Resolving the Interrelationships of the Spathebothriidea

Rebecca Pearson

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Science of Imperial College London and the Diploma of Imperial College

Contents

List of Figures

List of Tables

Abstract

Tapeworms of the order Spathebothriidea are unique in exhibiting serially repeated reproductive organs without the accompanying segmentation that is the hallmark of the Cestoda. The order comprises five monotypic genera disjunctly distributed across the northern hemisphere and found in a range of freshwater, euryhaline and marine teleosts and condrostean hosts suggesting they are a relictal group of a once more diverse clade. This study represents the first investigation of intra specific variation in three of the genera and the first attempt at molecular characterization of the genus *Bothrimonous*. Nuclear ribosomal gene sequences (28S and ITS2) were amplified from a number of individuals and combined with all pre-exising sequences of individuals from the order. Attempts were made to characterize mitochondrial COI data, but were not successful. In addition, a short fragment (103 bp) of 18S sequence from an archival, formalin-fixed specimen of *Bothrimonous* was characterized using spathebothriid-specific primers newly designed for this study. Data were aligned using MUSCLE, corrected by eye, and models of nucleotide substitution chosen using MrModelTest. Interrelationships of the taxa were estimated using parsimony and Bayesian inference. Results indicate that no significant intraspecific variation is present across a number of hosts and geographic localities, except for the presence of two crytic species previously shown to exist in the genus *Didymobothrium*. The first genetic data ever characterised from *Bothrimonous* shows *Cyathocephalus* to be its closest association, rejecting previous synonymisation of the genus *Bothrimonus* with *Diplocotyle* and *Didymobothrium*. These data provide the first complete analysis of the Spathebothriidea.

Introduction

The Spathebothriidea (Platyhelminthes: Cestoda) contains five currently recognised monospecific genera *Bothrimonus* Duvernoy, 1842, *Cyathocephalus* Kessler, 1868, *Didymobothrium* Nybelin, 1922, *Diplocotlye* Krabbe, 1874 and *Spathebothrium* Linton, 1922. These five occupy a disjunct distribution across the northern hemisphere inhabiting the intestines of a variety of chondrostean and teleost hosts such as *Myoxocephalus scorpinus* (Scorpion fish), *Acipenser spp.* (Sturgeon) and *Solea lascaris* (Sand sole) (Gibson 1994). The tapeworms of the order are unique in exhibiting serially repeated reproductive units (proglottids) without accompanying external segmentation, which is the hallmark of the Cestoda (Olson et al. 2001) (Marques et al. 2007). This condition was previously hypothesised to have resulted from secondary loss of external segmentation (Gibson 1994) but molecular work involving nuclear ribosomal genes has placed the Spathebothriidea at, or near, the base of the cestodes repeatedly, suggesting that their condition represents an intermediate form in the evolution of segmentation (i.e strobilation) (Mariaux 1998; Olson et al. 2001; Olson et al. 2008) This coupled with their disparate host associations, disjunct geographical distribution and the progenetic behaviour (reproduction in the intermediate, invertebrate, rather than final, vertebrate, host) of some species suggests they represent a relictal group of a once more diverse clade (Marques et al. 2007) (Gibson 1994). Indeed Olson et al. (2001) suggest the small number of extant taxa when compared to over 600 strobilate genera illustrates the significance of the evolution of strobilation within the Cestoda, allowing development and fertilization to occur away from the adult and thus promoting cross fertilization. Olson et al. (2001) also included a diagrammatic representation of the evolution of strobilitation which highlights the important step the Spathebothriidea represent in this process (Fig. 1)

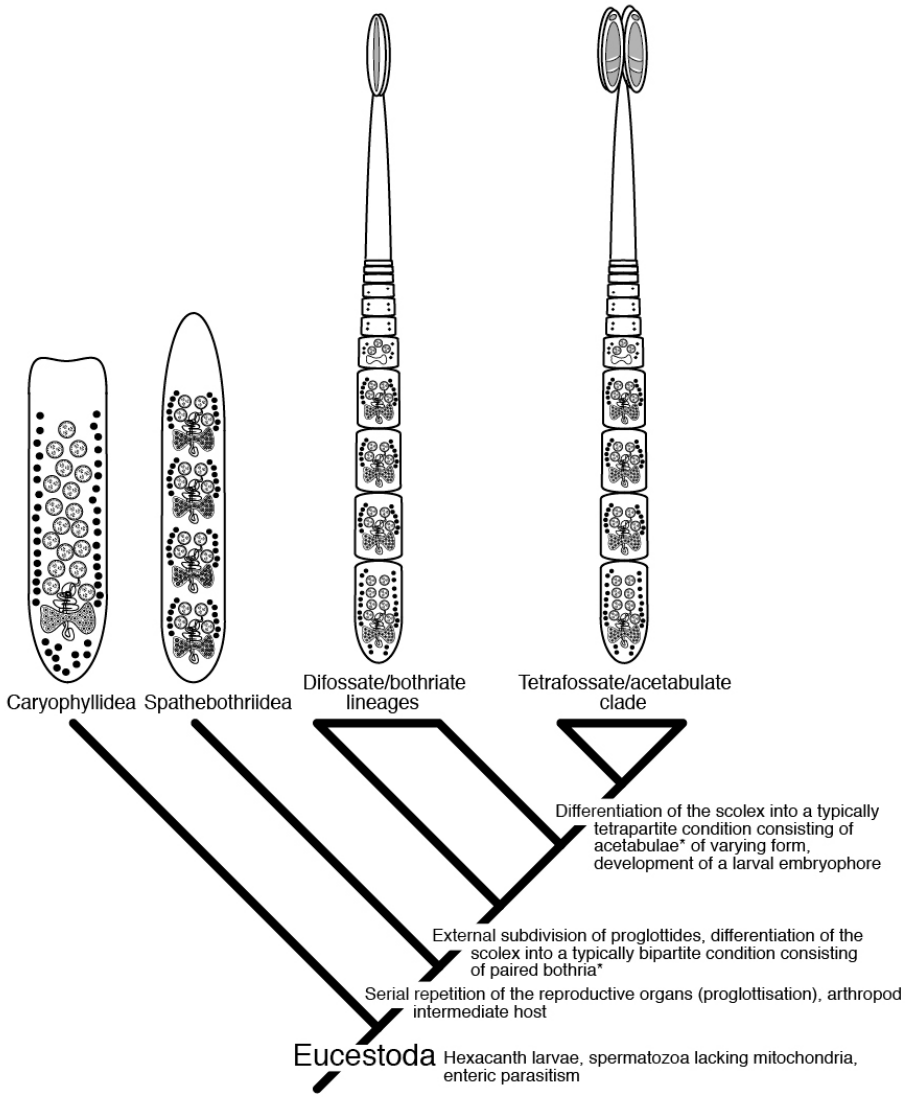

Figure 1: Diagramatic representation of the stepwise evolutionary pattern resulting in strobilate, tetrefossate/acetebulate condition characteristic of the higher eucestodes. From Olson et al. (2001)

Few studies have been carried out involving this poorly known order. Even morphological studies have been rare, the last detailed study of all the taxa is considered to have been that of Nybelin (1922) whose figures are still used widely used (Gibson 1994). Another study of note was that of Burt and Sandeman (1969) who presented a detailed review of *Diplocotyle, Didymobothrium* and *Bothrimonus* based on their morphology. They concluded that the differences used to separate the

genera were not valid, regarding *Diplocotyle* and *Didymobothrium* as junior synonyms to *Bothrimonous.* Recently a few detailed morphological studies have been carried out such as those of Poddubnaya et al. (2005, 2006) and Marques et al. (2007) but these focused on only *Diplocotlye olrikki* and *Didymobothrium rudolphii*..

The use of molecular data in investigating the interrelationships of tapeworms is scarce overall and has almost exclusively covered groups of medical or veterinary importance (Mariaux 1996) (Mariaux 1998) meaning that other than the inclusion of the Spathebothriidea (usually with only one or two representatives) in investigations of the wider interrelationships of the Cestodes (such as (Olson et al. 2001; Olson et al. 2008) little work has been done involving the order. The only investigations of note are those of Renaud and Gabrion (1998) who first identified the crypic spcies of *Didymobothrium* in Sole using allozymes and Marques et al (2007) who corroborated their findings using 28S and ITS2 sequences of multiple specimens of this genus. Marques et al. (2007) also went on to compare numerous Didymobothrium sequences with single sequences from *Cyathocephalus truncates*, *Spathebothrium simplex* (amplified and accessioned by Olson et al. 2001) and *Diplocotlye olrikki* (generated during the study) to allow an investigation of the relationships among these four genera. Other than this work no other investigation of the interrelationships of the Spathebothriidea has been carried out, no study has included representatives of all five genera, or indeed multiple representatives of any species other than *Didymobothrium rudolphii*..

The genus *Bothrimonous* has not been included in any molecular study to date. The hosts for this genus are several chondrostreans including *Acipenser stellatus* (Sterllate sturgeon) and *Acipenser nudiventris* (Fringebarbel sturgeon) (Gibson 1994) which are rarely encountered and heavily protected. *Bothrimonous* is mentioned as a parasite in sturgeon species by Mokhayer as part of a Phd dissertation on Surgeon parasites in the southern Caspian Sea in the 70's but it is difficult to find other records of this genus (the problem being further exacerbated by a number of publications being in Russian and proving inaccessible) (Mokhayer 1972). A recent review also mentions

occurrences of *Bothrimonous* but again details are scarce and the records are from the 60's and 70's (Bauer et al. 2002). A further study of parasitic worm intensity in sturgeon in the Caspian in 2005 did not record any specimens of *Bothrimonous* among the 102 samples examined and the study found decreased diversity of parasites overall (Sattari and Monkhayer 2005). This suggests population decline in the parasites which is unsurprising considering the alarming fall in abundance of Sturgeon due to overfishing, lack of management, habitat alteration and pollution (Bauer et al. 2002). Indeed Sturgeon can be considered amongst the most threatened of all fish species and all living species are listed under CITES (Convention on International Trade of the Endangered Species) (Lenhardt et al. 2006).

Inclusion of *Bothrimonus* in a molecular study would be interesting both because it would mean all species in the order were represented, and it would add additional hosts and geographical areas for comparison. It could potentially also resolve the issue of whether Burt and Sandleman's (1986) consideration of either *Diplocotlye* or *Didymobothrium* as junior synonyms of *Bothrimonus* was correct, something which was disputed by Gibson (1994). Gibson identified morphological characteristics which he believed could separate the 3 genera meaning the names should be recognised (Burt and Sandeman 1969; Marques et al. 2007; Sanmartín Durán et al. 1989) In addition to this morphological ultrastructure work on the vitelline material has identified characteristics which can be used to distinguish *Diplocotyle* and *Didymobothrium* (as well as *Cyathocephalus*) (Poddubnaya et al. 2006). Gibson (1994) also identified differences in host associations between the species; *Bothrimonus* being from holarctic chondrosteans, *Diplocotlye* from teleosts particularly salmonids and pleutonectids and *Didymobothrium* from in marine teleosts particularly soleids. Clarification of whether this evidence for recognising all three names is corroborated by molecular evidence would be interesting.

This project aims to conduct the first comprehensive molecular phylogenetic analysis of the Spathebothrridea based around newly available material representing all five monotypic genera

from a number of localities and hosts. This should enable the clarification of pre-existing taxonomic issues as well as for examination of the intra generic variation related to host association and locality.

Newly acquired samples for this study included a number, most from the genus *Bothrimonus*, which had come from museum collections and been preserved using formalin (an aqueous formaldehyde solution). Use of such samples from museum collections for molecular work is becoming increasingly common in molecular systematics as researchers seek to gain access to the genetic information and diversity tied up in collections created to reference morphology (Herniou et al. 1998). These collections may represent the only available material for a species, such as in the case of *Bothrimonus* where new occurrences of the parasite in the wild have not been recorded for considerable time.

In such material autolytic decay will have begun at the time of preservation, causing rapid degeneration of DNA (Carter 2003) and loss of the phosphodiester links which are essential for the maintenance of its macromolecular structure (Eglinton and Logan 1991). The DNA in preserved biological material is therefore subject to a number of degradative processes such as denaturation to single stranded DNA (then more susceptible to chemical degradation), breakage of the sugar phosphate backbone leading to fragmentation, chemical modification of nucleotides and crosslinking of the DNA with other molecules (Carter 2003) processes which are influenced by the way the specimens are preserved, their age and the environmental conditions in which they are stored (Carter 2003).

Formaldehyde was once frequently used in the fixing or preservation of museum specimens and is known to have a detrimental effect on the genetic material of specimens. Formaldehyde and its effects on DNA have been studied extensively (see references in (Carter 2003), (Zimmermann et al. 2008) and whereas it was previously thought to make PCR amplification of DNA difficult because

it led to extensive fragmentation, it is now largely accepted that DNA is instead difficult to extract due to the formation of crosslinks between the DNA and protein. Fig. 2 is a graphical representation of the effects of formalin fixation on DNA, illustrating the formation of crosslinks particularly in material which has been exposed to formalin for extended periods (from Qiagen.com).

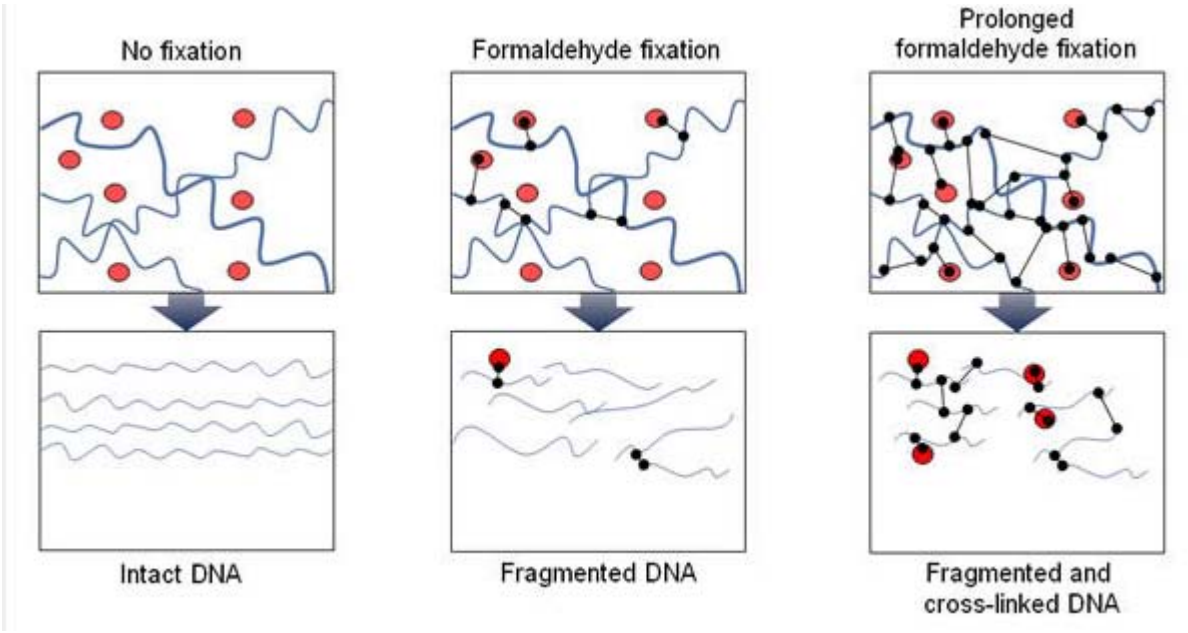

Figure 2: Effect of formalin fixation on DNA; fixation causes fragmentation of DNA and formation of crosslinks to other biomolecules, from Qiagen webpage.

The isolation of crosslinked nucleotides from formaldehyde treated DNA seemed to have been achieved back in the 1980's (Chaw et al. 1980). More recent work identified the early extraction of DNA as the difficult step in formaldehyde treated material and suggested that DNA was extensively crosslinked to the surrounding tissue explaining why extraction was so difficult. Base lesions expected in crosslinked DNA were not present, which the authors suggested indicated crosslinked DNA could not be extracted, although they did state that the results they saw (i.e. difficulty in amplifying DNA fragments from formaldehyde exposed tissues leading to low concentrations of DNA and small fragment sizes) could also have been due to heavy degradation

(Zimmermann et al. 2008). What is clear is that extracting DNA from formaldehyde treated tissues is challenging but there are no hard and fast rules to extrapolate which specimens will or won't work (Herniou et al. 1998). However, differing methods and practices appear to be required for any chance of success. Studies such as that of Herniou et al. (1998) which achieved some success in DNA extraction from formaldehyde treated helminth specimens, states a number of recommendations and considerations necessary for such work such as contamination, verification of sequences and extraction techniques. A more recent review expands on such advice listing extensively 'pitfalls and precautions' for such work, but also highlights the importance of using museum specimens (Zimmermann et al. 2008).

Methods and Materials

Taxon sampling

The new material available for this study included twenty six specimens of spathebothriid tapeworms across the five genera provided by R.Kuchta (Univ. South Bohemia, Czech Republic) including both fresh material and museum specimens (for which the full history was unknown). Samples came from a range of localities and hosts, details of which can be found in Table 1: Appendix 1. In addition to this, pre-existing sequences of spathebothriids were also included at the analysis stage from (Olson et al. 2001) and (Marques et al. 2007)), details of these are listed in Table 2: Appendix 1. Four additional specimens were also included as outgroups for the Spathebothriidea from its sister groups.

A subset of the samples that came from museum collections were not ameanable to DNA amplification by standard molecular techniques and required the use of different methods in order to optimise the chance of successful amplification. These samples, TS04.10, Got221, Stoc88.20, TS01.1 and Bfax which represent individuals of *Bothrimonus fallax* and one specimen of *Diplocotyle olrikki* from a previously unsampled location, shall hence forth be referred to as formalin fixed samples. Any work with these samples was carried out in a different laboaratory and using different equipment wherever possible in order to reduce the likelihood of any contamination by PCR products.

The samples arrived stored in 100% ethanol and the tissue was transferred (either the whole piece or a subsection thereof depending on how much material was provided) into two changes of T.E (Tris EDTA) buffer prior to extraction.

DNA extraction, PCR and sequencing

Genomic DNA of the majority of specimens was extracted using a DNeasy™ Blood and Tissue Kit (Qiagen) with a homogenisation step using a hand-held homogenizer with a sterile plastic pestle attached prior to incubation with protinase K.

The formalin fixed samples were thoroughly homogenised in 1ml of water using a Precellys 24 lysis and homogenization machine (Bertin technologies) with ceramic beads at 6500rpm for 23seconds prior to extraction. Following homogenisation the solution containing the tissue was spun down in order that water could be separated from the tissue. Genomic DNA extraction was carried out using a QIAamp™ DNA FFPE Tissue kit (Qiagen) designed specifically for formalin fixed and/or paraffin embedded tissue. As the samples in this case were not parafiin embedded, the protocol was adapted slightly, omitting the steps intended to break down the paraffin. Following extraction DNA concentration was quantified using a NanoDrop spectrophotometer (Thermo Fisher Scientific).

Three nuclear ribosomal regions were characterised in the study 28S (the LSU), 18S (the SSU) and ITS2 (internal transcribed spacer), these having proven to be informative in this group previously (Marques et al. 2007; Olson et al. 2001) as well as a mitochondrial gene COI (cytochrome oxidase I) which would be subject to different selective pressures and thus provide an independent source of data. For the main group of specimens complete ITS2 and COI and partial 28s genes were amplified by polymerase chain reaction (PCR) using the primers listed in Appendix 2, Table1 and cycle sequencing conditions listed in Appendix 2, Table 2 (Littlewood and Olson 2000; Littlewood et al. 2008; Moszczynska et al. 2009; Olson et al. 2001) All PCR reactions utilised PuRe Taq™ Ready-To-Go™ PCR beads (GE healthcare) in a reaction volume of 25µl.

In the formalin fixed samples, 18S and 28S were targeted as these regions had proven to be straight forward to amplify in other specimens and a number of primers were available. Nested PCR of both genes was carried out, initially targeting large fragments and then using a number of internal

primer combinations to counteract the effects of low DNA concentrations in the samples. Table 4 (appendix 2) details the combinations of initial and nested primers used and Table 5 (appendix 2) gives details of the primers themselves (Littlewood and Olson 2000) Initial PCR's were carried out using the programme listed for LSU5-1200R listed in Table 2 (appendix 2) and nested PCRs were carried out under the following cycling conditions: 95°C for 5 minutes, 25x 95 °C for 30seconds, 56 °C for 30 seconds, 72 °C for 1 minute and 72 °C for 5minutes final extension.

A series of new short fragment (100-200bps) 18S and 28S primers specific to the Spathebothriidea were also designed, for use with the formalin fixed specimens in the course of this study. Based on an alignment from Olson et al. (2008), regions that were conserved within the Spathebothriidea but variable within their closest related taxa were identified by eye as potential priming sites. From among these sites those which flanked regions that were the most variable within the spathebothriid species were further investigated to see if a non variable regions (within the Spathebothriidea) 18-25 nucleotide in lengths preferably with a G/C content of around 50-60% was present. Potential primer sequences were then tested virtually using Primer3 (Untergasser et al. 2007) through the Geneious platform (Drummond A.J. et al. 2010). Suitable reverse primers were also designed and tested within Primer3. In total 5 pairs of primers for each of 18S and 28S were designed, listed in Table 6. These primers were then tested in the laboratory using a positive control across a suitable temperature gradient (based upon the melting temperature (Tm)) on a Veriti Thermocycler (Applied Biosystems) using the Veriflex™ step in order to optimise the annealing temperature for each. The cycling conditions were as listed for LSU5-1200R in table2 but with an annealing temperature gradient of 55-60°C. PCR amplification of the formalin fixed specimens was then carried out using the same programme with the annealing temperatures displayed in Table 6 (chosen based on the result of the primer test).

Successful amplification of the correct product size was confirmed in all cases by the use of agarose gel electrophoresis, although in the case of the short fragment amplification a higher percentage gel

(3.5%) was used and running time was increased to ensure separation of bands. Samples which showed visible bands on the gel were retained for downstream applications.

PCR products were then either purified directly using a QIAquick® PCR purification kit (Qiagen), gel excised and purified following the protocol of the QIAquick® Gel extraction kit (Qiagen) or in the case of the short fragment products using Microcon YM30 centrifugal filter devices. These columns ensured that the small fragments would be retained. DNA concentrations of purified products were examined using a NanoDrop spectrophotometer (Thermo Fisher Scientific).

Sequencing reactions for both strands were carried out in a 10µl total volume; 0.5µl Big Dye™, 3.5µl buffer, 1µl primer (1pmol) and DNA template plus sterile water to 5µl. Two nanograms of DNA per 100 base pairs of fragment was required and the correct volume of purified product to add was calculated based upon NanoDrop readings for the samples. Reactions were carried out on a Veriti™ thermocycler (Applied Biosystems) using the following cycling conditions: 25 cycles of: 96 °C for 10 seconds, 50 °C for 5 seconds, 60 °C 4 minutes final hold 10°C. Additional sequencing primers were used for some of the reactions, listed in Appendix2, Table 3.

Sequencing reactions were cleaned either, via ethanol precipitation or in the case of the short fragments using Edge Bio Performa DTR Gel Filtration Cartridges, again to ensure the fragments were retained despite their small size and were run on an ABI 3730 DNA Analyser (Applied Biosystems)

Sequence alignment and phylogenetic analysis

Contiguous sequences were assembled and edited manually using Sequencher™ (GeneCodes Corp) then verified for identity via a BLASTn search (McGinnis and Madden 2004).

Verified sequences were aligned (by eye and pairwise alignment to the existing Spathebothriid alignments of Marques et al. (2007) using MacClade (Maddison and Maddison 2005). Sequences for the outgroup taxa either generated during the project (ITS2) or taken from an existing

alignment (28S) ((Olson et al. 2008) were added to the alignments prior to phylogentic analysis. Regions of the alignment where data was missing for most taxa, such as the beginning and end of sequences were excluded from the analysis in MacClade (Maddison and Maddison 2005). Also as outgroup sequences were seen to be highly dissimilar to those of the ingroup, regions of the alignment where there were large gaps in the ingroup sequences, generated by indels in the outgroup taxa, were also excluded. Concatenation of the two alignments (28S and ITS2) was carried out using a Perl script (Written by D. Chesters, NHM) to allow an analysis of a combined dataset including both loci. A combined dataset was produced including only taxa which had sequence data from both loci. Additionally a dataset was also produced including all individuals for which sequence data was available for one or both loci. Amplification of COI sequences was unsuccessful for all taxa using the primers desribed (Appendix 2, tables 1 and 2) and thus analysis conducted includes only the nuclear ribosomal genes.

The sequences obtained from the formalin fixed specimens were for short fragments and this meant more difficulty was encountered in recovering contiguous sequences as often overlap was not present in the sequences. Calling of bases was also done cautiously, so if a base could not be confidently identified an ambiguity code was utilised. These fragments were also compared to each other, where they represented the same region, and to the same region in the other four species, in a Muscle alignment (Edgar 2004) created using Geneious (Drummond A.J. et al. 2010) to identify base positions at which the new sequences differed. Any variable bases were then checked again in the original sequences to ensure that any bases that were different were absolutely confidently called. Any sequences which differed from those of the other 4 species in the case of *Bothrimonus* specimens (or showed the same or a very similar sequence to *Diplocotyle* in the case of TS00.5) when aligned with them were retained for phylogenetic analysis. Usable fragments were aligned with the other four species using a Muscle alignment (Edgar 2004) corrected by eye in Geneious

(Drummond A.J. et al. 2010). All bases outside of the fragment (including the primer region) were removed so analysis was of the equivalent 100bp fragment in all taxa.

Phylogenetic trees were constructed using both maximum parsimony and Bayesian inference. In the case of the 28S and ITS2 data the two genes were analysed separately but also in a combined dataset. Maximum parsimony analysis was performed using PAUP* v4.0b(Swofford 2001). A heuristic search of 1000 replicates with TBR (tree bisection reconnection) branch swapping and random addition sequence was carried out but with only one tree was saved per replicate (multrees=no) and maxtrees was set at 10000. All characters were unweighted and unordered. These trees were then swapped on again, this time allowing multiple trees and a strict consensus of these trees was produced. In order to determine branch support 1000 bootstrap replicates were also carried out but again only one tree per replicate was saved. All trees were rooted using the outgroups. The short fragment data was also analysed in PAUP* but due to the small number of characters and taxa an exhaustive search was used, results were viewed as an unrooted network. Patristic distance between the taxa were also calculated on PAUP*.

Bayesian analysis was carried out in MyBayes v3.2 (Ronquist and and Huelsenbeck)through the bioinformatics platform Geneious (Drummond A.J. et al. 2010). Prior to analysis the appropriate nucleotide substitution model for the data was selected using MrModelTest2 (Nylander 2004). The model block is executed in PAUP* (Swofford 2001) where the likelihood scores of 24 models are calculated for neighbour joining trees. These scores are then evaluated in the ModelTest programmes which ranks then based on the Akaike information criterion (AIC) taking into account the likelihood score of the model and its number of free parameters. The model with the lowest AIC is the appropriate model to use for the dataset. For the 28S and ITS2 datasets the GTR (general time reversible) model was selected with either gamma distributed among site rate variation (+G) and invariant sites (+I), in the case of Combined dataset including all individuals and ITS2, or just GTR+G, in the case of the combined datset for individuals with both loci and 28S. The short

fragment data required the use of a much simpler model, HKY (Hasegawa-Kishino-Yano). Analyses were run for 1,100,000 generations with sampling ever 200th generation and 10% burn in (100,000 generations**).**

Results

For the 26 new samples available for this project, 28S sequences (300F-ECD2) were obtained for 12 individuals, and ITS2 sequences (ITS2.3S-ITS2.2) for 8 samples due to the apparently degraded condition of the DNA of many of the samples. Four individuals that were sequenced for 28S (TS09.282, TS09.275, TS09.283 and TS09.280) are likely to have worked for ITS2 also, but as these individuals were the same species (*Diplocotyle olrikki*) and from the same locality as other specimens, they were not included. No sequence for either of these fragments was obtained for individuals of the genus *Bothrimonus*. These sequences were combined with 25 28S and 24 ITS2 sequences which had been obtained in previous studies (Marques et al. 2007; Olson et al. 2001), as well as 4 28S sequences (Olson et al. 2008) and 5 ITS2 sequences (sequenced during the project) from outgroup taxa. No reliable sequence was obtained for COI despite the use of a number of primer pairs both for other tapeworm species and for universal application.

Trees obtained with both Parsimony and Bayesian inference are shown in Figs. 3-6 where the trees from both analyses are displayed together for each of the datasets 28S, ITS2, combined with both loci and combined all individuals (see Table 1 for details of datasets). When comparing the trees a degree of topological similarity is observed between them with the main exception of the ITS2 tree (Fig. 4) where rather different sister relationships are recovered. The tree recovered for the combined dataset closely resembles the 28S tree (Fig. 3) suggesting that this larger fragment (1162bps c.f 457) has the greater influence over the topology of the tree, despite the ITS2 region being more variable with 52% parsimony informative characters (compared to 24%) (see Table 1).

Table 1: Parsimony tree statistics

Data set	Length (of included sites)	Constant sites	Variable but parsimony uninformative sites	Parsimony informative sites	Percentage informative sites	Tree length	Number MPC's
ITS2	457	91	128	238	52%	689	2
28S	1162	415	464	283	24%	1239	10000 (max trees)
ITS2 and 28S individuals with both loci	1651	545	605	501	30%	1887	675
ITS2 and 28S all individuals	1675	551	594	530	32%	1960	10000 (max trees)

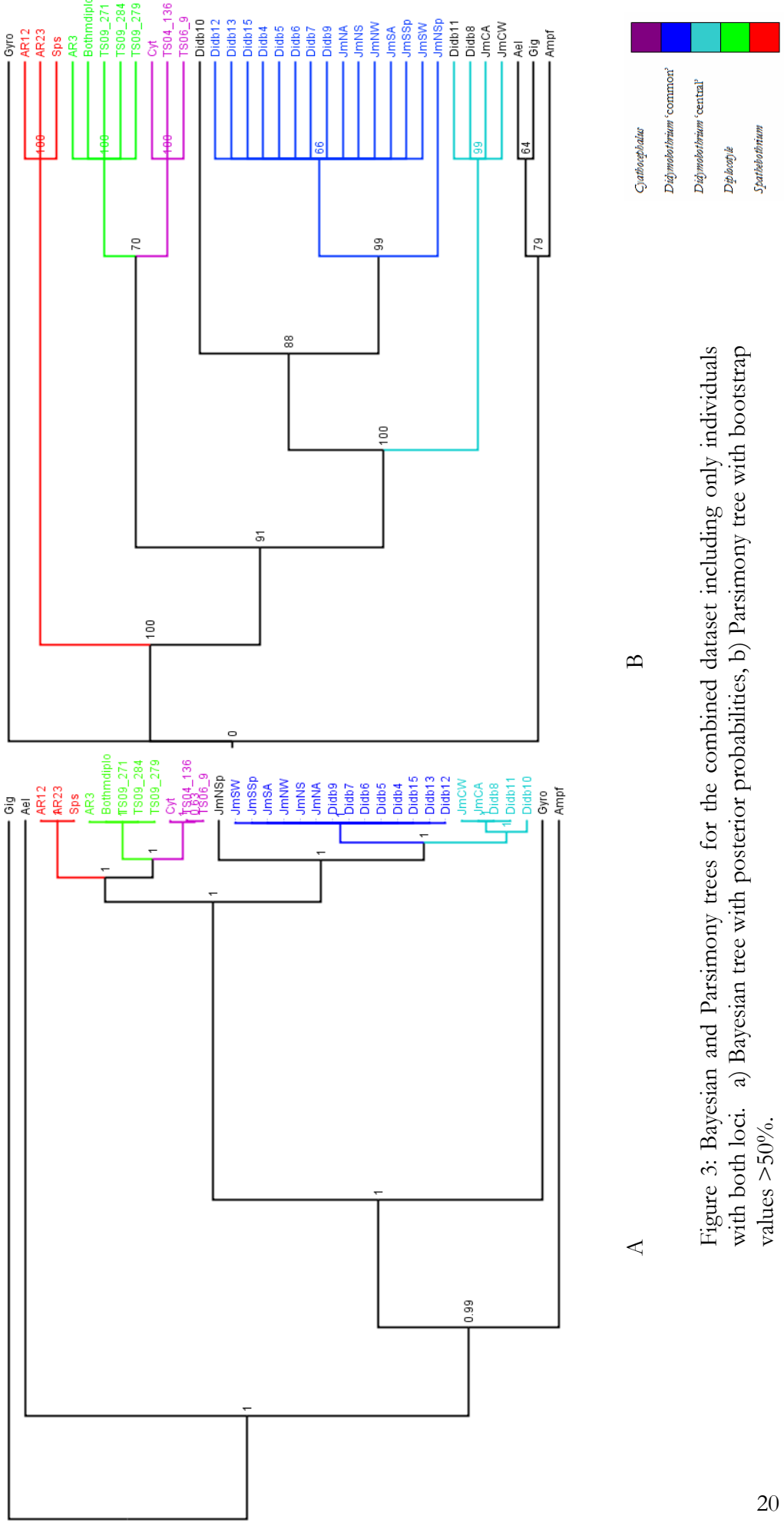

Figure 3: Bayesian and Parsimony trees for the combined dataset including only individuals with both loci. a) Bayesian tree with posterior probabilities, b) Parsimony tree with bootstrap values >50%.

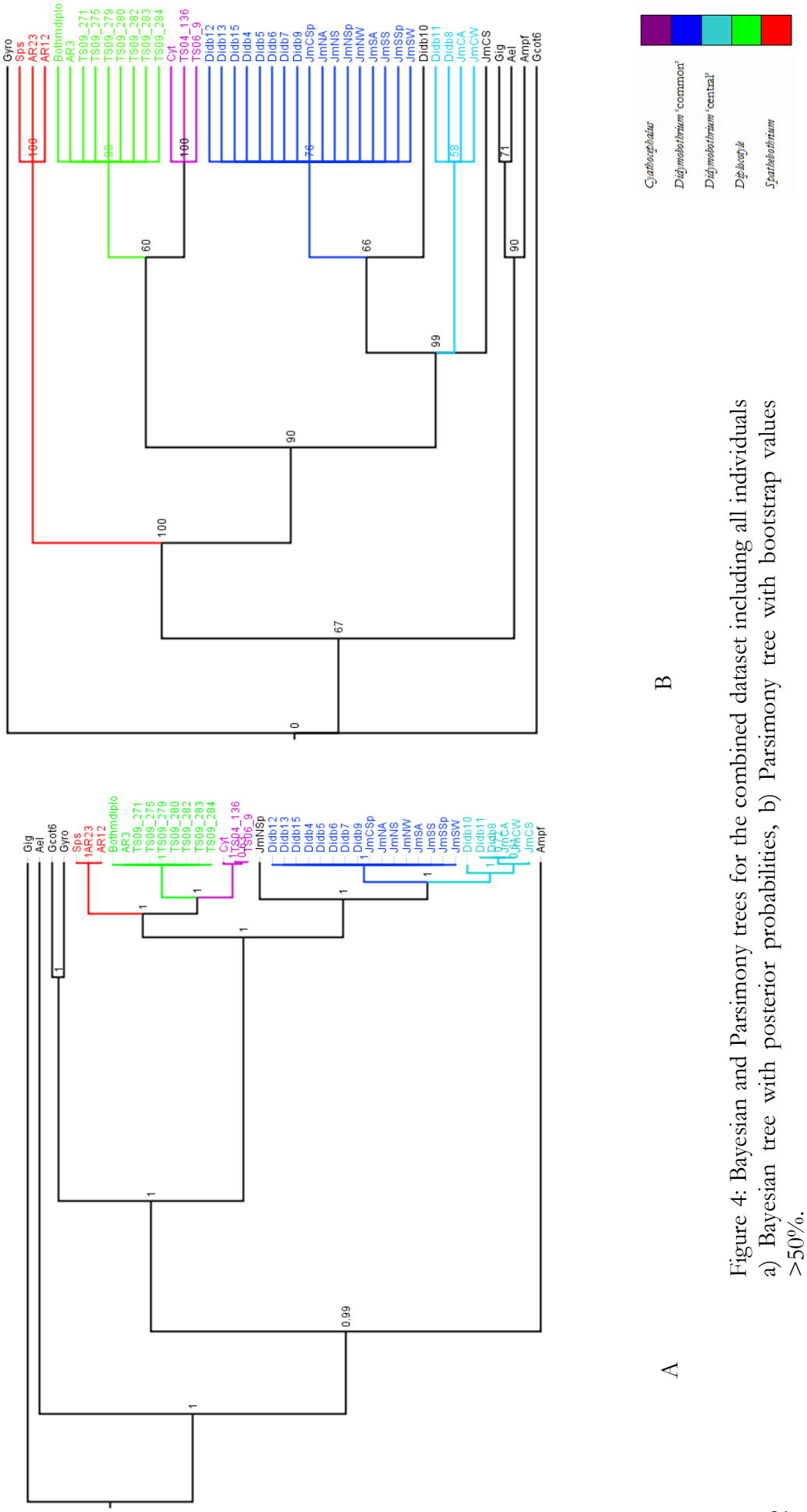

Figure 4: Bayesian and Parsimony trees for the combined dataset including all individuals a) Bayesian tree with posterior probabilities, b) Parsimony tree with bootstrap values >50%.

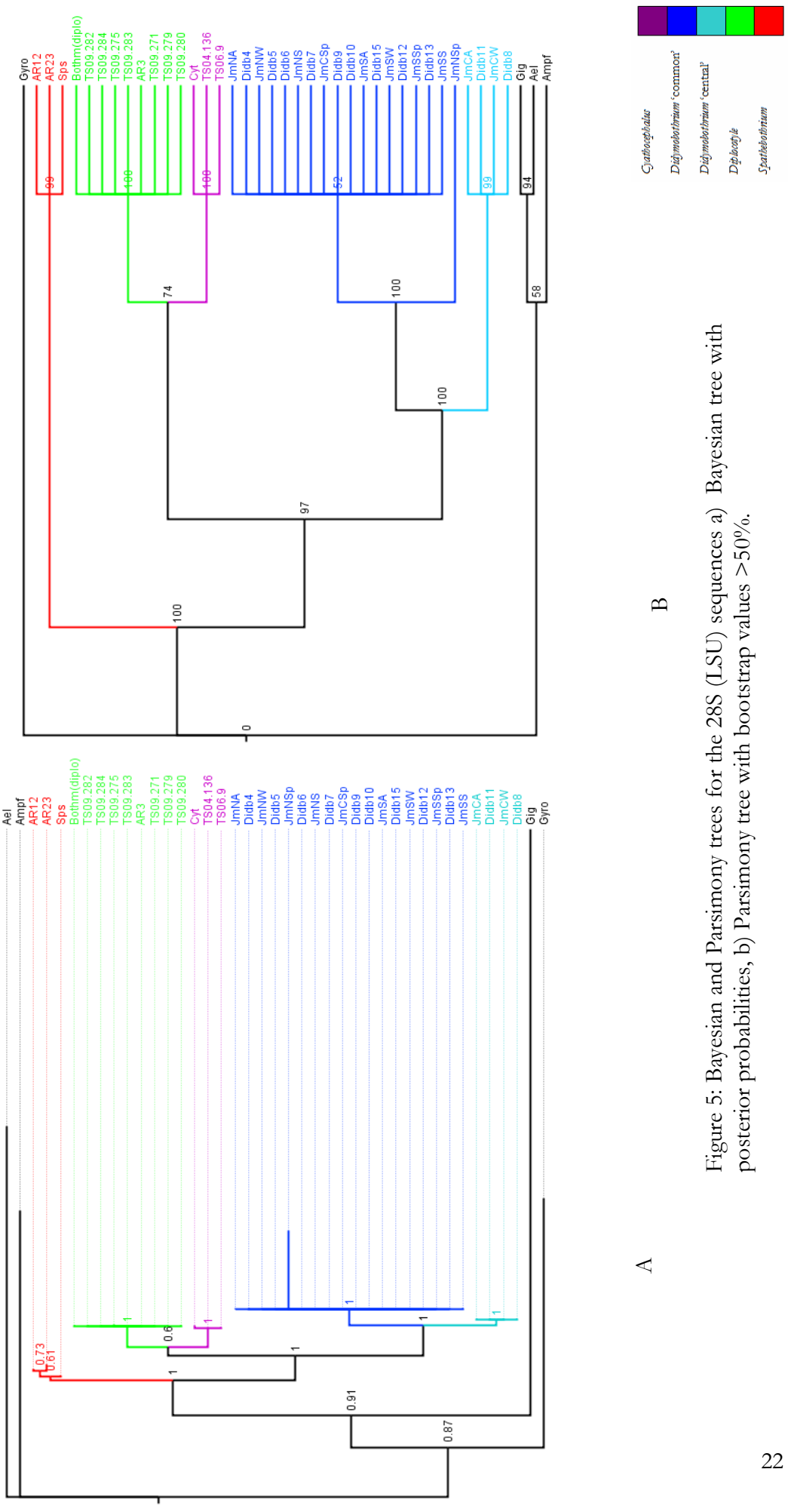

Figure 5: Bayesian and Parsimony trees for the 28S (LSU) sequences a) Bayesian tree with posterior probabilities, b) Parsimony tree with bootstrap values >50%.

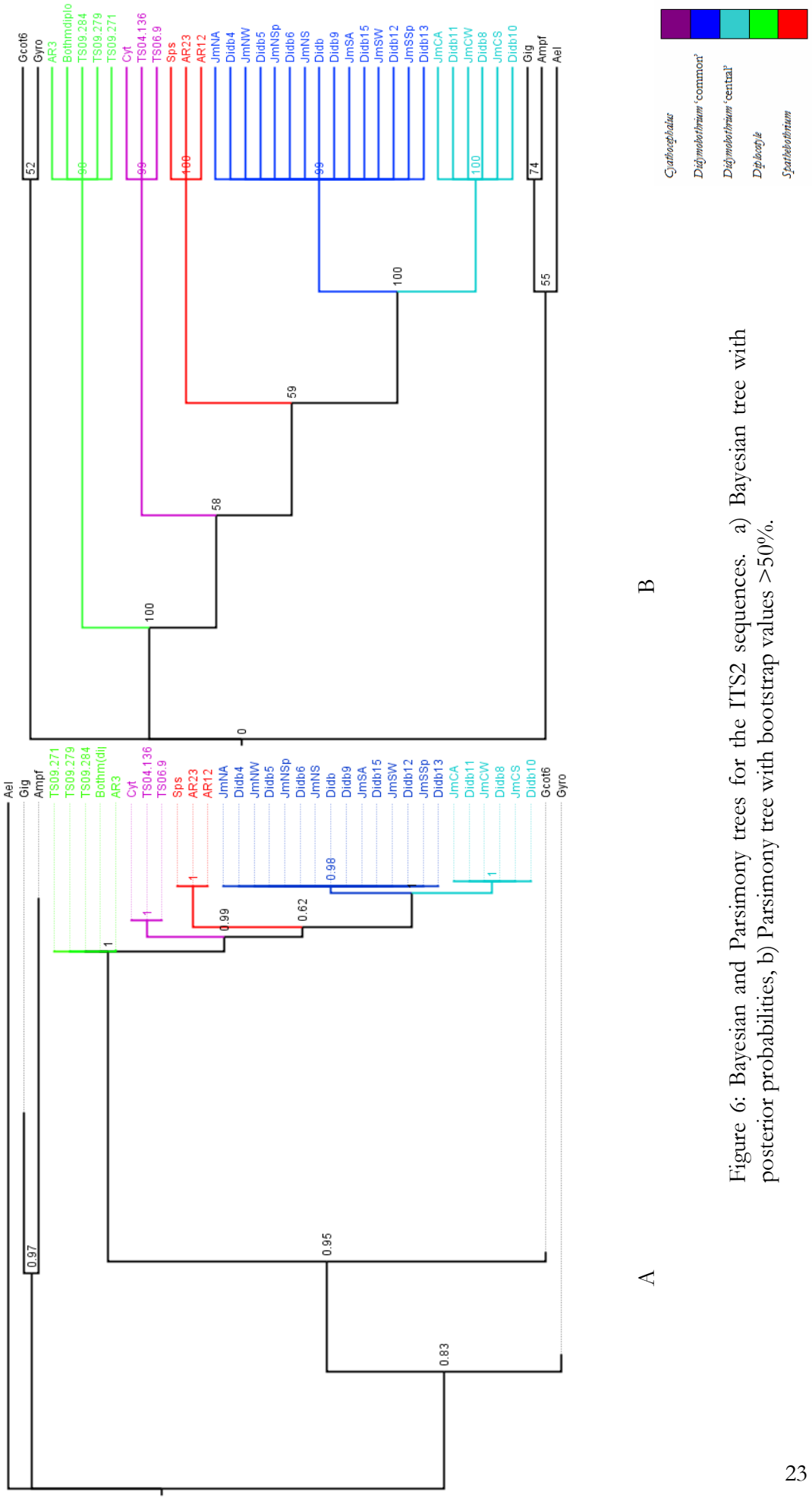

Figure 6: Bayesian and Parsimony trees for the ITS2 sequences. a) Bayesian tree with posterior probabilities, b) Parsimony tree with bootstrap values >50%.

All the trees recover each genus as a monophyletic clade with relatively high levels of support, see Table 2 for values. Within *Didymobothrium* 2 monophyletic groupings, termed 'common' and 'central' by Marques et al. 2007 are also recovered representing 2 cryptic species.

Table 2: Posterior probability and bootstrap support values for recovery of species as monophyletic groups

Dataset	Monophyletic recovery of all species in a clade with support values (Posterior probabilities/bootstrap values)				
	Diplocotyle	***Cyathocephalus***	***Didymobothrium* (central)**	***Didymobothrium* (common)**	***Spathebothrium***
ITS2	1/99	1/99	1/100	0.98/99	1/100
28S	0.98/100	1/100	1/100	1/99	0.61/99
ITS2 and 28S individuals with both loci	1/100	1/100	1/99	1/99	1/100
ITS2 and 28S all individuals	1/99	1/100	1/58	1/76	1/100

In the case of the trees recovered via Parsimony individuals within a genus clade form polytomies (with the exception of *Didymobothrium* which contains 2 species groupings) suggesting there is little or no genetic variation within each genus. The exceptions are several individuals of *Didymobothrium* (Dibd10 and JmNSp) which fall outside of the main clade on a number of occasions. However these taxa appear to be unstable as their position on the tree is variable. In the Bayesian trees some support for groupings of individuals within species groups is recovered. The 28S (Fig. 3) tree shows support for a subgrouping with the *Spathebothrium* . AR12 and AR23 are grouped with a support value of 0.73 to the exclusion of Sps, when Sps is added support for the group as a whole falls to 0.61. One other recovered grouping is within *Cyathocephalus* on the combined dataset trees

(Figs. 1 and 2) where TS04.136 and TS06.9 are recovered with support of 0.63, increasing to 1 when Cyt is included.

In terms of the sister group relationships among the Spathebothriidea the same relationships are not recovered consistently across the trees. Table 3 details the relationships recovered with support values obtained. The most common topology among the trees unites the 2 groups of *Didymobothrium* (common and central) with high support values and indeed this grouping is recovered in all trees. Also commonly recovered are *Cyathocephalus* and *Diplocotyle* as sister groups. This is recovered with medium levels of support in the majority of the trees, both Bayesian and parsimony, the exceptions being the ITS2 tree (Fig. 4) where *Diplocotyle* falls as the sister to all the other species. Another commonly recovered relationship is *Spathebothrium* as sister to all other species. This is recovered consistently in the Parsimony trees and in some of the Bayesian trees although in the case of combined data sets *Spathebothrium* is instead recovered as sister to the *Cyathocephalus, Diplocotyle* grouping only. In the ITS2 tree *Spathebothrium* is recovered as sister to the 2 *Didymobothrium* groups instead and indeed it is clear that it is this dataset which produces trees most markedly different from the others.

Table 3: Recovery of sister grouping with support values (Posterior probability and Bootstrap).

Dataset	Recovery of sister groupings with support values (Posterior probabilities/bootstrap values)							
	***Didymobothrium* common and central**	***Diplocotyle* and *Cyathocephalus***	***Didymobothrium, Cyathocephlaus and Diplocotyle***	***Diplocotlye, Cyathocephalus* and *Spathebothrium***	***Spathebothridium* as sister to other species**	***Didymobothrium* and *Spathebothriium***	***Didymobothrium, Spathebothrium* and *Cyathocephalus***	***Diplocotlye* as sister to other species**
ITS2	0/99					0.62/59	0.99/58	0/100
28S	1/100	0.6/74	1/97		1/100			
ITS2 and 28S individuals with both loci	1/100	1/70	91	1	100			
ITS2 and 28S all individuals	1/99	1/60	90	1	100			

Results for formalin fixed specimens

Nested PCR of the formalin fixed specimens was unsuccessful and only one of the sequences obtained from the use of the short fragment Spathebothriidea specific primers was reliable enough to be used. Other fragments which amplified successfully appeared to be contamination as primers targeted the most variable regions of the genes, but sequences were almost identical to the equivalent region in one of the other species. The usable fragment was from the Spath18s_1470F, Spath18s_1600R which amplified a variable region at of 18S around 100bps long. The successfully amplified sample was TS04.10 one of the *Bothrimonous* samples from Sturgeon. A BLASTn (McGinnis and and Madden 2004) search of this fragment recovered the following top 4 matches: 94% *Cyathocephalus*, 87% *Diplocotyle*, 82% *Spathebothrium* and 77% *Didymobothrium* and visual inspection revealed it to differ from *Cyathocephalus* (identified as most similar) at 8 reliably called base positions. Parsimony analysis of the fragment using an exhaustive search produced the following unrooted network (Fig. 7) with the Patristic distances between each of the taxa are displayed in Table 4 to indicate levels of sequence divergence.

The TS04.10, *Bothrimonus* fragment is seen to be closest to *Cyathocephalus* having a much shorter distance between it and this species than any of the species. With *Didymobothrium* being the most divergent of the taxa for this region of the gene. The unrooted phylogram produced from the Bayesian analysis (Fig. 8) suggests the same position for *Bothrimonus* as the one produced by parsimony. Although the position of the taxa appears slightly different *Bothrimonus* still has *Cyathocephalus* as its closest association and *Didymobthrium* is the most divergent taxon, this is also reflected in the number of substitutions per site which are displayed on the phylogram and the pairwise similarity values given in Table 5, overall similarity across all 5 sequences is 73.3%

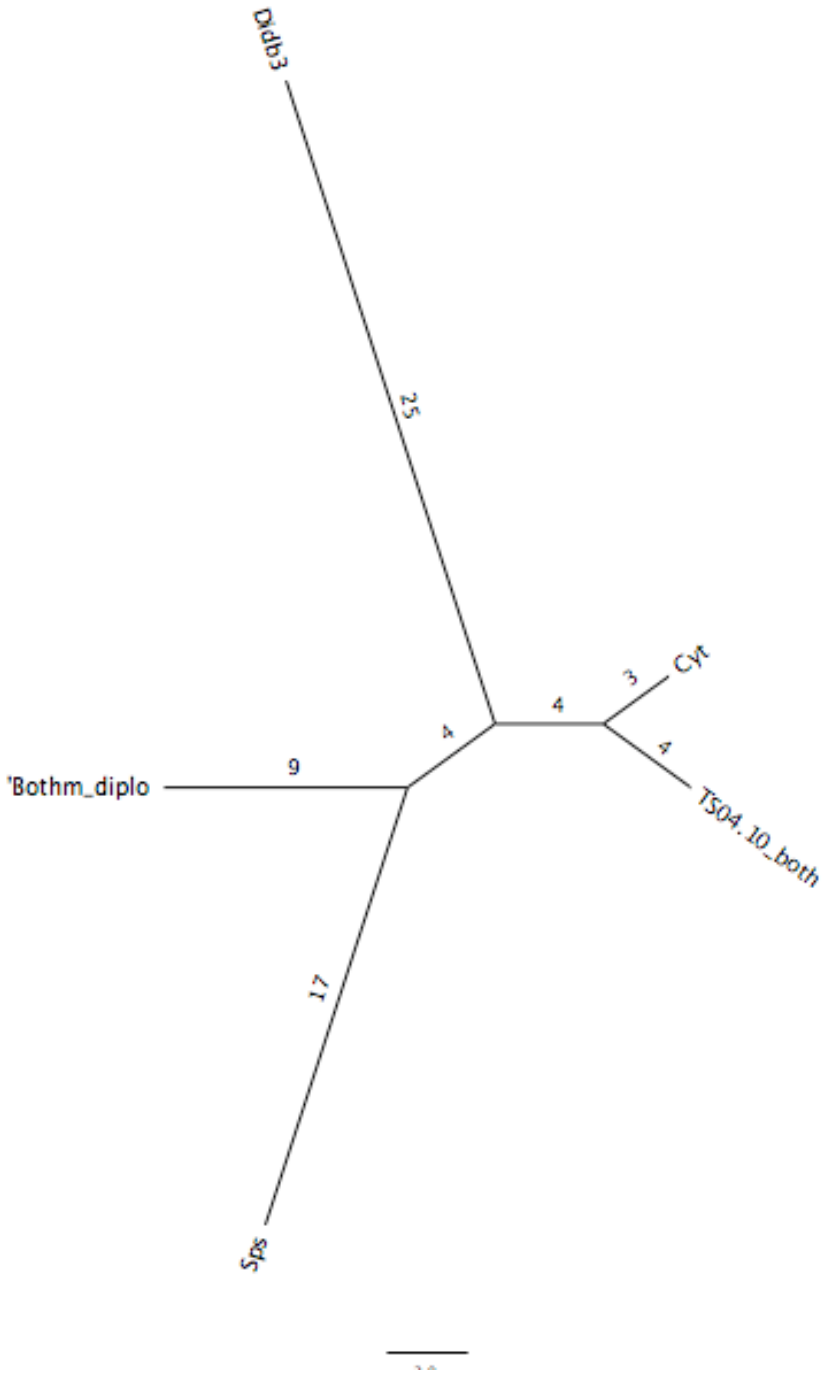

Figure 7: Parsimony, unrooted phylogram of Spathebothriidea species based on a 100bp 18S fragment. Values given are Branch lengths. Didb3=*Didymobothrium* 'common', Cyt=*Cyathocephalus*, TS04.10_both= *Bothrimonous*, Sps=*Spathebothrium* and Bothm_diplo=*Diplocotyle*

Table 4**:** Patristic distances between spathebothrrid taxa in the unrooted phylogram of a 100bp fragment of 18S.

	TS04.10_ both	**Sps**	**Dibd3**	**Cyt**	**Bothm_ diplo**
TS04.10_both	**-**				
Sps	**29**	**-**			
Didb3	**33**	**46**	**-**		
Cyt	**7**	**28**	**32**	**-**	
Bothm_diplo	**21**	**26**	**38**	**20**	**-**

Table 5: Percentage pairwise identity of 18S fragment sequences

	TS04.10_ both	**Sps**	**Dibd3**	**Cyt**	**Bothm_ diplo**
TS04.10_both	**-**				
Sps	**75.7**	**-**			
Didb3	**68.9**	**54.4**	**-**		
Cyt	**92.0**	**76.7**	**69.0**	**-**	
Bothm_diplo	**80.0**	**71.8**	**65.0**	**81.0**	**-**

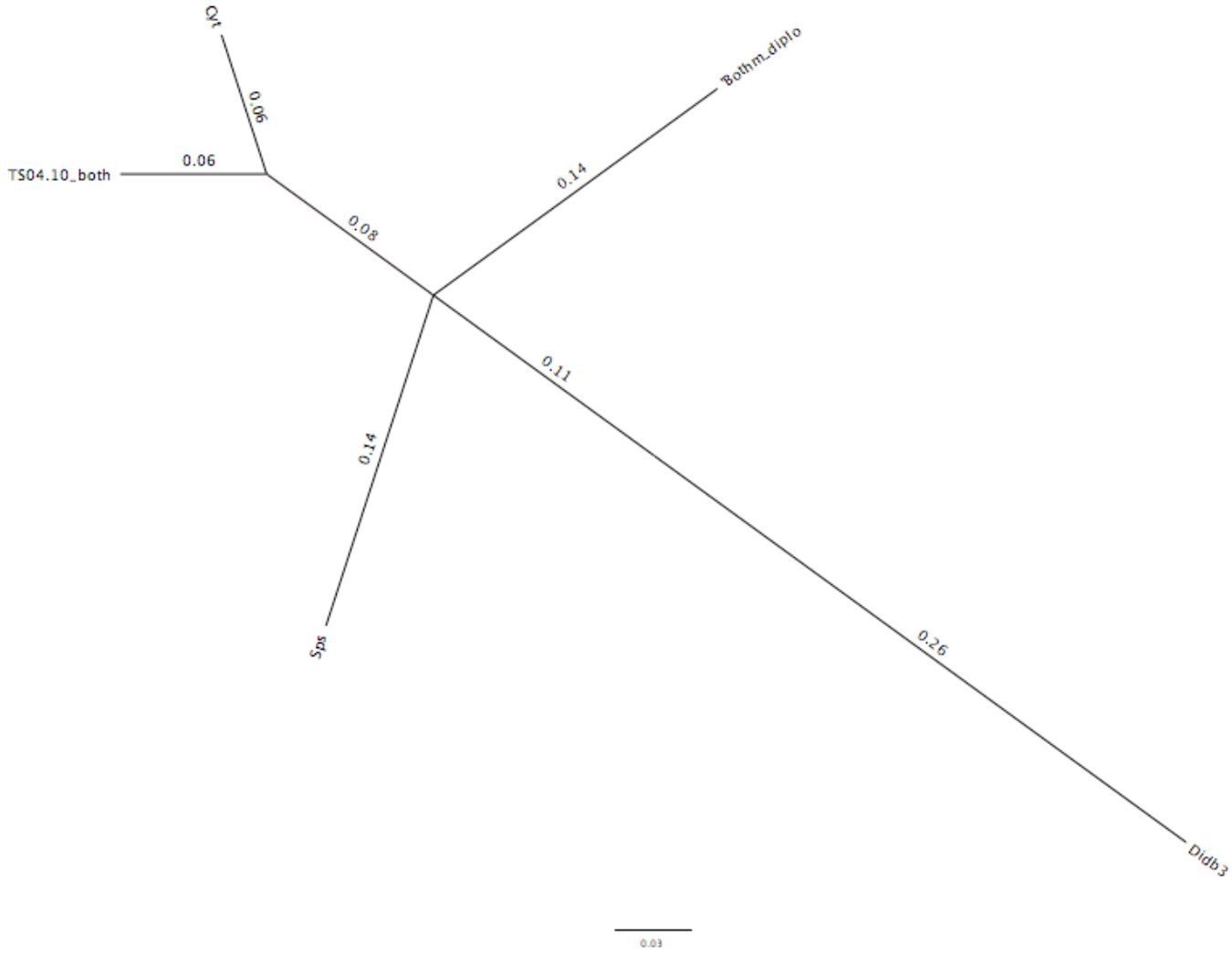

Figure 8: Bayesian, unrooted phylogram of Spathebothriidea species based on a 100bp 18S fragment. Values given are substitutions per site. Didb3=*Didymobothrium* 'common', Cyt=*Cyathocephalus*, TS04.10_both= *Bothrimonous*, Sps=*Spathebothrium* and Bothm_diplo=*Diplocotyle*

Discussion

This project represents the first opportunity to investigate intra generic variation in the 28S and ITS2 regions of *Cyathocephalus, Diplocotyle* and *Spathebothrium* and ascertain how it correlates with both host association and geographic distribution. Renaud and Gabrion (1988) and Marques et al. (2007) illustrated that the genus *Didymobothrium* contained two genetically distinct crypic species but it was uknown as to whether cryptic species are also present in the other genera.

The trees produced during this study recover each of the genera as monophyletic and identify two species present in *Didymobothrium*, as expected. This result shows that the other genera are indeed monotypic and that there is little or no variation below the species level. In *Cyathocvephalus* some variation is present in the sequences but overall they show 99.7% similarity so the level of variation is extremely low. The variation itself is correlated with geographic locality as 2 specimens from the same locality (Russia) are seen to group together on some of the Bayesian trees at the exclusion of a third from another locality (Switzerland). Individuals within the other genera all show no sequence variation. Despite this 2 individuals of *Spathebothrium* are recovered as a group, however this is a result of missing data due to differences in fragment length. Other individuals which are seen to fall outside of the monotypic genera (not produced during this study) do so due to sequence errors rather than genuine variation. Therefore it must be concluded that the other genera do not contain crypic species. Any intra specific variation present was at a level significantly lower than that found in the *Didymobothrium* species, only 0.3% compared to 2.7%.

While this project made use of all available *Spathebothriid* material, sampling of hosts and localities was still somewhat limited and often samples that amplified successful were from the same locality. Availability of material from a broader range of localities and hosts could identify further intra specific variation present in these taxa but such material simply does not exist.

Marques et al (2007) constructed the first molecular phylogeny of the Spathebothriidea using a number of 28S sequences from *Didymobothrium* individuals and one sequence from each of *Cyathocephalus, Spathebothrium* and *Diplocotlye.* They identified *Spathebothrium* as sister to the other genera with (relatively good support (68%)) as well Cyathocephalus as sister to a monphyletic *Didymobothrium* containing the two cryptic species (77% support) (see Fig.7). Unrooted networks were also produced for both 28S and ITS which identified *Didymobothrium* and *Cyathocephalus* as separated from *Diplocotyle* and *Spathebothrium* although with short internal branches whilst also highlighting the high level of divergence between the two *Didymobothrium* groups.

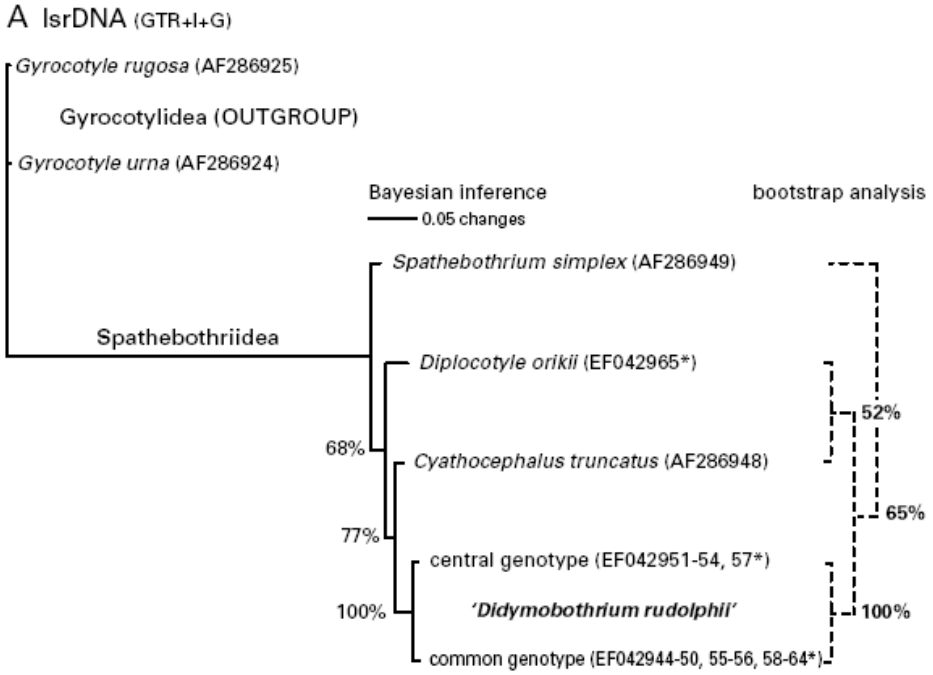

Figure 9: Rooted Phylogram with relative branch lengths based on Bayesian inference (left) and parsimony based bootstrap analysis (right) of 28S data, from Marques et al. (2007)

In this study *Spathebothrium* was recovered as sister to the other genera in the majority of the parsimony trees whilst being placed as sister to only to *Diplocotlye* and *Cyathocephalus* in many of the

Bayesian trees. ITS2 however produced a tree with a markedly different topology in which *Spathebothrium* is sister to *Didymobothrium*. This difference in the relationships suggested by the 2 genes is not something that was recovered by Marques et al. (2007). In addition *Cyathocephalus* as sister to *Didymobothrium* is not recovered in any trees in this study. A much more common sister group relationship in the current data is that of *Cyathocephalus* and *Diplocotlye*. Overall the trees cannot be said to recover the same relationships as Marques et al. (2007) and clearly the addition of more taxa has impacted on tree topology as would be expected.

While this study represents the most comprehensive molecular analysis of the order to date, clearly the picture is not yet complete. The 2 gene regions utilised in this study are both nuclear ribosomal genes which will have been subject to the same selective pressures. The inclusion of sequences from another gene region, such as a mitochondrial gene would allow corroboration of relationships as the data would be from an independent source, due to the different selective pressures acting on mitochondrial genes. This additional evidence would be useful in finding a more consistent topology for the trees. In the course of this study despite using a number of tapeworm specific primer combinations and universal primers amplification of mitochondrial COI (cytochrome oxidase I) was unsuccessful. The difficulty in amplifying this region in flatworms is well documented and most studies have used ribosomal data when studying flatworm phylogenetics (Moszczynska et al 2009). Moszczynska et al. (2009) made an attempt to design primers to amplify the barcoding region of COI in diverse platyhelminthes. When the region was successfully amplified it was seen to yield good species level resolution. However in a number of the included taxa, sequence amplification was not successful due to the high level of sequence divergence among the species.

In this study a 103 base pair fragment of *Bothrimonous* fallax 18S sequence was successfully amplified, representing the first molecular data ever collected from this species. The fragment was amplified using primers designed during the study to be specific only to the Spathebothriidea. Analysis of this fragment clearly shows *Bothrimonous* to be most similar to *Cyathocephalus*, having 92% pairwise identity to it and a branch length of 7 on the unrooted parsimony generated phylogram. The unrooted phylograms produced both via parsimony and Bayesian inference (Figs. 5 and 6) indicate the same relationships with *Cyathocephalus* as its closest association and *Didymobothrium* being the most divergent taxon. BLASTn (McGinnis 2004) results also identify *Cyathocephalus* as the most similar, with the other 4 spathebothriid taxa making up the top four hits. This result, showing the sequence to be most similar to the other Spathebothriidea coupled with the fact that the 103bps are unique provides strong evidence that the sequence is genuine. The finding that *Cyathocephalus* is the closest relative to *Bothrimonous* also appears to make sense in terms of host association. *Bothrimonous* is known from freshwater chondostreans in the holarctic region and *Cyathocephalus* from freshwater and euryhaline teleosts (mainly salmonids and percids) also from the holarctic region, thus placing the species both in freshwater hosts in the same general region. In contrast, the other species have marine or euryhaline species as their hosts and have not been reported from the holarctic region. (Gibson 1994). This finding is particularly inmportant as previously *Bothrimonous* had been synonymized with *Diplocotyle* and *Didymobothrium* by Burt and Sandeman (1969) who stated that the differences used to separate these genera were "neither sufficiently constant nor of sufficient stature to stand as valid criteria at the generic level". The data generated in this study reject the synonomy of these genera and indeed indicate that neither *Diplocotlye* nor *Didymobothrium* is the closest lineage to *Bothrimonus*.

Characterizing this short piece of sequence data from *Bothrimonous* required considerable effort due to the apparently highly degraded and fragmented condition of the DNA in the available samples. However as such museum material, whatever its condition, represents the only available material

for this species this is the only option by which sequence data can be acquired. It is highly unlikely that fresh material will become available again for this species due to the precarious position of many of its sturgeon hosts, which are among the most threatened of all fish species (Lendhardt et al. 2006). In the Serbian part of the Danube river, for example the last reported findings of *Acipenser nudiventris* (a known host for *Bothrimonus fallax*) were of 5 individuals between 1948 and 1954 and one finding in 2003 (Lindhardt et al. 2006). Indeed this species is considered to have been on the path to extinction in this region from the middle of the 20th Century (Lenhardt et al. 2005). A number of different factors have contributed to decline of sturgeon, such a dam building which restricts access to spawning groups. In addition the fish are affected by the high levels of water pollution and infected by diseases passed on by farmed fish. The result has been a dramatic decline in yearly catches (Bauer et al. 2002, Lindhardt et al. 2006) which must mirror a massive population decline. The now protected and endangered status of almost all sturgeon species means if these fish are encountered investigations of parasite fauna in the fish would not be permitted. In addition it appears that parasite numbers in sturgeon have also undergone a dramatic decline. The last available reports of a *Bothrimonus* specimen being found in a sturgeon were in the 70's (Bauer et al. 2002, Mokhayer et al 1972) and more recent investigations did not find any of the parasites and indeed reported an overall decline in parasite diversity (Sattari and Monkhayer 2005). This suggests that the parasite may now be extinct in Sturgeon in the Caspian Sea area.

When material can only be found in museum collections the genetic material will be highly degraded and damaged due to the age of the specimens and the use of fixing and preserving chemicals such as formaldehyde. As was discussed previously, formaldehyde is known to make DNA difficult to amplify by causing extensive degradation of the DNA and crosslinking it to proteins, making it difficult to extract from the surrounding tissue (Carter 2003, Chaw et al. 1980, Zimmerman et al. 2008). The effect of this is that work with such material is technically demanding (Wandeler et al. 2007). All but very short fragments (100-200 base pairs) of DNA prove extremely

difficult to extract and concentrations of DNA are low, either due to degradation or material being made inaccessible by crosslinking Low DNA concentrations in the samples also mean that chances of contamination by exisiting PCR products in the laboratory environment are high. Therefore, in this study a number of steps were taken and considerations made in order to maximise the chance of successful amplification from the samples. Nested PCR of the samples proved unsuccessful presumably as fragment sizes targeted were too large. Nested PCR is known to increase the chances of amplification of contaminants and products that were amplified by this procedure were indeed found to be contamination. To avoid this, short fragment, Spathebothriid specific primers were designed. The use of such primers is often highlighted as a key step when working with such material (Wandeler et al 2007). Work was also carried out in a laboratory separate to where previous spathebothriid samples had been characterised. Additionally PCR reactions were set up within a hood and equipment was treated with UV light prior to use. Negative and positive controls were used extensively at all stages. Such steps are considered crucial in order to avoid sample contamination and ensure confidence in the products amplified. Further steps that could be taken include repeating the amplifications and sequencing reactions a number of times, compete isolation of pre and post PCR laboratories, and repetition of experiments at another location to show reproducibility (Wandeler et al. 2007). However Wandeler et al. (2007) highlight that such steps may represent the ideal situation whilst in reality small quantities of precious DNA mean that numbers of PCR's that can be completed are limited, something which was an issue in this study.

The results of the short fragment amplification of museum specimens indicate a low success rate for this procedure despite all the care and considerations involved. This highlights the difficulty in amplification of such specimens and indeed previous studies on comparable material have also encountered difficulties. A study by Herniou et al. (1998) used a range of formalin-fixed, ethanol-stored helminth material from the Natural History Museum. Amplification of a number of

overlapping fragments of 18S was successful but they noted that non-target products were common (although cosmopolitan primers were used) and cloning was required in order to produce products clean enough for direct sequencing. Another study involving *Lingula* (a pseudophyllidean tapeworm) used formalin fixed specimens that had been in collections up to 24 years. Laboratory work was carried out where no tapeworm material had been handled previously and under sterile conditions. Quality and quantity of DNA recovered was variable, but in the majority of specimens, DNA was of low molecular weight and concentration. However amplification of sequences for material which had been preserved for 24 years was successful (Jingjing et al. 1999). A study involving frog specimens attempted to amplify large fragments from both 16S and nuclear rhodopsin gene found that either amplification was unsuccessful or very low DNA concentrations could be obtained from formalin fixed specimens especially as storage time increased.. The study suggested that the problems came in extracting the DNA initially from the surrounding tissue due to crosslinking (Zimmermann et al. 2008). In the present study consideration of the difficult in extracting the DNA meant that material was extensively homogenised using a bead beater prior to extraction to improve chances of amplification (although this may have further fragmented the DNA) and a specialised kit available from Qiagen was used (QIAamp™ DNA FFPE Tissue kit) which included a high temperature incubation step in order to partially reverse formaldehyde modification of nucleic acids.

Despite the challenges of working with formalin-fixed material the first fragment of *Bothrimonous* DNA was characterized in this study. The amplification of this fragment has allowed for the first time a molecular comparison between all 5 genera of Spathebothrridea and identified *Cyathocephalus* as the most similar genus to *Bothrimonus*, a conclusion which is plausible in terms of host associations and geographic ranges of the parasites. In addition it provided evidence for the rejection of the synonymization of *Diplocotyle* and *Didymobothrium* under *Bothrimonous*. This study also allowed an investigation of intra specific variation in *Diplocotyle, Cyathocephalus* and *Spaptebothrium*

concluding that *Didymobothrium* is the only genus which contains cryptic species. These results have thus provided new information about the interrelationships of the Spathebothriidea.

References

Bauer O. N., Pugachev O. N. and Voronon V. N. 2002. Study of parasites and diseases of sturgeons in Russia: a review. Journal of Applied Ichthyology **18**: 420-429.

Burt M. D. and Sandeman I. M. 1969. Biology of *Bothrimonus (=Diplocotyle)* (Pseudophyllidea: Cestoda). Part1. History, description, synonymy and systematics. Journal of the Fisheries Research Board of Canada **26**: 975-996.

J. D. Carter. 2003, The effects of preservation and conservation treatments on the DNA of museum invertebrate fluid preserved collections, M.Phil Thesis, University of Wales, Cardiff.

Chaw Y. F., Crane L. E., Lange P. and Shapiro R. 1980. Isolation and identification of cross-links from formaldehyde-treated nucleic acids. Biochemistry **19**(24): 5525-31.

Drummond A.J., Ashton B., Buxton S., Cheung M., Cooper A., Heled J., Kearse M., Moir R., Stones-Havas S., Sturrock S., Thierer T. and Wilson A. 2010. Geneious v5.1, Available from http://www.geneious.com.

Edgar, R.C. (2004), MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Research **32**(5), 1792-97.

Eglinton G. and Logan G. A. 1991. Molecular Preservation. Philosophical transactions of the Royal Society of London. Series B, Biological sciences **333**: 315-328.

Gibson D. I. 1994. Order Spathebothriidae Wardle and McLeod 1952. *In* Keys of the Cestode Parasites of Vertebrates. CAB International, Wallingford, p. 15-19.

Herniou E. A., Pearce A. C. and Littlewood D. T. 1998. Vintage helminths yield valuable molecules. Parasitology Today **14**(7): 289-92.

Jingjing, L., Xianghau, L. and Yang, H. 2000. Molecular characterization of a parasitic tapeworm (*Lingula*) based on DNA sequences from formalin-fixed specimens. Biochemical Genetics. **38** (9/10): 309-322.

Lenhardt M., Hegediš A. and and Jaric I. 2005. Action plan for sturgeon species management in fishery waters of Republic Serbia. *In* Institute for Biological Research "Siniša Stankovic",. Developed for Ministry of Science and Environmental Protection of Republic Serbia.

Lenhardt M., Jaric I., Kalauzi A. and Cvijanovic G. 2006. Assessment of extinction risk and reasons for decline in Sturgeon. Biodiversity and Conservation **15**: 1967-1976.

Littlewood D. T. and Olson P. D. 2000. Small subunit rDNA and the Platyhelminthes: signal, noise, conflict and compromise. *In* Interrelationships of the Platyhelminthes, Littlewood D. T. J. and Bray R. A. (eds.).

Littlewood D. T. J., Waeschenbach A. and Nikilov P. N. 2008. In search of mitochondrial markers for resolving the phylogeny of cyclophyllidean tapeworms (Platyhelminthes, Cestoda) - a test study with Davaineidae. Acta Parasitologica **53**(2): 133-144.

Maddison W. P. and and Maddison D. R. 2005. MacClade: Analysis of Phylogeny and Character. Evolution .

Mariaux J. 1998. A Molecular Phylogeny Of The Eucestoda. Journal of Parasitology **84**(1): 114-24.

Mariaux J. 1996. Cestode Systematics: Any Progress? International Journal of Parasitology **26**(3): 231-43.

Marques J. F., Santos M. J., Gibson D. I., Cabral H. N. and Olson P. D. 2007. Cryptic species of *Didymobothrium rudolphii* (Cestoda: Spathebothriidea) from the Sand Sole, *Solea lascaris*, off the Portuguese coast, with an analysis of their molecules, morphology, ultrastructure and phylogeny. Parasitology **134**(7): 1057-72.

McGinnis S. and and Madden T. L. 2004. BLAST: at the core of powerful and diverse set of sequence analysis tools. Nucleic Acids Research **32**: W20-25.

B. Mokhayer. 1972, Recherches sur le parasitisme des esturgeons de la mer Caspienne (côte iranienne). [Researches on the parasitism of the sturgeons of the Caspian Sea (Iranian coast).], Phd Thesis, University of Paris.

Moszczynska A., Locke S. A., McLaughlin D., Marcoglieses D. J. and Crease T. J. 2009. Development of primers for the mitochondrial cytochrome c oxidase gene in digenetic trematodes (Platyhelminthes) illustrates the challenge of barcoding parasitic helminths. Molecular Ecology resources **9**(Suppl. 1): 75-82.

Nylander J. A. A. 2004. MrModelTest, program distributed by the author. Evolutionary Biology .

Olson P. D., Littlewood D. T., Bray R. A. and Mariaux J. 2001. Interrelationships and evolution of the tapeworms (Platyhelminthes: Cestoda). Molecular Phylogenetics and Evolution **19**(3): 443-67.

Olson P. D., Poddubnaya L. G., Littlewood D. T. and Scholz T. 2008. On The Position Of Archigetes And Its Bearing On The Early Evolution Of The Tapeworms. Journal of Parasitology **94**(4): 898-904.

Poddubnaya L. G., Gibson D. I., Swiderski Z. and Olson P. D. 2006. Vitellocyte ultrastructure in the Cestode *Didymobothrium rudolphii* (Monticelli, 1980): possible evidence for the recognition of divergent taxa within the Spathebothriidea. Acta Parasitologica **51**: 255-263.

Renaud F. and Gabrion C (1988) Speciation in the Cestoda: evidence for two sibling species in the comples *Bothrimonus nylandicus* (Schneider 1902) (Cestoda: Cyathocephalidae). Parasitology (97): 139-147.

Ronquist F. and and Huelsenbeck J. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19**: 1572-1574.

Sanmartín Durán M. L., Quinteiro Alonso P., Rodriguez A. and Fernándeez A. 1989. Some Spanish cestode fish parasites. Journal of Fish Biology **34**: 977-978.

Sattari M. and Monkhayer B. 2005. Occurence and Intensity of some Parasitic Worms in *Acipenser gueldenstaedti, A.nudiventris* and *Huso huso* (Chondrostei: Acipenseridae) from the Southwest of the Caspian Sea. Turkish Journal of Veterinary and Animal Sciences **29**: 1279-1284.

Swofford D. L. 2001. PAUP*. Phylogenetic Analysis Using Parsimony (*and other Methods). Version 4.

Untergasser A., Nijveen H., Rao X., Bisseling T., Geurts R. and Leunissen J. A. M. 2007. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Research **35**: W71-W74.

Zimmermann J., Hajibabaei M., Blackburn D. C., Hanken J., Cantin E., Posfai J. and Evans T. C. J. 2008. DNA damage in preserved specimens and tissue samples: a molecular assessment. Frontiers in Zoology **5**(18).

Acknowledgements

Firstly I would like to thank my supervisor Pete Olson for his invaluable support and advice during this project. Additionally my thanks also go to the other members of his research group for their assistance as well as our collaborator R.Kuchta for providing the specimens. Staff in the molecular biology unit and sequencing laboratory of the NHM also played an important part in completion of this project. I would particularly like to thank Pat Dyal for her advice on molecular biology techniques and Sam Mohun for advice on working with formalin fixed samples as well as Julia Llewellyn-Hughes and Claire Griffin for sequencing of samples. In addition I would like to thanks Peter Foster for the use of the Beagle cluster and Robert Goodall for assistance with using it.

I would also like to acknowledge NERC for the financial support they have provided without which I could not have undertaken my MSc.