps there are some common biological, physiological or immunological traits among these fishes, currently unrecognised, that may restrict capsalids in this clade to host switch only within these fish groups.

Ecology and biogeography have been shown to explain the diversification of parasites across their host range in other parasite-host associations (Weckstein 2004). Capsalid taxa in Clade 2 (Chapter IV) also show no coevolution with their hosts but do not show the same possible phylogenetic association as in Clade 1. The host range for Clade 2 is far broader including fishes from 8 orders including Chondrichthyes (Myliobatiformes, Rajiformes, Torpediniformes, and Carchariniformes), basal bony fishes (Acipenseriformes) and derived fishes (Perciformes, Pleuronectiformes and Tetraodontiformes). There is also a wide geographic distribution including Atlantic, Indian and Pacific oceans but there is no grouping according to these distributions. There are however some host ecological associations amongst parasites in Clade 2. The Capsalinae, a well supported clade (Chapter IV), exclusively infect cosmopolitan pelagic fishes such as the Istiophoridae, Molidae, Scombridae and Xiphidae. While it is entirely possible that this subfamily has a broader host range than is currently recognised these worms are typically large and conspicuous during dissections allowing collections by parasitologists worldwide (Chisholm and Whittington 2007). It is known that the tetrahedral eggs of some capsalines have a short appendage at each pole believed to keep the eggs suspended in the water column and increase their chances of encountering a pelagic host. This strategy would be even more effective if the hosts formed permanent associations e.g. schools in scombrids, or temporary associations e.g. breeding aggregations (Kearn 1986). It is possible this adaptation may restrict host switching of this group among pelagic fishes. Why they are restricted to this subset of pelagic hosts is unclear. There may be an example of sympatric speciation in *Capsala* because *Capsala laevis* and *C. pricei* are sister taxa and co-occur on *Tetrapturus audax* but occupy different microhabits i.e. gills and skin, respectively (Chapter IV).

Based on phylogenetic analyses in Chapter IV the Entobdellinae is polyphyletic and contains five taxa currently assigned to two other capsalid subfamilies nested within it. Traditionally entobdellines are considered to parasitise rays and flatfish teleosts (Pleuronectiformes). However, with the inclusion of Capsalidae sp. 5 (unassigned to subfamily), *Macrophyllida* sp. (Trochopodinae), *Benedeniella postercolpa* (Benedeniinae), *Trimusculotrema* sp. (Benedeniinae) and *Calicobenedenia polyprioni* (Benedeniinae), the host range of entobdellines is expanded to include two shark species and a perciform, the wreckfish (*Polyprion americanus*; Table 2). There is also one report of *Entobdella hippoglossi* from a round bodied teleost, *Sebastes glaucus*, but this is from only a single specimen and awaits verification (Whittington 2004). All these fish share the ecological trait of being demersal. Numerous entobdellines, along with some of the taxa now grouping within the Entobdellinae, show adaptations that may restrict them to infecting demersal hosts and therefore may host switch more frequently within this ecological confine. Many entobdellines lay eggs that bear glue droplets that allow them to stick to the bottom substrate (Kearn 1963; Kearn and Whittington 2005). This could restrict these parasites to contact with hosts that share this benthic habitat and provide the opportunity for host switching or ecological transfer between bottom-dwelling teleosts and elasmobranchs (Llewellyn 1982; Whittington 2004). Opportunities to encounter hosts are not restricted only to monogenean larvae. Most juvenile and adult monogeneans remain permanently attached to a single host individual. Among capsalids, Kearn (1988) determined that adults and juveniles of *Entobdella soleae* may transfer between hosts when two fish make contact. Kearn and Whittington (1991) reported that an entobdelline from a ray, now known to be *Neoentobdella natans*, can swim, a phenomenon identified now more broadly within the genus (*N. parvitesticulata*, see Kearn and Whittington 2005 and *N. garneri*, see Whittington and Kearn 2009). The functional significance of swimming in capsalids is not understood, but as well as permitting transfer between host individuals of the same species, it may result in occasional „mistakes" that could initiate host switching if adults persist on the new host species. Further biological studies are required to investigate this possibility.

A conspicuous feature of monogeneans is their strict host specificity (Whittington 2004), which contrasts with our observation of little or no evidence for coevolution across the Capsalidae. There has been little investigation into what

mediates and maintains this host specificity and what features may allow them to switch hosts is not understood. Host switching requires a parasite to contact a host, overcome host defences, use the host as a resource and maintain the connection long enough to reproduce (Gerado and Caldera 2007). A host can evolve defences against a parasite but a parasite can also evolve ways around these defences. This arms race can lead to adaptations that may prevent a parasite from switching to novel hosts (Yoder 1997). Such adaptations have not been explored in capsalids but the majority of capsalids are strictly host specific. It has been demonstrated in other parasitic groups that parasites can respond to host-derived cues such as chemicals or behaviour in order to find a susceptible host (Haerberlen and Haas 2008; Mordue and Birkett 2009). Due to a long evolutionary relationship between the parasite and the host, the parasite may have adaptated to cope with the host defences, to maintain contact with the host in order to reproduce and evolved features of egg biology to keep larvae close to potential hosts. Egg laying and hatching rhythms have been demonstrated in several monogenean parasites (e.g. Whittington et al. 2000; Mooney et al. 2008) and it is thought that these rhythms may correspond to host behaviour and increase the chances of larvae encountering the specific host. Perhaps the most extreme examples of this are among the eggs of some monogenean parasites of rays. Eggs may embryonate fully but do not hatch unless supplied with the correct cue. Cues vary between species but examples are exposure to specific factors such as host mucus and to non-specific stimulants such as mechanical disturbances and shadows (Glennon et al. 2006). Among entobdelline capsalids included in our analyses, eggs of *Entobdella soleae* may hatch spontaneously and rhythmically (Kearn 1973) and in response to host mucus (Kearn 1974) and eggs of *Neoentobdella diadema* hatch in response to shadows (Kearn 1982). Little is known about hatching biology for other entobdellines but eggs of *N. parvitesticulata* and *Trimusculotrema* sp. fail to hatch spontaneously (Whittington and Kearn 2005; Whittington unpublished). For other capsalids, eggs of *Encotyllabe caranxi* and *E. caballeroi* fail to hatch spontaneously and can still contain active, viable larvae after incubation for 70 and 83 days, respectively (Kearn and Whittington 1992). More broadly among capsalids, rhythmical hatching is reported for *Benedenia seriolae* (see Kearn et al. 1992) and *B. rohdei* (see Ernst and Whittington 1996). What adaptations other capsalid species may have is unknown as these cues are rarely. The direct monogenean lifecycle on

fish hosts amenable to maintenance in aquaria provides a tractable model to explore further the range of adaptations for parasites to find their hosts.

There are many biological variables interacting in the close association between a parasite and its host making it difficult to understand the evolution of the parasite in the absence of detailed information about host biology and evolution. In the highly connected marine environment, the opportunity for a parasite, such as a monogenean, to encounter novel hosts is high with eggs and larvae generally free in the water column. The processes involved in achieving infection success and ultimately leading to host switching are unknown. Further studies into the biology of these parasites and their hosts are required to unravel the details of these complex relationships.

ACKNOWLEDGMENTS

This study was funded by an Australian Research Council Discovery Grant (DP0556780) awarded to IDW and SCD and a Sea World Research and Rescue Foundation Inc. grant awarded to SCD. EMP was supported by an Australian Postgraduate Award during PhD candidature in the School of Earth and Environmental Sciences at the University of Adelaide. We thank our generous colleagues K. Buchmann, L. Chisholm, F. Garibaldi, N. Hirazawa, K. Hutson, B. Jalali, S. Jones, G. Kearn, J. Kearns, M. McGrouther, P. Montero, J. Ovenden, T. Pietsch, S. Tavakol, D. Vaughan and J. Werry for supplying fish tissues.

Supplementary data

Supplementary data associated with this chapter is shown in Appendices XI and XII.

CHAPTER VI

GENERAL DISCUSSION

GENERAL DISCUSSION

Sequencing advances

My study has used molecular genetic approaches to explore relationships within the Capsalidae and examine the radiation of the family through time and in relation to their marine fish hosts. My analyses using three unlinked nuclear loci (Perkins *et al*. 2009) provide the first comprehensive insights into relationships in the Capsalidae, far extending the study by Whittington *et al*. (2004). I have shown that the current morphological classification of the family does not define monophyletic subfamilies or genera in many cases (Chapter II; Perkins *et al*. 2009). This is largely due to insufficient morphological characters and the use of apparently homoplasious characters. I also used three unlinked nuclear and two mitochondrial (mt) genes combined in phylogenetic analyses to examine capsalid relationships with broader taxon sampling but this approach has not been powerful enough to fully resolve relationships (Chapter IV). Molecular dating analyses showed the Capsalidae has its origins in the Triassic, approximately 235 million years ago (mya) with divergence of the family steady through time (Chapter IV). Due to the age of the family, saturation of nucleotide substitutions in genes is a significant problem and more nuclear genes with slower mutation rates may be required to resolve the deeper nodes more robustly. In addition to investigating the evolution of the Capsalidae, I assessed the monophyly of Monogenea and examined the evolution of diet across the parasitic Platyhelminthes using complete mt genomes (Chapter III). The addition of unlinked nuclear genes and greater taxon sampling across the Monogenea and Neodermata, especially to include Aspidogastrea, will provide further insights into the evolution of, and relationships within, the parasitic Platyhelminthes.

There have been major advances in sequencing techniques recently making gene sequencing for such phylogenetic analyses far more efficient, cost effective and less reliant on published primers or sequences. The advances of second generation sequencing, characterised by technology like 454 sequencing, have offered means to produce vast amounts of genetic data (albeit from few individuals so far), but these will soon be replaced by third generation sequencing which will generate complete genome data in a cost-effective manner for numbers of individuals typical of comprehensive phylogenetic assessments (Rusk 2009). Such technology makes multi-gene datasets far easier to produce which should lead to more robust phylogenies and a move away from the reliance on linked ribosomal RNA genes

commonly used in monogenean phylogenies and indeed in molecular phylogenetics of most other flatworms. The first two complete nuclear genomes of parasitic Platyhelminthes (the digeneans *Schistosoma mansoni* and *S. japonicum*) were sequenced recently (Berriman *et al*. 2009 and Zhou *et al*. 2009, respectively) also providing valuable data to identify slowly evolving genes that are not saturated and can resolve deeper nodes within the Capsalidae and further examine the evolution of the parasitic Platyhelminthes. The sister group to Capsalidae remains unresolved in my analyses (Chapters II and IV). My work using mt genomes confirmed paraphyly for the Monogenea (Chapter III). It is clear that a comprehensive molecular phylogeny of the Monogenea is required using multiple unlinked nuclear genes to understand the relationships between the different families within the Monopisthocotylea and Polyopisthocotylea. In particular, broader sampling across all monopisthocotylean families will help to resolve where the Capsalidae fit and identify their sister group. With sequencing advances, such studies can now become a reality.

Advances in sequencing will not only provide ways to produce data for higher level phylogenetic analyses but also to investigate shallower relationships and species boundaries, which was beyond the scope of my study. There are known instances of cryptic species within the Capsalidae particularly *Neobenedenia „melleni'*. This one „species" has been described from more than 100 host fish species representing five orders and 30 families (Whittington 2004). It is also recorded as an important pathogen in aquaculture and aquaria worldwide (Kaneko *et al*. 1988; Deveney *et al*. 2001). This apparent lack of host specificity has been recognised as an artefact of our inability to distinguish cryptic species (Whittington 2004). There have been few studies of *Neobenedenia* species on wild hosts at natural prevalences and intensities. The ability to distinguish species that cause disease from those that do not is important for the management of these parasites in aquaculture and aquaria and the biology and genetic relationships of *Neobenedenia* species is important to study further.

Future directions

Undoubtedly, increased taxon sampling and the addition of more unlinked nuclear genes would improve the phylogenetic hypotheses for capsalids that I have proposed. However, just as biological observations in the absence of a phylogenetic

context have limited meaning, phylogenetic relationships without biological context also have limited value. While my study has provided extensive insights into the relationships within the Capsalidae, there is a significant lack of biological, ecological and morphological information about the parasites and hosts to draw confident conclusions about their evolution and diversity in a unified biological context. My comparison of capsalid and host phylogenies showed that there is no evidence of cophylogeny and comparison of estimated molecular divergence dates for capsalids and their hosts also showed no temporal congruence with divergence dates for the parasites tending to be older than their hosts (Chapter V). However, host switching within phylogenetic or ecological confines may explain the distribution of capsalids across their diverse marine fish hosts. Despite their mostly strict host specificity, it seems capsalids have still been able to host switch over evolutionary time when niches or hosts become available. Much is unknown about the biology, reproduction and infection ability of most capsalids and these areas need to be explored to understand how they affect the ability of different parasite species to find, attach to and then survive on their hosts.

Egg morphology and larval characteristics are not described often with new parasite species largely because studies of live material adds complexity to collecting material and optimal, fresh specimens are not always obtained during host examination. These stages of the parasite lifecycle may be important to obtain phylogenetically informative morphological characters and may be vital to understand interactions between capsalids and their hosts (Whittington 2004). Capsalids display a range of egg morphologies. The majority are tetrahedral but others are urn-shaped and they can also have a variety of appendages, depending on species, from short, stout prongs to long or short filaments and some are associated with cement that bind them to the substrate (Kearn 1963, 1986). These different egg morphologies may be adaptations to increase their chances of encountering a potential specific host in preference to a novel host. The basic ground plan of larval morphology is relatively conserved but there are some highly variable elements among capsalid taxa. Most are ciliated and can swim but non-ciliated larvae are known (e.g. *Encotyllabe* species, see Whittington and Kearn 1992). Presence or absence of cilia may change how a capsalid oncomiracidium finds and contacts its host. Non-ciliated larvae are incapable of swimming and so must contact the host directly. *Entobdella soleae* is arguably the most well studied monogenean and is

certainly the most intensively studied capsalid. Investigations have shown that when larvae of this species are in the close vicinity of their flatfish teleost host, *Solea solea*, they reduce swimming speed and rotate on a fixed point believed to aid in contacting the host (Kearn 1981). Host finding behaviours in other capsalids are unknown.

Egg laying and hatching

Further biological information is required on capsalids to understand what maintains their host specificity yet allows host switching. Capsalids, like other monogeneans, are thought to be extremely host specific but my analyses and comparisons of parasite and host phylogenies show that host switching has probably been extensive within this family. The connectivity of the marine environment no doubt plays a significant role in this ability to host switch with many eggs and larvae likely contacting novel hosts by chance. My analyses, however, suggest this host switching has not been random and some capsalids may have strategies to increase the chances of contacting a specific host that could limit host switching. Egg laying rhythms and egg hatching strategies have been investigated in some monogenean parasites. The capsalid, *Benedenia seriolae*, displayed no egg laying rhythm but a polyopisthocotylean species from the same host carangid species, *Seriola quinqueradiata*, showed a well defined egg laying rhythm suggesting these two unrelated parasite species have different infection strategies for the same fish host (Mooney *et al*. 2008). Glennon *et al*. (2006) demonstrated three different egg hatching and host finding strategies in three unrelated monogenean species infecting the same host ray species. Eggs of one species hatched spontaneously but with a diurnal rhythm while others required specific cues to hatch such as mechanical disturbance and some displayed a mixture of both strategies. Larvae also differed in survival times with some surviving less than four hours but others able to survive up to two days (Glennon *et al*. 2006). These different strategies allow the larvae to find the same host species yet may also allow some to encounter and make contact with some novel hosts. Those that hatch spontaneously and whose larvae survive longer could potentially contact a wider range of hosts than those that hatch in response to specific cues. Studies of *E. soleae* eggs show they hatch mainly during the first two hours of daylight believed to coincide with when their nocturnally active host, *Solea solea*, settles on the bottom to rest for the day but eggs will hatch anytime when they come into contact with sole mucus (Kearn 1973; 1974). This strategy could limit the

types of hosts that *E. soleae* eggs may contact (e.g. demersal fish). Laying and hatching rhythms and the longevity of larvae may all influence the opportunities for a parasite species to encounter a novel host and therefore affect the number of opportunities to host switch.

Host immunology

Once a parasite has contacted a host, it is likely that host defences will play a role in the survival of the parasite and therefore the success or failure of the infection. Host defences and their immunological responses to monogenean parasite infections are not well studied. Immunological response obviously differs between host individuals due to genetic variation as in aquaria, there can be hosts with heavy infections alongside those with very few parasites (Kearn 2002). Studies on *Gyrodactylus derjavini* have shown that the anti-parasite response acts through the skin via the mucus (Buchmann 1999). Mechanical or chemical activation of the fish skin caused the production of cytokines resulting in mucus secretion. The mucus secretions include factors that can react with the parasite or even affect its behaviour (Buchmann 1999). Studies of *E. soleae* have also revealed that fish are not only infected by either eggs or larvae but that adults can also transfer between host individuals when soles come into contact with each other. Infection experiments involving *E. soleae* and novel hosts have shown that these different life stages of the parasite may be affected differently by the host immune response. Invading oncomiracidia were able to attach to a novel teleost host but did not survive to maturity whereas adults could not invade the host at all. However, infection experiments involving a novel ray host showed the opposite pattern with oncomiracidia unable to attach but adults survived for several days (Kearn 1967). Clearly further study of host immunological responses is required along with studies on how often parasite larvae may contact a novel host and the success of that contact in terms of infection. Once some of these factors are understood, the broad affect of this on the evolution of parasitic groups can be investigated. Many capsalid species are also very site specific and once attached to their host, post-larvae may migrate to the definitive site. Studies on *E. soleae* indicate that they may have a sense of touch which aids in them orientating themselves and migrating to their site of choice (Kearn 1984; 1988; Kearn *et al*. 1993). Whether tangoreception has a role in recognising a novel host is unknown.

Host evolution

Given the close association between host and parasite, there are also many host traits that could drive or restrict radiation of the parasites across their hosts. Unfortunately we do not know enough about the ecology, physiology and behaviour of many of the marine fish hosts to further understand this in relation to capsalids. What is very apparent from my studies of these parasites is that we can not understand their diversity, radiation and evolutionary history in ignorance of their hosts. My phylogeny of fishes provided the first examination of relationships between many fish orders and shows that many of the recent orders may be polyphyletic (Chapter V). This phylogeny allowed a robust hypothesis on the placement of host taxa for coevolutionary analyses. There was no signal of coevolution between capsalids and their diverse fish hosts but instead there may have been extensive host switching within ecological and phylogenetic confines of the hosts. Molecular dating placed the origins of the Capsalidae before the most recent fish groups which are the predominant hosts for the extant capsalid taxa sampled (Chapter V). This suggests that the origins of the Capsalidae lie on more primitive fish groups and the majority of early capsalid radiations either went extinct, perhaps along with their hosts, or are yet to be recovered from extant primitive host taxa.

Many of the estimated capsalid divergence dates precede the estimated diversification dates of their extant hosts. This indicates that these parasite lineages may have undergone multiple unknown host switches to end up on their current hosts, invading these new fish hosts as they evolved. Regrettably the number and nature of these likely host switches can never be determined. Without my independent robust study of host phylogeny and fish evolution through time, such evolutionary interactions could not be known. Host behaviour could also be very important for the transfer of parasites between individuals and between species perhaps leading to host switching. Some species of fish school often with fish of the same species but others may school in multispecies groups. Other fish can be largely solitary but then aggregate for spawning. Some fish can be sedentary while others undergo regular seasonal or annual migration. These different behaviours will likely all impact on the survival of their parasites and potentially reduce or increase opportunities for host switching. Without an understanding of this biology and behaviour, it is difficult to further understand the evolution of capsalids in relation to their hosts.

As discussed previously, some of the capsalids have also evolved adaptations to biological and behavioural facets of their hosts. A comprehensive and informal knowledge of capsalid diversity, radiation, speciation and evolutionary history can only be achieved by detailed, holistic studies on the biology of the parasites and their hosts. Indeed this principle can be applied to all parasite and host interactions. A reason why the association between chewing lice (Phthiraptera) and pocket gophers (Rodentia: Geomyidae) is so well known (Hafner and Page 1995) is that biological and behavioural features of each partner in the system are well understood. The resulting constraints of flightless, host specific, obligate parasitic lice with a direct lifecycle on burrowing gophers that rarely interact with other mammals has inevitably led to strong coevolution. The scenario for host specific marine parasites with a direct lifecycle infecting a diversity of marine fish species in openly connected oceans globally for more than 200 million years has, perhaps inevitably, led to a far more complex and difficult to unravel evolutionary association, probably due to loss of evidence through extinction of hosts (and necessarily their parasites) or parasites (without host extinction).

Conclusions

My project has delved within the phylogenetic relationships of the Capsalidae and studied their evolution through time and in respect to their diverse marine fish hosts. Resolution of the phylogeny of this monogenean family will be an ongoing task given its huge diversity, much of which is not discovered yet. However, what knowledge is really missing now is extensive biological, behavioural and morphological information to further interpret these phylogenetic relationships in a biological context. My study revealed that the current capsalid classification requires some revision due to extensive polyphyly in some subfamilies and genera. However, there currently exists no biologically meaningful way by which to divide the family. To redefine classifications within the Capsalidae requires significantly more biological and morphological information as to base the classification solely on the current molecular data may provide an unstable and problematic classification that will need to be continually revised. The evolution and diversification of this family is no doubt a result of complicated interactions between the parasites, their hosts and the complex and continuous marine environment, a challenge to unravel. Perhaps

what my study highlights most importantly is the need for an integrative approach to study the evolution, diversity and relationships of parasitic groups.