UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS

School of Ocean and Earth Science

Phylogenetic Systematics of the Phylum Nematoda – Evidence from Molecules and Morphology

by

Birgit Hildegard Maria Meldal

Thesis of the degree of Doctor of Philosophy

June 2004

Graduate School of the

Southampton Oceanography Centre

This PhD Dissertation by

Birgit Hildegard Maria Meldal

has been produced under the supervision of the following persons

Supervisors

Dr Alex Rogers

Dr John Lambshead

Prof Paul Tyler

Chair of Advisory Panel

Prof Patrick Holligan

Member of Advisory Panel

Dr Damon Teagle

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE & MATHEMATICS

SCHOOL OF OCEAN & EARTH SCIENCE

Doctor of Philosophy

PHYLOGENETIC SYSTEMATICS OF THE PHYLUM NEMATODA – EVIDENCE FROM MOLECULES AND MORPHOLOGY

by Birgit Hildegard Maria Meldal

The systematic relationships of nematodes have been investigated for over 100 years but phylogenetic reconstructions were impeded by the diversity of morphological characters and subjective interpretations by the investigators. However, knowledge of these relationships is imperative for all applied fields of nematology, such as biodiversity and ecology, diagnostics and pest control. Furthermore, nematodes are used as model organisms and systems for developmental mechanisms, the analysis of genetic pathways and population dynamics. Only since the introduction of molecular techniques, which offer the promise of a more objective analysis, have systematists been able to establish a potential consensus of the phylogenetic relationships of nematodes.

A growing database of small subunit ribosomal RNA (18S) sequences is available for nematodes but most sequences come from commercially important parasites. In order to fill the gaps in the database the 18S gene was sequenced for 40 species of marine nematodes, and the sequences aligned to the current database. Additionally, for the first time in the history of nematology, a cladistic analysis was performed on up to 180 morphological characters of 167 species covering the whole of the phylum.

Using only molecular characters, de Ley & Blaxter's three main clades, Enoplia, Dorylaimia and the Chromadoria were all recovered and well supported as was the monophyly of the Nematoda itself. The root relationship between the three clades could not be resolved. Most classic relationships and in particular families were recovered as monophyletic groups. Some species (and families) of uncertain placement could be assigned a firm place within the phylogeny (e.g. *Alaimus* sp., *Prismatolaimus intermedius*, *Campydora demonstrans*, *Xyzzors* sp., *Cyartonema elegans*, Comesomatidae, *Setaria digitata*, *Globodera pallida*, *Deladenus* sp.) whilst other relationships remain uncertain (e.g. *Teratocephalus lirellus*, *Brevibucca* sp., *Philonema* sp., *Gnathostoma turgidum*, *Myolaimus* sp., *Bunonema franzi*). Major changes compared to classic phylogenies include the placement of *Alaimus*, *Campydora* and *Prismatolaimus* in the Enoplia and *Cyartonema* and the Comesomatidae in the Monhysterida. The internal relationships of the Enoplia remain unresolved. The polyphyly of the Rhabditidae and Pratylenchidae has been confirmed.

A parsimony analysis of 180 morphological characters recovered many of the traditional families but the relationships of higher ranks bear no resemblance to older systems or to the molecular data. Many clades rather resemble habitat and trophic relationships; for example, the Dorylaimia are frequently found as sister taxon to the Rhabditida whilst they are generally understood to be the sister group to the Enoplia. The results from this analysis show that homology assessments of morphological characters are inherently difficult to determine accurately.

Whilst the whole data set contains RNA sequences from over 200 nematode species, many more sequences need to be obtained to resolve the root relationships of the Nematoda. A more slowly-evolving gene may be more suitable than the 18S gene. Conversely, faster-evolving genes, e.g. mitochondrial genes, are needed to establish many of the intra-familiar relationships. However, the Tardigrada are confirmed as the closest metazoan group to the Nematoda. A refined phylogeny of the phylum is proposed.

Table of Contents

List of Tables

List of Figures

List of Accompanying Material

CD-ROM containing the following files:

Appendix II

Table A2.1: <TableA2_1.rtf>

Table A2.2: <TableA2_2.rtf>

Appendix IV

Table A4.2: <TableA4_2.xls>

Table A4.3: <TableA4_3.rtf>

Table A4.4: <TableA4_4.rtf>

Appendix V

Table A5.1: <TableA5_1.rtf>

Figure A5.1: < FigureA5_1.rtf>

Table A5.2: <TableA5_2.rtf>

Figure A5.2: < FigureA5_2.rtf>

Appendix VI

Table A6.1: <molecular.nex>

Table A6.2: <morphology.nex>

Table A6.3: <combined.nex>

Declaration of Authorship

I, Birgit Hildegard Maria Meldal,
declare that the thesis entitled
"Phylogenetic Systematics of the Phylum Nematoda – Evidence from Molecules and Morphology"
and the work presented in it are my own. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission.

Signed: .....

Date: .....

Acknowledgements

First of all, I would like to thank my parents for enabling me to come to the UK to study for my undergraduate degree and to become a marine biologist. Without this start I would not have been able to study for this PhD.

I would like to thank my supervisors, Dr John Lambshead (The Natural History Museum, London, NHM), Dr Alex Rogers (British Antarctic Survey) and Professor Paul Tyler (Southampton Oceanography Centre, SOC), for their continuing advice, support and encouragement on a journey that at times felt incredibly lonely. Thanks also go to Dr Mark Blaxter (University of Edinburgh) for adding his comments to the final draft.

I would like to thank Nicola Debenham (NHM) for the identification of the marine nematodes and providing access to the Free-living Nematode Collection of the NHM. Thanks also go to Eileen Harris (NHM) for providing access to the Parasitic Nematode Collection and the Host-Parasite Database of the NHM, and Janet Rowe (Rothamstead Research, Harpenden) for providing access to their nematode collection. The analysis of the morphological characters would not have been possible without their corporation.

I would like to thank Ben, Bruce, Cath, Francisco, Jo, Kez, Maria, Marie and Rodrigo of the 'late' Molecular Ecology Research Group at the SOC and Adam, Natalie, Nicola and Tim of the Nematode Research Group at the NHM for their help and advice on practical issues as well as many fruitful discussions. A special thanks goes to Natalie who was always at hand to offer some floor space or even a bed when I stayed in London collecting morphological data and to Marie who kept encouraging me to keep on going when there were only the two of us of the original research group left in the SOC.

Thanks go to Dr Mark Dixon for providing access to the DNA sequencer at the School of Biological Sciences of the University of Southampton and to Dr Candida Rogers (formerly School of Biological Sciences, University of Southampton) for providing live *C. elegans* cultures.

I would like to thank Dr Mark Blaxter and Peter Mullin (University of Nebraska) for providing me with unpublished DNA sequences without which my data set would not have been as comprehensive as it is. I would like to thank Dr Mark Wilkinson, Dr Peter Foster and Dr Robert Hirt from the Department of Zoology at the NHM for expert advice on different aspects of phylogenetic analysis and especially Peter for providing me with external access to the LINUX computer cluster of the Molecular Biology Group. Thanks also go to Dr David Gibson of the Parasitic Worm Group at the NHM for his advice on parasitic nematode morphology.

I would like to acknowledge the School of Ocean and Earth Science of the University of Southampton, The Department of Zoology of The Natural History Museum, London, and the Society of Nematologists for proving funding for my research and the attendance of two international conferences.

Last but not least, very special thanks go to my husband Tor Ole. He always supported me throughout the years of my PhD. He was always there when things were not going smoothly and although he knows nothing about nematode systematics he continuously told me that I was doing a great job and that the hard work and sacrifices would pay off in the end. Let's hope he was right!

In short, if all the matter in the universe except the nematodes were swept away, our world would still be dimly recognizable, and if, as disembodied spirits, we could then investigate it, we should find its mountains, hills, vales, rivers, lakes, and oceans represented by a film of nematodes. The location of towns would be decipherable, since for every massing of human beings there would be a corresponding massing of certain nematodes. Trees would still stand in ghostly rows representing our streets and highways. The location of the various plants and animals would still be decipherable, and, had we sufficient knowledge, in many cases even their species could be determined by an examination of their erstwhile nematode parasites.

Nathan August Cobb, 1915

1. Introduction

1.1 The old, the new, and the controversial – nematode systematics throughout the ages

As so poignantly observed by the godfather of nematology, Nathan August Cobb, in 1915 (see previous page), nematodes are ubiquitous organisms. They form the third largest taxon in the animal kingdom by number of species preceded only by the arthropods and molluscs and they are the most numerous one in number of individuals. Nematodes are of economic importance as pests in agriculture and as animal and human parasites. Free-living species, on the other hand, play an important role as detritivores, bioturbators and bioindicators (e.g. pollution indicators). Nematodes are also widely used in experimental biology and medicine as model organisms for developmental studies and genetics. For example, the entire genome of the model nematode *Caenorhabditis elegans* (Maupas, 1900) is known (Okimoto *et al.*, 1991; The *C. elegans* Sequencing Consortium, 1998).

Hugot *et al.* (2001) made an inventory of all known nematode species. They counted 26,646 described species, of which 4,070 are marine and 6,611 terrestrial free-living species, 4,105 plant parasites, 3,501 invertebrate parasites and 8,359 vertebrate parasites. Estimates of the total number of nematodes vary from a conservative 500,000 (Hammond, 1992) to an average 1 million (May, 1988) to an upper limit of 100 million (Lambshead, 1993) depending largely on the mathematical model and species concept employed (Hugot *et al.*, 2001; Coomans, 2002). This means that no more than 5% (0.03% for Lambshead's estimate) of nematode species are known to science, the lowest ratio for any group of animals (Malakhov, 1994). At an average rate of 385 new descriptions per year it would take at least 1300 years to describe all nematode species (Hugot *et al.*, 2001). From the above figures it becomes apparent that more research is being conducted on vertebrate parasitic nematodes than all other groups. Therefore, it is important to redress the balance of research into all groups of nematodes. The following sections provide a historical overview of nematology (the study of free-living and plant parasitic nematodes) and helminthology (the study of all animal parasitic worms) which will shed some light onto the discrepancies of the above figures and give an insight into the troubled state of nematology today.

1.1.1 The early years

The oldest record for any nematode is a mention of the intestinal roundworm *Ascaris lumbricoides* Linnaeus, 1758, from China, 4690 years ago. The "Papyrus Ebers",

discovered by Ebers in 1872 and dating back to 1553 - 1550 BC, mentions two parasitic nematode species, the afore mentioned *Ascaris lumbricoides* and the guinea worm, *Dracunculus medinensis* Linnaeus, 1758. A number of other references to parasitic nematodes and their diseases are known from ancient times (Hippocrates, Aristotle, The Bible). However, it took another three millennia before the first free-living nematode, the vinegar eelworm *Vibrio aceti*, was described by Borellus (1656). Needham (1745) pictured and described the first plant-parasitic nematode, *Vibrio tritici*, and Müller (1786) described the first fresh-water and marine nematodes (e.g. *Vibrio gordius* and *Vibrio anguillula marina*). Between the 15th and 19th century most nematologists devoted their time to descriptions and investigations of the life cycles of parasitic species.

Only in the 19th century did the taxonomic descriptions of nematodes, including free-living species, become more wide spread. Major contributions to the description and classification of nematodes in the 19th century came from Rudolphi (1808), Oken (1815), Dujardin (1845), Diesing (1851, 1861), Leydig (1854), Eberth (1863), Bastian (1865), Schneider (1866), Bütschli (1873, 1874, 1976), Villot (1875), von Linstow (1876 - 1892), Örley (1880, 1886) and von Daday (1894 - 1910). Dujardin (1845) was the first to conduct detailed research into the free-living nematodes to which he gave the suprageneric name Enopliens. He established four genera, *Enoplus, Oncholaimus, Rhabditis* and *Dorylaimus*, all of which are still valid, including some of his original species. Bastian (1865) described 100 new species of 22 new genera of free-living nematodes. All of his genera are still in use, although some have a revised content, and many of his species remain valid to this day. However, he grouped all free-living nematodes under one family, the Anguillulidae, and divided these into soil and fresh-water inhabiting species and marine species. Bastian also made significant contributions to the anatomy of free-living nematodes. Schneider (1866) concentrated on the parasitic nematodes and barely touched on the free-living species in his monograph. He provided anatomical descriptions, histological information and detailed illustrations of the species he studied. Bütschli (1873 - 1876) worked with all types of free-living nematodes. He corrected many of Bastian's mistakes, described more species and added more information on anatomical characters. He discussed tentative relationships between genera but did not advance to higher levels of relationships. Örley (1880, 1886) constructed the first comprehensive taxonomical survey and the first usable key for free-living nematodes. He divided the order Nematoda into the Parasita, Rhabditiformae and Anguillulidae. Örley explicitly stated the marine ancestry of nematodes with a secondary colonization step of terrestrial and limnic genera back into the sea. Despite his realization that the free-living nematodes were paraphyletic he still

grouped all aquatic and many terrestrial genera (excluding the Rhabditidae) in the group Anguillulidae.

During the late 19th and early 20th century, two contemporaries made the probably biggest contribution to the science of nematology: the Dutch, Johannes Govert de Man, and the American, Nathan August Cobb (1859 - 1932). De Man proved to be Europe's most prolific nematode and crustacean taxonomist. Typically for European scientists of his time he published most of his papers in French and German, some in his native Dutch but none in English. Between 1876 and 1922 de Man described a wealth of new genera and species and included excellent drawings of a standard that has never been surpassed using light microscopy. In 1886 and 1904 he published two volumes devoted entirely to the detailed anatomy of nine marine nematode species, one of which was new to science. He also devised a formula (the so-called "de Man formula") that describes the body dimensions of each species and he introduced the dichotomous key. From a taxonomic and ecological point of view, his work was pioneering for Northern European nematode fauna surveys. However, despite his extensive work on taxonomy and anatomy he never attempted the classification of these disjunctive genera beyond the eight families he introduced in his 1876 paper (Ironidae, Dorylaimidae, Tylolaimidae, Odontosphaeridae, Ptychopharyngiae, Tripylidae, Monhysteridae and Odontopharyngidae). Five of these families (Ironidae, Dorylaimidae, Tylolaimidae, Tripylidae and Monhysteridae) are still valid today although with a revised content of genera.

Cobb (1888 - 1935) was de Man's counterpart in the New World. Unlike de Man, who stayed in Leiden (The Netherlands) for his entire career, the native American frequently travelled the world with extensive visits to Europe and Australia. Interestingly, he started his career in Jena, Germany, where he passed his PhD in 1888. He published a large number of taxonomic works on soil and plant-parasitic nematodes. Many of his works were collected in the "Contributions to a science of nematology" (1914 - 1935). Unfortunately, for newly described species he often provided only brief descriptions without drawings, which made recognition difficult in many cases. However, where he provided drawings they were of extremely detailed nature. Cobb introduced an alternative formula to de Man's that expresses the length of various parts of the body as percentage of the total body lengths and provided the corresponding relative body width. Cobb (1919) also provided the only key describing all free-living nematode genera known at the time. He was the first to classify the nematodes as an independent phylum but his use of the

morphology of the buccal cavity as the single important character for his classification forced many previously recognised natural groups apart.

1.1.2 The coming-of-age of nematode systematics

The early 20th century saw an explosion in the number of nematode surveys covering most of Europe, Russia and North America (e.g. G. Schneider, 1906, 1926, 1927; Jägerskiöld, 1901 - 1909; Ditlevsen, 1911 - 1928; Schwarz, 1911, 1912; Menzel, 1912 - 1929; Hofmänner, 1913 - 1915; Steiner, 1914 - 1945; Stefanski, 1913 - 1954; Micoletzky, 1913 - 1927; Filipjev, 1929; Kreis, 1928, 1929; de Coninck, 1931 - 1965; Schuurmans Stekhoven, 1931 - 1954; Wieser, 1951 - 1967). Most remarks on the relationships of the species covered did not advance further than the family level.

Filipjev (1918/1921, English translation 1968/1970, and 1934) drew up the first comprehensive classification of free-living nematodes with specific emphasis on the marine species. He disagreed with Cobb's (1920) assumption that nematodes are an independent phylum and ranked them as a class with eleven orders of which five contain free-living nematodes (Enoplata, Chromadorata, Desmoscolecata, Monhysterata, Anguillulata) and the remaining six are entirely parasitic (Oxyurata, Ascaridata, Spirurata, Filariata, Dioctophymata, Trichurata). He rejected the division into two independent subclasses of free-living and parasitic species because of their homogenous body form. However, Filipjev suggested that if there was any division of the orders into subgroups it should be between the Anguillulata (containing most free-living rhabditids and tylenchids) and the remaining four orders of free-living nematodes rather than between the Anguillulata and the six orders of parasites.

Micoletzky (1922) provided a detailed taxonomic key of the free-living soil nematodes including limnic species. He was forced to exclude the marine species as the Second World War prevented him from doing the relevant fieldwork. Micoletzky criticised Cobb's approach to using the buccal cavity morphology as a systematic criterion but nevertheless regarded a different interpretation of the same characteristics as the most suitable basis to a comparative analysis of anatomy. Micoletzky, however, did not attempt a reclassification of the nematodes. Baylis & Daubney (1926) attempted the classification of the whole group of nematodes but as they had little knowledge of the free-living genera their attempt found little support. Schuurmans Stekhoven (1935) divided the nematodes into two subclasses, the Nematoda errantia and Nematoda parasitica. He stated explicably that it was nearly impossible to compare the two groups because of the profound morphological differences and logistical complications as few nematologists were working

on species of both groups reflecting the deep split between nematologists and helminthologists.

Chitwood laid down the basis of the modern system of nematode taxonomy: First published in 1933, and revised in 1937 and 1950 (Chitwood & Chitwood), he divided the phylum Nematoda into two (apparently natural) classes, the Phasmidia and Aphasmidia, named after caudal sensory organs (phasmids) that are only present in the former group. His works included both, parasitic and free-living genera. Most revisions of his system focused on the discussion of whether the Nematoda is a phylum in its own right or a class, and on the number of subclasses, superorders or orders nematodes are divided into including the assignment of certain families into their respective orders (de Coninck, 1965). In 1958, Chitwood followed his colleagues and changed the terms of Phasmidia and Aphasmidia back to von Linstow's (1905) older terms of Secernentea ('secretors', species with an excretory system with lateral canals) and Adenophorea ('gland-bearers', species with caudal glands) respectively, thus avoiding the confusion with an order of insects called Phasmida. It should be noted that the extent of von Linstow's Secernentea and Adenophorea differs greatly to the modern classification.

Andrássy (1976) provided the first extensive and comprehensive approach to the classification of all nematodes. He rejected the two class/subclass division because he saw many common, positive characters and suggested three subclasses of the class nematoda: the Secernentia (almost identical to Chitwood's Secernentea), the Torquentia (instead of Chromadorida, but slightly different in extent) and the Penetrantia (instead of Enoplida, but slightly different in extent). However, he did not regard all three groups as of equal evolutionary age but proposed that the Penetrantia derived from an ancestor of the Torquentia and that the Secernentia evolved most recently from within the Plectoidea (Torquentia). The latter relationship is based on similarities between species of the Rhabditidae and Plectoidea, a connection already made by Chitwood & Chitwood (1950) and others. Coomans (1977) reviewed Andrássy's work and listed a great number of factual errors. He pointed out inconsistencies in the ranking orders, most pronounced in the case of the ancestry of the subclass Secernentia from the lower ranked superfamily Plectoidea. He also stated that Andrássy was by far not the first to criticise the classification of the Adenophorea as unnatural. In principle, Inglis (1964) had already set out the same three-way division of the nematodes as Andrássy and had questioned the monophyly of nematodes altogether. Other authors (Maggenti, 1963, 1970; Gadéa, 1973; Drozdovskiy, 1980; Adamson, 1987; Malakhov, 1994) also agreed on the three-way

classification of nematodes with a closer affinity of the Secernentea (Rhabditia *sensu* Malakhov; Rhabditea *sensu* Adamson) to the Chromadoria. J. Goodey (1963a) rejected Chitwood's systems in the revision of his father's monograph on soil and fresh-water nematodes (T. Goodey, 1951) based on the presence of phasmids in *Euteratocephalus crassidens* (de Man, 1880), an otherwise 'aphasmidian' nematode.

Lorenzen (1981) criticised the three-way division of nematodes based on extensive studies of aquatic free-living nematodes. A revision and subsequent translation into English of this work appeared in the same year as Malakhov's review (1994). Lorenzen accepted Chitwood's system of two main groups of nematodes and gave a detailed account on the plesiomorphies and holapomorphies on which he based his revised classification of the Adenophorea. In contrast to Inglis (1964) and Andrássy (1976), he argued that, within the Adenophorea, the Enoplia are more primitive than the Chromadoria. Lorenzen used twelve characters for his analysis, two of which had not been used previously for a classification of nematodes: the presence of metanemes in the Enoplida and the number of gonads and their position relative to the intestine. Lorenzen (1994) and Malakhov (1994) made the last major revisions of the phylogenetic systematics of nematodes using morphological characters.

1.1.3 The shortcomings of morphology-based classifications

It becomes obvious that nematode systematics is inherently problematic and controversial. Almost all classic attempts to establish a system for the Nematoda were based on the assessment of similarity of a small selection of characters, often a single character alone (e.g. Diesing, 1851, cephalic setae; Bastian, 1865, cuticular annulation; Schneider, 1866, somatic musculature; Cobb, 1919, stoma armature; Micoletzky, 1922, stoma structure (and ecology); Maggenti, 1963, 1983, pharynx structure and excretory system).

Andrássy (1976) slightly advanced on this system and compared the number of diagnostic characters shared between different groups of nematodes and between nematodes and other higher taxa. The shortcomings in his work lie in the lack of comparability, as he used different suites of characters for different pairwise comparisons. Hence, any numerical similarities established by Andrássy are relative and not comparable between different pairs of taxa. Nevertheless, as will be evident from the current study, his system introduced phylogenetic relationships that have stood the test of time and proved to be as valuable as other works that claim to have chosen a more 'suitable' approach (de Ley & Blaxter, 2002).

In his revision of the Adenophorea, Lorenzen (1981/1994) was the first to explicitly adapt Hennig's (1966) principles of phylogenetic systematics to the classification of the Nematoda. Whilst his approach was extensive and very detailed, its recognition of uncertain relationships was also its failure. For many groups, mainly at the family level, Lorenzen did not find symplesiomorphies with other higher taxa. Instead of providing possible tentative links to monophyletic groups he combined such taxa into potentially paraphyletic 'bucket groups' (de Ley & Blaxter, 2002), leaving its resolution to later generations of nematologists. The establishment of these 'bucket groups' also led him back to Chitwood's outdated two-class system instead of recognising the paraphyly of the Adenophorea.

As can be seen from the above discussion, all the classic morphology-based phylogenetic systems of nematodes are pragmatic and do not necessarily reflect the evolutionary relationships of the taxa. There are two main reasons for the controversy over the 'true' phylogeny of nematodes. Firstly, nematodes have a very simple body plan with few apparent characters above the ultrastructure level. Most species are also very small (>100 μm to a few millimetres), with the exception of some vertebrate endoparasites that can reach several metres in length. [The longest nematode known is *Placentonema gigantisma,* from the placenta of the sperm whale, at up to eight meters in length.] Many characters show gradual changes without providing clearly defined states (relative measurements, longer/shorter than a certain value), other characters exhibit diverse variations and homology assessments become very difficult (e.g. buccal cavity). Lorenzen (1981/1994) attempted to give each character a value and to explain its evolutionary history. However, several times he characterised the lack of a feature as a primary loss (e.g. the presence of metanemes in only the Enoplida) while defining the nematodes as a monophyletic phylum. Since evolution is the adaptation of a species to a change in environment featuring as either specialization or loss of characters, a missing character within a monophyletic group can only be interpreted as secondary loss. These secondary losses are of no phylogenetic validity because they can occur several times during evolution. Because comparison should only be made between homologies, i.e. features of the same evolutionary origin, similar looking features have to be tested for homoplasy. Many distantly related nematode species live in similar environments and similar-looking characters might just be the product of convergent evolution or parallelism, in which case they are analogies rather than homologies. This leaves very few useful common characters across the whole range of nematodes rather than only for closely related species. Section 1.2 explores the theoretical background of homology assessments in more detail.

The second problem is the lack of a significant fossil record for nematodes. Only few nematodes have been preserved in amber and they belong almost exclusively to the group of mermithids and tylenchids (Poinar, 1985, 2003). In the presence of a fossil record the evolutionary history of a character can be analysed much easier but without a comprehensive fossil record for nematodes homology assessments of morphological characters are based entirely on data from extant species.

In the second half of the last century, new technologies emerged that aided the observation and analysis of morphological characters. Microscopy advanced from the use of simple illumination of the 19th century to include a variety of filters (e.g. phase contrast, Nomarski filters) that can highlight various structures, often aided by specific staining procedures. A major revolution in microscopy occurred with the introduction of the electron microscope (EM) in the 1960s (Ferris, 1994). EM technology led to the discovery of a wealth of additional and rudimentary structures and characters, especially in the field of ultrastructure studies on, for example, the homology of the buccal cavity morphology and on the cuticle. However, this technique is relatively expensive and has been used mainly in the study of economically important parasitic nematodes (Dorris *et al.*, 1999). It is still little used in the investigation of many groups of non-parasitic species including most marine nematodes (exception: Urbancik *et al.*, 1996a, 1996b).

At the same time, selected nematodes were studied for embryological and post-embryological development and the history of cell lineages. Ontogenetic data can reveal the history of the evolution of a character, but until today studies in this field are still rather rare and focused on selected species (*Caenorhabditis elegans*, *Pristionchus pacificus*, *Enoplus* spp., *Pontonema vulgaris, Anoplostoma vivipara*) (Malakhov, 1974; Malakhov & Cherdantsev, 1975; Malakhov & Akimushkiva, 1976; Malakhov, 1977; Malakhov & Cherdantsev, 1977; Alberch, 1985; Malakhov, 1998).

1.1.4 The molecular revolution

Parallel to the technological development in the study of morphological characters, a new type of data has emerged: molecular characters. The main advantage of using molecular markers is that the mode of molecular evolution is better understood than that of morphological evolution (see Section 1.2). The rate of molecular evolution is thought to be the same as the rate of speciation and molecular techniques can give statistical support for the degree of divergence between species. In cases where fossils exist the genetic distance can also be converted into evolutionary time, establishing a molecular clock. The latter is not possible for nematodes because of a lack of fossils. Biochemical methods were

introduced into nematology in the 1970s, later than for other animal taxa (Hussey, 1979). The first molecular characters were based on the analysis of protein differences in the fields of serology, gel electrophoresis of isoenzymes and restriction site analysis (Nadler, 1990; Ferris, 1994). The protein-based techniques were successful in the study of population genetics only for large, mainly endoparasitic nematodes but were too insensitive for the small plant-parasitic and free-living species.

The discovery of the polymerase chain reaction (PCR) in the 1980s (Mullis & Faloona, 1987) and the availability of direct DNA sequencing technology in the 1990s allowed the analysis of population genetic structures and systematics of all groups of nematodes. Recent studies on the deep phylogeny of nematodes have used the small (18S) (Aleshin *et al.*, 1998; Blaxter *et al.*, 1998; Kampfer *et al.*, 1998) and the large (26/28S) (Litvaitis *et al.*, 2000) subunits of the nuclear ribosomal RNA (rRNA) genes. Other studies, looking at differences between species and strains of parasitic and terrestrial nematodes have used different mitochondrial genes and the two internal transcribed spacers (ITS) of the rRNA; the latter evolve faster than the rRNA genes themselves because they are non-coding (e.g. Hyman, 1988; Thomas & Wilson, 1991; Gasser *et al.*, 1993; Powers *et al.*, 1993; Campbell *et al.*, 1995; Zarlenga *et al.*, 1998; Hoberg *et al.*, 1999; Watts *et al.*, 1999; Nadler *et al.*, 2000).

Three landmark studies on the phylogeny of nematodes using the 18S gene were published in 1998. An Austrian group (Kampfer *et al.*, 1998) compared 18S sequences and 17 morphological characters of 13 'adenophorean' and 15 'secernentean' nematodes. Only the 13 'adenophorean' species were included in the morphological analysis. All 28 sequences were included in the molecular analysis that strongly supported the monophyly of each of the two groups. However, all but two of the 'adenophorean' species belong to the same family, the Desmodoridae (*sensu* Lorenzen 1981/1994), and all but one of the 'secernentean' species belong to the order Rhabditida (*sensu* Lorenzen 1981/1994). It is therefore of little surprise that the two main groups appear to be sister taxa when rooted on a non-nematode outgroup taxon.

A consortium of European and American nematologists (Blaxter *et al.*, 1998) sequenced over 50 nematodes from a wide variety of trophic groups. They concluded that (a) nematodes form five distinct clades, (b) the Chromadoria are paraphyletic, (c) the Secernentea form three clades and (d) parasitism is of multiple evolutionary origin within the nematodes. In more detail, Clades I and II are most basal to the tree and both consist of a mixture of parasitic and free-living Enoplia, with the marine Enoplida found in Clade II.

Clades III, IV and V are sister clades containing all Secernentea. Clade III contains three groups of vertebrate parasites and one group of invertebrate parasites previously not linked this closely together. Clades IV and V both contain a mixture of parasitic and free-living species. Representatives of the Rhabditida are found in both of these clades; therefore the Rhabditida are paraphyletic. The position of the Chromadoria could not be established definitely. However, the family Plectidae clustered most basal to the Secernentea, forming a "Secernentea + Plectidae" clade. The remaining Chromadoria were found basal to the Secernentea + Plectidae clade forming a "Chromadoria + Secernentea" clade. The division of the Chromadorida and Monhysterida within the Chromadoria could not be established because only one monhysterid nematode was sequenced. The gross picture that emerged from this study is that the Nematoda are divided into two major clades, the Enoplia and the Chromadoria plus Secernentea, with a deep division within the Enoplia. However, the authors themselves believe that the Chromadoria are more closely related to the Enoplia than to the Secernentea and that the secernentean ancestor can be found within the Chromadoria, possibly within the Plectidae. The multiple origin of parasitism was interpreted further by Dorris *et al.* (1999) and Blaxter *et al.* (2000).

The third study, by a Russian group (Aleshin *et al.*, 1998), came to a similar conclusion as Blaxter *et al.* (1998). Using fewer species but testing a variety of outgroups, they concluded that the nematodes are monophyletic and can be grouped into four clades: Clade I consists of part of the (paraphyletic) Rhabditida and the Strongylida, clade II of the Monhysterida, Plectida and all Secernentea (including Group I) and clade III of the Chromadorida, Desmodorida and Group II. The fourth clade, the Enoplia, are the most basal group to the tree, with a deep division that could not be resolved further. The overall division of the nematodes became most apparent in an analysis of "Hairpin 49", one of the characteristic loops in the secondary structure of the 18S gene that is normally very conserved. The authors found a region of an accumulation of 35 substitutions that they called the "secernentean stem". This differentiation was also found in plectids and monhysterids, was rudimentary in the remaining Chromadoria but not present in any representatives of the Enoplia.

Litvaitis *et al.* (2000) compared over 40 nematode sequences of the D3 extension segment of the 26/28S gene. Using constrained analysis on the major groups they came to the following conclusions: (a) the Secernentea are monophyletic, (b) the Rhabditidae are monophyletic after constraint analysis despite polyphyletic appearance, (c) the Adenophorea are paraphyletic and the synapomorphies are actually plesiomorphies for all

nematodes, (d) the Chromadoria are monophyletic after constraint analysis despite polyphyletic appearance, (e) the Chromadorida, Monhysterida and Desmodorida (*sensu* Lorenzen, 1994) are all monophyletic and together with the Desmoscolecida (and Leptolaimida?) form the Chromadoria, (f) there is a deep split within the Enoplia and (g) the closest living taxon to the Secernentea lies within the Plectidae (Chromadoridae) or Comesomatidae (tentatively Monhysterida, unresolved). Overall, they suggested the use of the term Adenophorea only in a historical sense and to establish two classes of Nematoda, the Enoplia and the Chromadoria.

These four studies exemplify the trend in the interpretation of molecular data of nematodes toward a two-class division of Enoplia and Chromadoria with the Secernentea as a group derived from the Chromadoria. While the relationship between the Enoplia and Chromadoria could not be resolved entirely there seems to be a tendency that the common ancestor of the Chromadoria may be found within the ancestors of the Enoplia and that the former class is more derived than the latter class. This interpretation is also supported by Adamson (1987) who performed a cluster analysis of 12 morphological characters and a few studies comparing features of the embryological development to the tree of Blaxter *et al.* (1998) (e.g. Voronov *et al.*, 1998; Voronov, 1999).

In 2002, de Ley and Blaxter wrote the introductory chapter to a new book on nematology and established a new system of the phylum Nematoda including all morphological and molecular data to date. Relationships that were strongly supported under bootstrap of molecular data were used as the basis of the new classification. Taxa for which no molecular sequences were available were added to the system based on established morphological characters. The phylum Nematoda was divided into two classes, the Enoplea and Chromadorea, with two subclasses in the Enoplea (Enoplia and Dorylaimia) and a single subclass for the Chromadorea (Chromadoria). The root is placed somewhere between the three subclasses, potentially rendering Enoplea paraphyletic. The major change to any previous systems is the new ranking of many 'secernentean' taxa. In order to preserve true phylogenetic nestedness, i.e. a derived taxon cannot be of higher rank than its ancestral group, the former (sub-) class 'Secernentea' has been downgraded to the order Rhabditida. As a consequence of the loss of at least one ranking level in this group, de Ley and Blaxter introduced the rank of infraorder to preserve the then currently accepted superfamilies and lower ranks.

In all above sections, the original taxon names according to the author under discussion were chosen. The classification and nomenclature of the remaining part of this

thesis is taken from de Ley and Blaxter (2002) unless it is otherwise quoted. References to the old taxa 'Adenophorea' and 'Secernentea' are only made in the historical sense. Similarly, owing to the advances in technology and especially in molecular character analysis, the status of the Nematoda as a phylum is now widely accepted (de Ley & Blaxter, 2002) and treated as such in this thesis.

1.1.5 A plea for integration

Whilst this thesis is focused on the systematics of nematodes it has wider implications for many fields of nematology, namely diagnostics, ecology, biodiversity and biogeography. In the last decade, a number of forum articles have been published in the journal "Fundamental and Applied Nematology", now "Nematology", highlighting the lack of future taxonomists, especially in the field of nematology (Ferris, 1994, Hyman, 1996; Hugot *et al.*, 2001; Coomans, 2002). Modern nematology attracts most funding for research into diagnostics and studies related to conservation issues with a focus on the application of molecular technologies to these fields. However, species can only be identified correctly by well-trained taxonomists unless barcodes exist (Floyd *et al.*, 2002), and estimates for conservation purposes have to be made on the base of a sound understanding of the relationships of species. Therefore, taxonomic background knowledge is of pivotal importance today as it has been in years gone by.

A second, and equally important point raised by the above mentioned authors is the lack of integration of scientists of the different fields of nematology. As discussed earlier in this chapter, this lack of dialogue is historical and experts in the different fields of nematology became more and more unable to look past the boundaries of their own discipline. With the advancement of molecular techniques renewed interest has been raised in the classification of the whole phylum (e.g. Blaxter *et al.*, 1998). It has become clear that scientists from all fields of nematology and helminthology need to work together to resolve this problem. The effort into integrating the different disciplines has shown limited success to date. In 2002, the Fourth International Congress of Nematologists saw for the first time a whole session devoted to aquatic and marine nematode biodiversity and with it the first attendance of delegates with a background purely in marine nematology. However, there were still a few "helminthologists", who held their own meetings, and the whole conference was dominated by sessions covering topics of agricultural nature.

The integration of nematologists from different disciplines is further hampered by the discordance in terminology. To give but one example: the anterior feeding apparatus is commonly known as the *buccal cavity* by aquatic nematologists, *buccal capsule* by

helminthologists and *stoma* by nematologists working with terrestrial nematodes and plant parasites. Some also include this region into the *pharynx*, the anterior digestive region which otherwise is known as the *oesophagus*. To complicate matters further, marine nematologists use the term *buccal capsule* for the specific strengthening and potentially doubling of the head cuticle. Following Hugot *et al.* (2001), there is urgent need for a clarification and standardisation of terminology if the classification of nematodes and the integration of nematology as a unified discipline are to be successful.

1.2 Phylogenetic systematics – theory and implementations

Ever since man had an interest in the natural world he has tried to understand the relationships and interactions of the different organisms living on this planet. Initially this was purely for survival reasons; if the hunter or gatherer was able to predict the next move of prey he was able to be a step ahead and had a better chance of success. Later on, philosophers of the ancient world (Egypt, Greece, and Rome) started to construct more advanced theories about mankind and its surroundings. But it took the great modern naturalists to understand the details of how our world works. Linnaeus (1767) established the first classification system of all animals and plants known by his time and created the binomial system of genus-plus-species names that is still used today. Ever since Darwin's theory of evolution and Mendel's analysis of inheritance scientists have improved our knowledge of the relationships among all living organisms on earth.

Phylogenetic systematics is the study of the reconstruction of relationships of a group of organisms in terms of their evolutionary history. Relationships are determined by similarity of comparable features. Initially these features were related to the outward appearance of the organism (morphology), physiological details and their behaviour and ecology. In the second half of the 20th century molecular characteristics were added to the pool of useful features. Molecular characters include physiological pathways, blood groups, protein characters and DNA, RNA or protein sequences. The most important part in the determination of relationships amongst different organisms is to compare the homologous characters.

1.2.1 Definition and treatment of characters

1.2.1.1 Character coding

Phylogenetic reconstructions are based on characters and their states. There has been some argument about the true definition of these in respect to a taxon and an individual organism. Jardine (1969) made an easy and logical distinction: characters refer to taxa while organisms are characterised by specific character states. Or, to put it into the context of a character matrix, characters are the independent entities in the data matrix (columns or rows) and character states are the values of the individual cells.

Characters can be divided into qualitative vs. quantitative and discrete vs. continuous variables. Most phylogenetic analysis techniques only accept discrete, qualitative data because continuous or quantitative characters are difficult to code as discrete entries in a

data matrix. Exceptions are those quantitative measurements that can be translated into qualitative characters such as relative lengths or ratios and similarities or distances because they can take on discrete numerical values. Overall, distinct, qualitative characters are those that distinguish groups and sister relationships unambiguously.

A character state that is thought to be ancestral to a sister-group relationship is termed *plesiomorphic* (shared and common) while a character state that is thought to have arisen from within this sister group is termed *apomorphic* (shared and derived). Furthermore, a character state that is found in the most recent common ancestor of one sister group, but not in its closest neighbour, is called a *synapomorphy* with respect to its closest neighbour but a *symplesiomorphy* with respect to any taxa within this sister group. A last situation occurs if a derived state is present only within one terminal taxon in a sister group, this state is called an *autapomorphy*. Autapomorphies are uninformative for phylogenetic reconstruction because they only define one terminal taxon and do not provide any information on sister-group relationships.

Only apomorphies are informative for cladistic reconstruction because they define specific sister groups, while plesiomorphies only give information about the commonality of a group, rather that its separation from its sister groups. Defining plesiomorphic and apomorphic character states relies on rooting the resulting cladogram (see Section 1.2.5), because these terms are dependent on the directionality of character change or more precisely, the evolutionary process itself. All the above principles were originally defined by Hennig (1966).

Characters that are used in phylogenetic reconstructions have to be of common evolutionary origin for all taxa in the data set, that is, they have to have been derived from a common ancestral character. Such characters are called *homologies*. Homologous morphological characters have to be derived from the same original structure; homologous molecular characters must originate from orthologous genes, i.e. genes that have the same evolutionary origin.

If the same character state is found in more than one terminal taxon, one would at first assume it to be a synapomorphy for a sister relationship of these taxa. However, with real data, there can be conflicting characters, where one character supports one sister-group relationship, and another an alternative. In this case, one of the characters cannot have arisen only once from a common ancestor of this sister group. This character must either have multiple origins (parallelism or convergence) or a later mutation must have changed it back into a previous state (reversal). Conflicting characters are called *homoplasies*.

Characters that support the same groupings are *consistent*. There are two types of consistent character: *compatible* characters are consistent within the same data set and are judged prior to cladogram construction, while *congruent* characters support the same phylogenetic hypothesis and are determined *a posteriori* (Sharkey's (1994) interpretation of Farris, 1971). Characters that are incompatible or incongruent, i.e. characters that are in conflict with other characters in the data set or established phylogeny, are *homoplasies*.

There are three different types of taxa within an established cladogram. In real examples, these groups often have Linnaean names. *Monophyletic* groups contain all descendants of a common ancestor and are defined by synapomorphies. *Paraphyletic* groups contain some but not all subgroups of a monophyletic group and are defined by a suite of symplesiomorphies; in other words, the lack of specific apomorphies defines the individual subgroups. *Polyphyletic* groups contain subgroups based on homoplasies and therefore do not contain true cladistic groups.

Synapomorphies are *homologies* that define monophyletic groups and are continually tested in a cladistic analysis. For true cladists, only synapomorphies are true homologies, and define true cladistic groups.

1.2.1.2 Linkage and covariation

Character coding has to be conducted with great care because most available algorithms for phylogenetic reconstructions interpret characters as independent evidence of relationships (Wilkinson, 1995). That is, all characters in the matrix are treated as independent entities. In biological reality however, there are numerous reasons why this assumption might be violated. The change from one character state to another in Character A might be directly linked to the change in Character B. This can be a result of functional dependence (such as the secondary structure of the ribosomal RNAs, see Section 1.2.3.4) or gene linkage. In an ideal situation, all characters can be clearly identified and are fully independent of each other. In this case coding would be simple and phylogenetic reconstruction would find the true tree, providing the reconstruction method is consistent. However, if there is insufficient evidence concerning character linkage, incorrect coding can result in construction of homoplasies and non-most parsimonious cladograms.

1.2.1.3 Binary versus multistate coding

There are two fundamental character-coding strategies: *semi-independent, binary character coding* (or reductive coding, Wilkinson, 1995) or *multistate character coding* (or composite coding, Wilkinson, 1995). In multistate characters all states (or combinations of

states) of a hypothetically homologous character are coded as independent states of one character (or all possible combinations), where one state denotes the absence of the character altogether and the remaining states identify the different phenotypes or its various combinations. Multistate coding minimises the effect of linkage, and hence the departure from independence (Pleijel, 1995) but the constituent binary characters of a multistate character cannot be tested independently. Alternatively, in semi-independent, binary coding the character states of the multistate character are treated independently. There are three variations for binary coding (Pleijel, 1995):

1) Each potentially independent feature of the multistate character is coded independently and has an additional state coding for the absence of this character altogether. Coding the absence of a character as an extra state can cause problems when the complex structure of a reductively coded character suite is entirely absent in a taxon. Duplicated coding of absence will overestimate the phylogenetic signal of the absent structure (Maddison, 1993).

2) One character codes for the presence/absence of the feature and additional characters code for the different potentially linked real characters. This method avoids the problem of duplicated coding but introduces questions about the treatment of missing data (see Section 1.2.1.5) because absence in the real characters has to be indicated as missing data ('?').

3) The last option is to code each state separately as presence/absence without attempting to account for potential character linkage. This option has the advantage that no account has to be made for missing data; however, no differentiation can be made between the three types of missing data.

Multi-character, binary state coding potentially gives more weight to the variation in the underlying independent character than multistate coding. If one change occurred in each of two independent characters in a binary coded matrix, it would add two steps to the total cladogram length whilst the same biological change would only add one step to the multistate-coded cladogram. Only if there is evidence of the underlying biological and evolutionary processes of a suite of characters, one could make an informed choice between the two coding strategies. Wilkinson (1995) suggested that if the character can be divided into subsets of characters and they do not covary, then it might be safe to assume biological independence and binary coding (or at least less complex multistate coding) should be used. If, on the other hand, two binary characters seem to vary directly with the states of their more complex multistate character, the more reduced characters are not

biologically independent and subsets of the complex character should be discarded to avoid overweighting of the phylogenetic evidence of the underlying variation.

Binary and multistate characters deal differently with homoplasy. If two different features of one structure are incongruent (see Section 1.2.7) binary coding will reconstruct the homoplasy on the cladograms. Multistate coding, on the other hand, cannot be incongruent within itself; thereby it cannot display any evidence of the underlying homoplasy of the components of the complex character.

1.2.1.4 Order and polarity

When conducting a parsimony analysis of multistate characters a decision has to be made on its transformation series (or optimisation). The character states may either change freely into any other state increasing the cladogram length by one step independently of the nature of the change. In this case, character change is *non-additive*, or *unordered.* Conversely, character transformation may be restricted according to a predefined order where a change from one state to another has to pass through all intermediate states. In *additive*, or *ordered*, transformation, the cladogram length is increased by the number of changes a character must have gone through even if this has not been observed directly. If the plesiomorphic state of the character is known, the character is *polarised. A priori* polarisation of characters requires additional information about the data, such as palaeontological or ontogenetic evidence. If character order and polarization are unavailable *a priori*, characters are treated as *unordered* and polarisation is read off the rooted cladogram (see Section 1.2.5). Until today, nucleotide sequence data are always treated as unordered, multistate characters. Differential weighting of transitions and transversions or successive weighting of all characters are used to apply constraints on a transformation series.

Characters can be fully or partially ordered (Wilkinson, 1995). In a fully-ordered character state network there is only one possible transformation series for each character change; the extreme case is the linear character state tree. If, on the other hand, a character state network has rejoining branches that allow alternative pathways for character transformations, the character is defined as incompletely or partially ordered. Wilkinson (1992) introduced the method of intermediates for the incomplete ordering of multistate, composite characters. The number of steps required to change from one state to another are coded in a step-matrix. Only incomplete ordering by the method of intermediates and the use of a step-matrix is analytically equivalent to reductive, binary coding and can retrieve the same most-parsimonious trees. Any other ordering or unordering of multistate

characters will lead to substantially different trees than those from binary coding of the same character suite.

Therefore, the choice between ordered and unordered character optimisation also makes an important assumption about the underlying evolutionary model (Mickevich & Weller, 1990, Wilkinson, 1992). Unordered characters simply ignore any similarity between its states and assume equal probability of all changes. Characters ordered by the method of intermediates make the assumption that neighbouring states are more similar to each other.

1.2.1.5 Missing data

As already mentioned above, there are cases where data are missing for some characters in some taxa. The underlying cause for missing data is threefold: the character state might be *unknown* (e.g. because of lack of sampling), the character might be *inapplicable* because it is difficult or impossible to score, (e.g. because a taxon does not possess this feature) or the character is *polymorphic* (i.e. two or more alternative states have been observed). In the case of a binary character (0 or 1), an unknown character can have states 0 or 1, an inapplicable character has neither state 0 nor 1 and a polymorphic character has states 0 and 1. The problem is that there are no algorithms available that can differentiate between unknown and inapplicable states but modern computer programs such as MacClade (Maddison & Maddison, 1992) and PAUP* (Swofford, 2002) can account for polymorphisms. An unknown character is substituted by all possible states but truly unknown data can only be resolved by additional sampling.

The debate continues on how to treat inapplicable data best. Characters containing unknown states for one or more taxa can be excluded altogether, thereby losing information for those taxa for which the character state is known. The unknown state can be coded as missing ('?'), which does not differentiate between the three possible interpretations. It can be coded as a new state or an arbitrary choice has to be made as to which alternative state is chosen. Alternatively, the taxon with the inapplicable character can be excluded from the analysis but this is a rather drastic measure and only applicable if many characters are missing (Nixon & Wheeler, 1992; Wilkinson & Benton, 1995).

The inclusion of missing data, coded as '?' in the matrix is likely to increase the number of equally best trees, i.e. equally parsimonious or most likely trees. In many instances adding more taxa to a data set will lead to a better resolution but this is not true in the case of missing data. As Platnick *et al.* (1991) showed, missing data increase the

amount of homoplasy and hence often reduces the resolution of the phylogenetic reconstruction. Missing data may also produce more resolved or more parsimonious reconstructions or spurious topologies all of which could not be recovered if the missing entry was replaced with all possible character states that would be allowed under plausible character evolution (Platnick *et al.*, 1991, Maddison, 1993).

Very recently, Wiens (2003) suggested that it is in fact the distribution of the missing data that produces the problems. If the missing cells are randomly distributed throughout the data matrix, as is common with molecular data, there are only a few complete characters available and the above problems apply. However, if the missing cells are concentrated within a few characters, such as in the case of a combined data set where a number of taxa are only scored for one of the individual sets, a large proportion of complete characters remain and the analysis can recover a reasonable phylogenetic hypothesis. These latest findings provide hope for the application of combined analyses with all available taxa even if some have not been scored for one of the data sets because of a lack of resources.

1.2.2 Character weighting

Differential treatment of characters can reduce homoplasy in phylogenetic reconstruction. As mentioned above, phylogenetic analyses often result in a number of equally parsimonious or most likely trees. One way of comparing the evidence of all these trees is by consensus analysis (see Section 1.2.7.1). However, consensus trees are always less powerful in explaining the data because they loose information compared to any of the fundamental trees (Mickevich & Farris, 1981; Farris, 1983; Miyamoto, 1985). An alternative to constructing consensus trees is character weighting. Character weighting assigns lower weight to those characters that introduce homoplasy compared to characters that are informative for phylogenetic reconstruction. Reducing the influence of homoplastic characters on the phylogeny also reduces systematic error (see Section 1.2.6). The weighting scheme can be employed before analysis starts (*a priori*) or after an initial run through the data (*a posteriori*).

A priori weighting is usually done by character analysis. Character analysis is simply the re-examination of the original data by character observation, definition of homologies and polarity estimation with the aim of weeding out homoplasies and inappropriate coding. It is an iterative process and the standard procedure for morphological data.

A priori weighting of molecular data is more diverse: Hillis *et al.* (1993) and Swofford *et al.* (1996) discussed the different weighting methods for molecular characters. The two most radical forms of character weighting are equal weighting or character deletion. However, assigning equal or zero weights is as much an assumption on the evolutionary history of the characters as assigning differential weights. Alternatively, molecular characters can be weighted across or within positions. Across position weighting entails partitioning of the data and treating the partitions under different assumptions. However, in recent years it has become evident that traditional partitions such as codon positions or structural features (e.g. loop vs. stem regions of ribosomal RNAs) do not evolve in clear-cut patterns and each partition has to be examined anew for each taxon (Dixon & Hillis, 1993; Källersjö *et al.*, 1999; Farris, 2001) (see Section 1.2.3.4).

Within-position weighting, or character-state weighting, differentiates between different base substitution rates of nucleotides in DNA sequences or synonymous versus non-synonymous changes of amino acids in protein sequences. For nucleotides, most commonly transitions are down-weighted compared to transversions (Miyamoto & Boyle, 1989) or differential weighting of base substitutions is incorporated in a substitution matrix (Swofford *et al.*, 1996). An effective application of character state weighting is the iterative search with weighted parsimony and likelihood. The substitution matrix is estimated using maximum likelihood, the tree space is then searched by parsimony and the resulting trees are then evaluated and optimised with likelihood. The advantage of this iterative search is computation time compared to a full likelihood search (Swofford *et al.*, 1996).

A posteriori weighing methods are the same for morphological and molecular data. There are two different types, successive approximations weighting (Farris, 1969, 1989; short: successive weighting) and implied weighting (Goloboff, 1993).

In successive weighting, characters are assigned the weight according to their level of homoplasy, i.e. their relative fit on the cladograms. The fit of the character can be measured as their (rescaled) consistency or retention index (see section 1.2.6.1). Successive rounds of weighting are used to stabilise the weights. Approximation is finished if two successive iterations yield no new weight estimates or new trees. Successive weighting does not reduce the number of cladograms recovered or choose between them. It only selects characters according to their fit to the cladogram. In reality, successive weighting may yield more cladograms than the initial analysis, may produce novel topologies and the new set of cladograms may also have a higher score. The

advantage of successive weighting is that the new trees will be more consistent and represent the best fit for all characters because ambiguous characters are down-weighted.

There are, however, two types of problems with successive weighting. Firstly, the iterative search for the best tree can be trapped in a local minimum in the same way as a heuristic search (see section 1.2.4.2) (Neff, 1986, Swofford *et al.*, 1996). Successive weighting can be effective if more extreme forms of weighting are used than the rescaled consistency index (Campbell & Frost, 1993; Swofford *et al.*, 1996). The second problem lies in the lack of an optimality criterion by which the degree of improvement can be measured (Maddison, 1990).

Implied weighting (Goloboff, 1993) uses the concave fitting function of cladistic consistency (Farris, 1969). Weights are based on implied homoplasy and a weight function and an optimality criterion is defined. The weights for all characters over the cladogram are summed up and the heaviest cladogram is selected (rather than the shortest). Uninformative characters are assigned zero weight compared to successive weighting where they are assigned maximum weight. An extension to successive weighting is dynamic weighting (Williams & Fitch, 1989). It combines character and character state weighting regimes and includes information on the relative frequency of the observed character change (Kitching *et al.*, 1998).

1.2.3 Specific issues related to the treatment of molecular characters

1.2.3.1 Mutations of molecular characters

The presence of different character states for the same character in different taxa is caused by mutation. Mutations either occur during faulty DNA replication or as a result of environmental factors such as ultraviolet light and gamma radiation or chemical genotoxicity. Only those mutations that occur in germ cells are inherited by the next generation. There are two general types of mutations: DNA mutations and chromosomal mutations. *DNA mutations* result in local defects in a gene caused by point mutations, insertions and deletions. *Chromosomal mutations* either change the structure of the chromosome by chromosome breakages or recombination or they alter the number of chromosomes.

Point mutations are the simplest type of mutation that lead to single nucleotide polymorphisms (SNPs). Point mutations are synonymous (or silent) if the complementary amino acid remains the same; they are non-synonymous if the resulting amino acid is changed. A change in the amino acid sequence or the structure of a functional gene (i.e.

RNA genes) can lead to serious deleterious effects. More severe cases of DNA mutations are deletions of sections of DNA and insertions of excised DNA into another region of the genome; *insertions and deletions* are collectively called indels. Insertions can occur near its original site or in a completely different chromosome. Indels in coding regions can cause frameshifts of the amino acid reading frame and often result in serious genetic diseases. SNPs and indels are the results of mismatching during DNA replication or of environmental factors.

Changes in chromosome numbers are the result of incomplete meiotic divisions. If the first division has not taken place at all the gamete cells contain the same number of chromosomes as the parental somatic cells giving rise to offspring with more than the 'normal' set of chromosomes. Polyploidy is more common in plants than in animals. Alternatively, in aneuploidy only a selection of homologous chromosomes fails to separate during the first meiotic division resulting most often either in trisomy (the existence of three homologous chromosomes) or monosomy (only one homologous chromosome). Aneuploidy can also cause serious genetic disease.

Chromosomal breakages cause four kinds of mutations: inversions, translocations, duplications and deletions. In each case the end of a chromosome breaks off and is re-attached in a slightly different fashion. Either the whole section flips over end-to-end (*inversion*) or it is attached to another chromosome (*translocation*). If the break occurs during replication and the loose end is reattached to the original chromosome the result is a *duplication* of part of the section than was broken off. During the detachment part of the DNA may also be lost resulting into a *deletion* of a chunk of DNA.

Whilst mutation rates can be rather high, most cells have repair mechanisms that reduce the effective number of mutations. One major exception is the mitochondrial genome that does not possess any repair system and therefore has a significantly higher mutation rate than the nuclear genome. In some instances, the DNA is repaired directly after DNA replication but this mechanism is rather rare. A common mechanism is excision repair where the damaged DNA is cut out of the strand and replaced using the complementary strand as a template. Mismatch repair is a type of proofreading where the DNA polymerase reads along the newly-synthesised DNA strand and checks for mismatches where A does not bind to T and C does not bind to G. Exonucleases then replace the faulty nucleotide with the correct one. The mismatch repair system in *Escherichia coli* reduces the mutation rate from ~1 bp per 10^8 bp after DNA replication to

~1 bp per 10^{10} bp after proofreading. Finally, homologous recombination (see below) can also act as a repair system.

Recombination is the result of mutations moving between chromosomes, both homologous and non-homologous ones. The result of recombination is that the progeny has genomic information that is different from both its parents. In eukaryotes, recombination usually takes place after the first meiotic division. It very rarely occurs during mitosis. Homologous recombination occurs while the newly-replicated homologous chromosomes lie adjacent to each other in preparation for the second meiotic division. During this time the non-sister chromatids cross over forming chiasmata. The chromosomes are split at the point of the chiasmata and a section of DNA is exchanged between the non-sister chromatids. Homologous recombination is the most common type, but quite frequently similar sections of DNA from homologous or non-homologous chromosomes are exchanged. Unequal crossing-overs are the result of a mismatch of genes during meiosis when similar, non-homologous genes are aligned next to each other. This commonly occurs in multigene families (such as nuclear RNA genes) and results in the duplication of similar DNA regions. Another special case of recombination and a mechanism for the evolution of multigene families is gene conversion. Here, one gene is replaced by a non-homologous but similar gene. Gene conversion does not lead to duplications but it can be biased towards certain genes that act as donor sequences more frequently. Gene conversion is the driving force of concerted evolution in multigene families.

Whilst many mutations are deleterious many others are neutral. Neutral mutations are passed on throughout the generations without any effect on its carriers. However, they may become advantageous when environmental conditions change giving its carrier a greater fitness. When neutral mutations become selective they can provide the mechanism for speciation.

All observable mutations are based on changes in the genome but only certain mutations are actually expressed in the phenotype. Character state differences in morphological characters are indirect indicators of genetic change. This is the reason why it is so difficult to code morphological characters for homologies because unless the exact underlying gene or genes are known there is no indication what kind of mutations have caused the change and whether there are one or several genes involved in coding a particular structure.

1.2.3.2 Sequence alignment

As mentioned above (see Section 1.2.1.1), one of the most difficult, and not very well understood, parts of phylogenetic analysis is the *a priori* decision of character homology. Whilst decisions of homology of morphological characters are essentially a problem of character coding, homology assessments for molecular data sets have to be made by a process usually referred to as sequence alignment. Even if the nucleotide sequences for the data set have been obtained using the same primers, the actual length of the sequenced segment might vary substantially as a result of insertion and deletion events. If a section of DNA is very variable and contains many indels, it is difficult to determine which bases are exactly homologous to each other and where to insert gaps to account for these mutations. Incorrect alignment of sequences will lead to homoplasies and incorrectly inferred cladograms. Whilst there are no perfect solutions, many different approaches have been suggested.

Determining homologies in a nucleotide sequence alignment is easiest when the topology has already been established. However, in most cases, the topology is the unknown variable that is the subject of the study. If the taxa that are being analysed are relatively closely related, homology assessment is often straightforward because not enough time has passed for long indels to accumulate. Most of the difference would result from SNPs and insertions of single gaps can even be done by hand. However, if more divergent taxa are analysed homology assessments are very difficult to do by hand and computer algorithms are usually employed. If alternative data sets are available for testing a sensitivity analysis can be carried out to assess character congruence between the alternative sets (see Section 1.2.7.2). From the indication of character conflict an alternative homology assessment of the alignment can be made and a new analysis can be carried out.

There are two types of alignments: pairwise alignments of any two sequences and multiple alignments of more than two sequences. All automated alignment procedures are based fundamentally on the Needleman-Wunsch (N-W) algorithm (Needleman & Wunsch, 1970). Full N-W methods are computationally very complicated and heuristic methods are used for their implementation in most multiple alignment algorithms and some pairwise alignment procedures. As with all heuristic search algorithms there are risks of only finding local optima, rather than global optima (Morrison & Ellis, 1997). Exact searches are implemented in automated algorithms for the pairwise alignment step but can also be used for multiple alignments of very small data sets.

All sequence alignment programs (for pairwise and multiple alignments) use an underlying evaluation criterion. The most common criterion is based on distance estimates (e.g. CLUSTAL_W, Thompson *et al.*, 1994; CLUSTAL_X, Thompson *et al.*, 1997; DIALIGN, Morgenstein, 1999; T-COFFEE, Notredame *et al.*, 2000; POA, Lee *et al.*, 2002). TREEALIGN (Hein, 1989a, 1989b) combines the distance criterion with a parsimony step and branch swapping. There are very few algorithms that use a pure parsimony criterion (MALIGN, Wheeler & Gladstein, 1994) or the likelihood criterion. Search algorithms employing the likelihood criterion are under development but so far they are computationally too expensive for the use on large data sets (Phillips *et al.*, 2000).

No matter which underlying criterion is employed, all multiple alignment algorithms are based on iterative methods to optimise a score function (Wilbur & Lipman, 1984; or minimum edit distance, Phillips *et al.*, 2000) that should reflect biological reality. During pairwise alignment costs are assigned to each step that is needed to change sequence A into sequence B. The score function is usually inversely related to the sum of all costs.

Gaps are inserted into the alignment to minimise the cost (or maximise the score function). There are usually two types of gap penalties: gap opening (GOP) and gap extension penalties (GEP). If the gap penalties are set high, the program will invoke more mismatches in order to keep the number of gaps small. Equally, if the gap penalties are set low (i.e. the mismatch cost is set high), the program will introduce many gaps in order to reduce the number of mismatches. The trick is to find the optimal combination of gap penalties to achieve the best alignment (Phillips *et al.*, 2000).

By assigning constant gap penalties to all types of gaps, one long gap would be weighted equally high as many small gaps of the same total length. In reality, one small or one long gap usually refers to a single indel event and a long gap should be downweighted compared to many short gaps of the same total length. However, length-dependent GEP are only implemented in a few alignment algorithms (e.g. Gotoh, 1982, 1996, 1999, fully automated; Kjer, 1995, by manual alignment; O'Brien *et al.*, 1998, including sequence weighting); most programs implement constant GEP independently of gap length because it is computationally less expensive.

Automated alignment algorithms, independent of the underlying search criterion (distance, parsimony, likelihood), are based on progressive programming (Feng & Doolittle, 1987, for distance-based algorithms): initially, a similarity score for each pairwise alignment is calculated and a guide tree (usually based on the neighbour-joining algorithm) is drawn based on these similarities. The multiple alignment is calculated by

starting with the most similar pair of sequences and subsequently adding all sequences according to the guide tree relationships. The similarity scores are the sum of all costs (or penalties). Gaps are introduced to minimise these costs (Morrison & Ellis, 1997). The guide tree topology depends on the sequence input order, but if the gap assignment is unambiguous, all resultant phylogenetic trees from the multiple alignment are the same. A drawback of many alignment programs (e.g. CLUSTAL_W, Thompson *et al.*, 1994) is that they are unidirectional, that is, they start at one end of the sequence alignment and only use the information available up to the point they have reached. Few programs (e.g. PRRN, Gotoh, 1999) go over the alignment again from the other direction to correct for mismatches.

Several studies have been carried out to test the performance of different alignment algorithms (Titus & Frost, 1996; Morrison & Ellis, 1997; Gotoh, 1999; Hickson *et al.*, 2000; Phillips *et al.*, 2000; Lassmann & Sonnhammer, 2002). The overall consensus is that the performance depends on the type of data at hand. With more and more divergent sequences, most alignment algorithms perform increasingly poorly, but the quality improves with increasing sequence length (Lassmann & Sonnhammer, 2002). To get the best estimate of the optimal alignment, it is advisable to use different alignment programs, to vary the parameters or even to change the *a priori* hypotheses (guide tree) and then to compare the resulting alignments and topology estimates. However, running multiple assessments is very time consuming, especially for large data sets, and it is often difficult to decide which of the resulting alignments should be selected for the phylogenetic analysis. Furthermore, the globally-optimal alignment is not necessarily the true alignment as the maximum similarity can result from divergence, parallelism or reversals rather than true homology (Morrison & Ellis, 1997).

A potential solution is an iterative approach of sequence alignment and tree search (Kjer, 1995; Wheeler & Gladstein, 1994; Gotoh, 1996, 1999; Lee, 2001). As described by Lee (2001), this method is the only way of truly estimating the gap penalty parameters for the alignment algorithms. Unfortunately, this procedure is also very time-consuming with moderate to large data sets because of the computational requirements of phylogenetic reconstruction modelling. Furthermore, no exact searches can be used for large data sets and therefore the reconstructed hypothesis might not be the correct one and alignment adjustment on a wrong tree does not improve the final hypothesis.

In reality, most investigators choose one program and use either the default parameter settings or their own set of parameters with few or no further rounds of

estimations. They would then use the one final alignment or make further manipulations by hand (e.g. in BIOEDIT, Hall, 2001). As mentioned above, long strings of gaps can cause problems and it is often feasible to check the alignment by hand and to make some manual alterations. Manual adjustment of sequence data has been criticised for its lack of objectivity, repeatability and objective criteria in homology assessments and data removal (Gatesy *et al.*, 1993; Wheeler *et al.*, 1995). However, there is one exception: If the data come from a gene with a strong secondary structure (e.g. ribosomal RNA), manual alignment is often the only option if the investigator wants to include secondary structure information (see Section 1.2.3.4 below).

1.2.3.3 The treatment of gaps in phylogenetic analysis

The most conservative approach is to exclude all positions from the analysis that have a gap for any one sequence in the alignment. However, for taxa that are closely related to each other, this section might be aligned correctly and contain a lot of information while the gaps were caused by the addition of more distantly related taxa. Smith (1994) suggested excluding only those positions for which no unambiguous alignment options could be determined. If only one or a few sequences per character contain an ambiguous state this could be coded as unknown ('?'). Such a situation could occur when outgroups are aligned to the ingroup sequences.

Lee (2001) and Phillips *et al.* (2000) emphasise that the exclusion of gapped positions is not justified because they are an integral part of the positional homology assignment process. Some positions that are excluded for the whole alignment would still be informative for a subset of taxa that are closely related to each other and the exclusion thus reduces the informative content of the data set for the phylogenetic analysis. Exclusion reduces ambiguity closer to the root of the tree but also makes branch lengths shorter and more prone to sampling error and reduces resolution close to the tip of the tree.

If gaps remain in the analysis they can be coded either as missing data or as an additional character (5^{th} base for nucleotide data, 21^{st} amino acid for protein sequences). However, the latter approach is problematic because independent gap positions would be treated individually when in fact they are caused by indels and therefore may not be entirely independent of each other.

1.2.3.4 The use of ribosomal RNA genes in molecular systematics and phylogenetic inference

Ribosomal RNA genes have been used in phylogenetic inference studies across the whole of the tree of life (Hillis & Dixon, 1991). Ribosomal RNAs (rRNAs) are vital constituents of ribosomes, the sites of protein synthesis in the cell. Because all life depends on proteins, rRNAs are found in all life forms and therefore can be used for comparisons amongst the most diverse groups of organisms. rRNA or rDNA also exhibits a wide range of evolutionary rates in different genes and different parts of the same gene. This versatility and the ease of amplifying rRNA genes, because of the presence of multiple copies in each cell, make it one of the most commonly used genes in molecular systematics.

There are generally four nuclear rRNA genes and two organellar (mitochondrial and chloroplast) rRNA genes. In Eukaryotes, three of the nuclear genes occur in an array starting with the external transcribed spacer region (ETS), followed by the small subunit of the rRNA (18S), the first internal transcribed spacer (ITS-1), the 5.8S gene, the second internal transcribed spacer (ITS-2) and the large subunit (28S). The ETS and both ITS regions contain signals for processing the rRNA transcript (Hillis & Dixon, 1991). They often show intragenomic variation and are usually only used for analyses of intra-specific relationships. The small subunit (18S) is generally the one with the slowest rate of evolution and therefore used for reconstructing deep phylogenies including the tree of life (Hillis & Dixon, 1991). The large subunit (28S) has some regions that evolve faster than the small subunit and some regions that evolve as slowly as the small subunit. The two smallest units (5S and 5.8S) have been used for phylum level analyses but in general are too small to be used on their own for robust phylogenetic reconstructions (Halanych, 1991). Organellar rRNAs have higher mutation rates and are therefore used mainly for the inference of relationships of closely related taxa.

A unique feature of rRNAs is the existence of multiple copies in the genome caused by unequal crossing-over. The multiple copies evolve in concert (Arnheim *et al.*, 1980) aided by gene conversion (see Section 1.2.3.1). Concerted evolution in rRNA genes also reduces the amount of intraspecific variation compared to that found in single-copy genes. This means that a smaller sample size for each species is needed to ensure that sufficient genetic information is compared between species. In practice, a single specimen per species is often used as a representative in an analysis among a wide range of species (Hillis & Dixon, 1991).

rRNAs are functional units of the ribosomes so their secondary structure is of primary importance. rRNA genes are folded in a predefined manner, forming paired regions (called stems or helices) and unpaired regions (called loops) (Noller, 1984). Although mutations of the primary structure occur, the secondary structure is often preserved over long periods of evolutionary time in a wide variety of organisms by compensatory mutations in stem regions (Mindell & Honeycutt, 1990; Hillis & Dixon, 1991). If the secondary structure of a gene can be determined, it is of enormous help in the alignment step of the analysis because homology estimates can be made with greater confidence and varying evolutionary rates in different parts of the gene can be taken into account (e.g. Kier, 1995; Hickson *et al.*, 1996; Titus & Frost, 1996; Page, 2000; Hickson *et al.*, 2000).

There are several publicly accessible databases that contain published sequences that are aligned using secondary structure information. These alignments can be downloaded from the internet and used as profiles for newly acquired sequences (e.g. The European RNA Database, Wuyts *et al.*, 2002, http://rrna.uia.ac.be/ssu/list/Eucaryota.html).

Compensatory mutations present difficulties when using rRNAs in phylogenetic inferences because they change the evolutionary rates. Studies on different genes and different organisms resulted in incompatible conclusions on the conservation and phylogenetic resolution of paired and unpaired regions of the rRNA genes. Some studies (e.g. Wheeler & Honeycutt, 1988) suggested that paired regions contain spurious compensatory changes and therefore reduced phylogenetic signal. It was suggested to eliminate either paired regions or at least to give them half the weight of the unpaired regions. On the other hand, the functionally less constrained loops might accumulate many spurious mutations and most phylogenetic signal would be contained in the stem regions alone (Morrison & Ellis, 1997). In such cases, the stem regions show phylogenetic topologies similar to those from other character sets (e.g. morphology) and to those from the analysis of all characters in the gene (e.g. Smith, 1989; Hillis & Dixon, 1991; Morrison & Ellis, 1997). Analyses of all characters or those from stem regions alone might also produce a more resolved phylogeny than using only characters from loop regions. Morrison & Ellis (1997) therefore suggested the use of separate objective criteria for assessing the phylogenetic informativeness of different regions within the sequence.

More recently it has become evident that conservation is not confined to stem regions alone. Some loops can be as functional as stem regions (Hickson *et al.*, 1996) and some highly conserved stem regions might contain hypervariable positions (Kjer, 1995).

Phylogenetic signal can be found in the same molecule in loop and stem regions because conservation is related to functional importance rather than structural position.

Hancock & Vogler (2000) suggested caution in the use of secondary structure information in alignments of hypervariable regions. Under high evolutionary rates indels are caused by slippage of single and di-nucleotides (mainly A and T) that can be self-complementary. Hancock & Vogler claim that the resulting stem formation is a mere by-product of these specific mutations rather than a result of the functional constraints of the secondary structure of the molecule. However, in most cases, the use of information on secondary structure has been proven helpful for aligning moderately to well-conserved regions (e.g. Kier, 1995; Hickson *et al.*, 1996; Titus & Frost, 1996; Hickson *et al.*, 2000; Page, 2000).

1.2.4 Phylogenetic reconstructions

1.2.4.1 Different types of trees

While homology assessment of the underlying characters is of critical importance to the reconstruction of phylogenies, the main goal for most researchers is the resulting hypothesis of the evolutionary history of the taxa under investigation. These reconstructions can be visualised as trees. There are three different types of trees. *Cladograms*, or non-metric trees, simply show the relative relationships of the taxa. The branch lengths on cladograms have no meaning, i.e. they are independent of the number of character changes that occur along them. On *phylograms*, also called additive or metric trees, the branch lengths represent the number of evolutionary changes encountered along that branch and give the investigator additional information about where character changes occurred. *Ultrametric trees* are a special case of metric trees. Ultrametric trees assume a constant molecular clock; all terminal taxa are equidistant from the root of the tree and the branch lengths represent actual divergence times.

In addition to the above-described properties of trees, cladograms and phylograms can be rooted and unrooted. Most computer programs initially construct unrooted trees. These contain information about the relative relationships of the taxa but not their evolutionary origin, i.e. they have no direction. Rooting of a tree provides information about ancestors (real or hypothetical). Rooting of a tree requires additional evidence that can be ontogenetic or palaeontological (from fossils) or trees can be rooted using outgroups. Tree rooting is discussed in more detail in Section 1.2.5. If such additional evidence is not available, the tree has to remain unrooted. Unrooted trees can be rooted

arbitrarily on any terminal taxon for display, but this is not a true root. Ultrametric trees are rooted by default.

In the absence of any homoplasy, a perfectly bifurcating tree is reconstructed; i.e. each node connects only three branches, one branch leading to its ancestor and two branches leading to the two daughter taxa. In the presence of considerable amounts of homoplasy the phylogenetic relationships cannot be determined completely and some nodes have more than two daughter lineages, they form (soft) *polytomies* (Kluge, 1989; Maddison, 1989). Sometimes, polytomies are interpreted as simultaneous speciation events (Maddison, 1989), but this interpretation is controversial as it is not likely that more than one new species arises at the same time. If the evolutionary rate of the characters under investigation is too slow for the true lineage divergence speed, hard polytomies can be constructed, but they do not represent the real evolutionary process.

Cladograms reconstructed under strong cladistic principles, i.e. minimum-length trees, do not represent the actual evolutionary history of a group of taxa that poses problems when interpreting such cladograms in terms of true character evolution. Furthermore, an actual ancestral taxon in a tree will always be reconstructed as a sister taxon in a cladogram but can be placed at the correct internal node in a phylogram reconstructed under non-cladistic principles.

1.2.4.2 Tree searching

For very small data sets (few taxa and characters) the cladogram can be constructed by hand but most data sets are too large and computer algorithms have been developed for the calculations. All automatic algorithms can be grouped into two types: those that search for all possible topologies (exact searches) and those that only sample a certain range of the tree space (heuristic searches).

For data sets of eleven taxa or less, *exhaustive searches* can be conducted. All possible trees are constructed and scored under an optimality criterion and the best tree is selected for output. For data sets with up to twenty taxa the branch-and-bound method (first applied to evolutionary trees by Hendy & Penny, 1982) can be used. This method starts with a random tree containing all taxa. It then backtracks one step disconnecting the last added taxon and adding it to all other internal branches. If the score of a new topology is larger than that of the initial tree, this topology is not pursued any further. If the score is lower, the new topology will be stored. When the best topology for this round is found, the procedure backtracks a further step. It retains the best tree with n-1 taxa, and reconnects

the second from last added taxon to all possible branches. Then, the n-2 tree with the best score is retained. This process is repeated until the algorithm reaches the initial three-taxon tree it started off with. In the worst-case scenario, all possible trees have to be examined, effectively conducting an exhaustive search. If the new topology has the same scores as the current best tree, then it is also further examined.

If there are many taxa, exact searches are usually too time-consuming and approximations have to be used. *Heuristic search algorithms* try to find the best score under an optimality criterion (e.g. minimum tree length under parsimony, or maximum likelihood score). Searches are conducted by constructing an initial tree, altering it (see below), and then comparing the scores for the alternative trees. If a tree with a better score is found, the search continues with this tree. There is no guarantee that the tree with the best score at the end of the search is actually the globally best solution because heuristic searches only accept trees with an equal or better score than the best tree at any time. The problem lies in the fact that the tree space can have several local optima but only one global optimum. If the tree search finds one of those local optima, any further rearrangements will not find a shorter tree and the search is stopped without finding the global optimum. However, since the tree space is not known *a priori* the investigator has no knowledge if the best tree at the end of the search corresponds to a local or the global optimum.

If a local optimum is found, only the acceptance of an intermediate worse topology can eventually find the global optimum but most algorithms do not allow the search to accept such topologies (for an alternative see Bayesian inference, Section 1.2.4.3.5). The only option is to run several searches, starting with a different topology each time. The results from each independent search can be compared and the best topology accepted. If the replicates repeatedly find the same best score or only one best score is found, the search is at least close to the global optimum. If the last repeat still finds new islands, then it is quite certain that the global optimum has not yet been found and more searches need to be conducted.

Apart from the chance of entrapment in local optima heuristic searches are 'greedy' algorithms. That is, once a taxon has been added to a particular branch its position cannot be changed anymore, even if at a later stage, this option turns out to be suboptimal. However, the branch-swapping process (see below) provides a means for reshuffling the taxa after the initial tree construction and thereby allows some rearrangement.

Most heuristic search algorithms are divided into two phases: *tree building* and *branch swapping*. During the first phase the taxa from the matrix are connected to each other in a certain order to construct the initial tree. In the second step, the branches of the initial tree are rearranged by disconnecting a terminal taxon or subtree from the initial tree and reattaching it to a different place in the tree. The second step is the important one because it provides the chance of finding a better tree than the one constructed initially.

The most commonly used tree-building method is *stepwise addition*. The algorithm starts of with three taxa from the matrix. Subsequent taxa are added to all possible branches and the topology with the best score is retained. This procedure is repeated until all unplaced taxa are added to the topology. There are a number of options for the choice of the initial three taxa and the decision of the addition sequence of the remaining taxa. The simplest option is to start with the first three taxa in the matrix and to add the unplaced taxa in the order they appear in the matrix. In a slightly more elaborated procedure (Farris, 1970), the search starts with a user-defined reference taxon and a matrix of relative similarities between taxa. The reference taxon and the two closest taxa are connected first, and then the other taxa are added according to the similarity measure. Swofford (1993) further developed Farris' procedure. The lengths for all three-taxon networks for the whole matrix are calculated and the triplet that gives the shortest tree provides the starting point. At each successive step, all remaining taxa are placed onto each branch and the taxon-branch combination with the best score is chosen. This procedure is repeated until all taxa are added to the tree. In comparison with Farris (1970), the procedure of Swofford (1993) calculates the similarity measure anew after each taxon addition. The latter method takes more computational time but theoretically increases the chance of finding the best tree. The most extensive option is to randomise the addition sequence and run several searches in sequence. This way, the investigator has some idea of how effective the search is and hopefully, one of the random replicates will find the global optimum. If the last random replicate still finds new islands (local optima) it is likely that the search has not yet found the global optimum and more replicates have to be performed.

An alternative to stepwise addition is star decomposition. Instead of starting off with a fully binary tree, all taxa are connected as a star tree to just a single internal node. In the second step, all possible pairwise terminal groups are constructed and the best option under an optimality criterion is chosen. Then, all pairwise joints are constructed again and evaluated and the topology with the best score is accepted. This process is continued until the tree is fully binary. Star decomposition is a clustering method (see Section 1.2.4.3.1).

The most common algorithm is the neighbour-joining method of Saitou & Nei (1987). Star decomposition is just as greedy as stepwise addition and there is no guarantee of finding the global optimum either.

As detailed above, both stepwise addition and star decomposition are likely to find only a local optimum unless the data set is small or homoplasies are rare. However, branch swapping can improve the chances of finding the global optimum by disconnecting a terminal branch or subtree from the best tree found during the tree building process and attaching it to another branch of the tree. If a tree with a better score is found the new tree is accepted, if not, branch swapping continues on the old tree. This process can be repeated many times until no better tree can be generated. Again, this is a heuristic method and there is no guarantee of finding the global optimum but it usually improves the initial tree. Branch swapping can be improved by swapping all trees with an equally best score instead of discarding any new trees of the same best score but different topology.

As with the tree building itself, there are several different approaches to the branch swapping process. During *local branch swapping*, a terminal branch from one end of an internal branch is swapped for a terminal branch from the other end of the internal branch. This procedure is called the 'nearest-neighbour interchange' (NNI). *Global algorithms* are a bit more complicated. In 'subtree pruning and regrafting' (SPR) a rooted subtree is clipped off the cladogram and attached to all other internal branches. If a better score is found for any of the new arrangements, the new topology is accepted. This process is repeated with all possible subtrees. During 'tree bisection and reconnection' (TBR), the tree is clipped into two or more subtrees. The subtrees are rerooted and reattached to give a new topology. TBR is the most sophisticated branch-swapping algorithm but it can be unfeasible to use for large data sets because of the computational time requirements.

1.2.4.3 Inferring phylogenies

Phylogenetic inference attempts to reconstruct the evolutionary history of a set of taxa. Methodologically, phylogenetic inference methods can take two different approaches to finding 'the best tree': *purely algorithmic methods* that combine tree search and evaluation and *criterion-based methods* that separate the two processes. Purely algorithmic methods are much faster, but theoretically less accurate than criterion-based methods because they only recover one best tree without any indication of whether this is the single best tree or one of many optimal trees.

Criterion-based methods divide the search for the best tree and its evaluation into two separate steps. In the first step, an initial tree is constructed and evaluated according to an *optimality criterion* (or objective function): The tree is given a score by which it can be compared to subsequently-constructed trees. In the second step, a specific algorithm is used to find more trees and compare them by their score (minimum or maximum, depending on optimality criterion). The trees with the best score are kept. Optimality criteria are parsimony, likelihood, minimum evolution and least-square-fit.

An advantage of the use of criterion-based methods is that alternative trees can be ranked according to their score. If there are many alternative trees with the best score, then under perfect search conditions these will all be recovered. A purely algorithmic method will only recover one tree and the investigator has no knowledge about its reliability.

Despite the disadvantages for purely algorithmic methods, they are still widely used because they require far less computational time. The application of criterion-based methods can become infeasible for large data sets (especially the likelihood criterion). In such circumstances, purely algorithmic methods are a good initial estimate of phylogenies that can then be further tested under an optimality criterion. This can significantly reduce computational time.

1.2.4.3.1 Algorithmic methods

There are two commonly used algorithmic methods, UPGMA and neighbour-joining. *UPGMA* (unweighted pair group method using arithmetic averages) is a clustering method that is based on the assumption that the underlying data are ultrametric, i.e. they assume the presence of a molecular clock. The molecular clock assumes a constant rate of evolution and in an ultrametric tree all terminal taxa are equidistant from the root of the tree (which can only be rooted). Cluster analyses start by constructing a pairwise-distance matrix of all taxa. The pair with the shortest distance is selected, their distances averaged and the two taxa are combined to a new taxon in the cluster. Then, all pairwise distances are calculated again and the pair with the smallest distance selected and combined. This process is continued until all taxa are included into pairs or larger groups and no single taxon is left.

The disadvantages of cluster analyses and the assumption of ultrametric distances is that no optimality criterion is defined that can be used to distinguish between alternative solutions and if evolution is not uniform across all taxa systematic error can be high. Cluster analyses have been used because of their ease of implementation and

computational speed. Today, more sophisticated, but still fast methods are available that make less strong assumptions than ultrametricity of the data.

The alternative to assuming ultrametric data is to make use of the additivity of a given tree. Under perfect conditions, the observed evolutionary distances and the branch lengths connecting two taxa on a tree are identical. All distance analyses (see Section 1.2.4.3.7) make this assumption. While more sophisticated distance methods also assume an underlying evolutionary model and use an optimality criterion, the *neighbour-joining method* (Saitou & Nei, 1987) is a distance method that only assumes tree additivity. It is conceptually related to cluster analysis but because it does not assume ultrametric data it reconstructs unrooted trees. The neighbour-joining procedure starts with a tree on which all taxa are joined to a single internal node, a so-called star-tree. Whilst cluster analysis uses raw distances between taxa, neighbour-joining adjusts the distance between two taxa according to their average distance to all other taxa, effectively normalising the divergence. Each time a minimum-distance pair is found, its ancestor is formed and the branches leading to the two original taxa are pruned from the tree. The process is complete when only one branch and two nodes remain. As with cluster analysis, neighbour-joining does not evaluate the tree according to an optimality criterion. However, it is a useful first approximation of the data and good starting point for branch swapping procedures under an optimality criterion (Swofford *et al.*, 1996).

1.2.4.3.2 Criterion-based methods

There are three fundamental, criterion-based methods: parsimony, likelihood and distance methods. All criterion-based methods assume an underlying model of character evolution. Under the parsimony criterion this is implicit. The best tree is the one with the shortest possible length measured by the number of character changes. Under the likelihood criterion and some distance-based methods, the model is explicit and the parameters, at least the starting values, are defined by the investigator. If one or more of the model assumptions are violated by the data, either because there is no knowledge of the correct parameter or because the pattern of character evolution in the data cannot be implemented correctly in the program, the resulting hypothesis is likely to be flawed. However, violation of model assumptions does not necessarily mean that the model is invalidated because it can still provide an estimate of the true phylogeny. If the model is inadequate, the data can reject it.

Irrespective of the algorithm, criterion and model, several assumptions are implicit to all methods: it is assumed that the true phylogeny has tree properties and that all characters

are inherited along the genomic lineage; genes that have undergone lateral gene transfer cannot be analysed under the above assumption and other methods have to be used. However, lateral transfer is uncommon in eukaryote nuclear genomes. Thirdly, all methods assume that all characters evolve independently from each other. This assumption is violated for most real data sets because character evolution is often linked by functional constraints (see Section 1.2.3.4). Unfortunately, character dependence is often only known to exist but it is not clear exactly which characters are linked. No computer programs are yet available that can fully implement character dependence. Lastly, changes in base frequencies and substitution rates are assumed to be equal over time and space. As shown below, unequal base frequencies in different taxa can be accounted for by the LogDet algorithm and different substitution rates amongst sites by the gamma distribution. Changes over time are much more difficult to implement and there are no algorithms available to date.

1.2.4.3.3 Parsimony methods

Parsimony (Farris, 1983; Sober, 1983, 1988) methods make the implicit assumption that evolutionary change is rare and therefore the best solution is the one that requires the fewest repeated character changes, i.e. it favours the solution that requires the least *ad hoc* hypotheses of change. In the case of character conflict, parsimony simply chooses the solution that requires the least changes in all characters. The actual amount of evolutionary change might be much higher if superimposed changes have taken place. To find the most accurate estimate of actual evolutionary change methods like maximum likelihood have to be employed (see Sections 1.2.4.3.4 and 1.2.4.3.5).

Parsimony is the only established method for the analysis of morphological characters and the only method accepted by cladists. As a result of its easy implementation, it is also the most widely used criterion-based method for the reconstruction of the evolution of molecular characters.

The 'most parsimonious' solution is that tree that implies the fewest changes across all characters and taxa resulting in the tree with the shortest total branch length. Often, a parsimony analysis recovers many alternative, most parsimonious solutions. There are three alternatives for deciding which solution to favour: making an informed choice based on additional information, such as the stratigraphic record, employing a character weighting scheme (see Section 1.2.2) or constructing a consensus tree (see Section 1.2.7.1) of all most parsimonious trees.

Parsimony discriminates between three different types of characters: characters that have the same state for all taxa (constant characters), characters that are either constant or for which only one taxon has a different state (uninformative characters), and characters that have at least two states for at least two different taxa. The last type can either be informative, i.e. lead to the correct phylogeny, or misinformative, i.e. mislead the analysis to infer the wrong solution (see below).

Parsimony does not account for heterogeneity in the rates of evolution across lineages or sites because it reconstructs the shortest possible tree. As Felsenstein (1978) showed, in the case of rate heterogeneity, parsimony becomes inconsistent, or "positively misleading" (Felsenstein, 1978), i.e. adding more and more characters to the data matrix will only increase the chance of recovering the wrong phylogeny. One suggestion for avoiding entrapment in the "Felsenstein Zone" is to sample as few characters as possible and to hope these are the informative characters that lead to the true tree. Since it is practically impossible to know which characters are misleading and which are informative, there is usually no way of knowing that the data are misleading (Swofford *et al.*, 1996). However, positively misleading data usually lead to taxa that are very divergent relative to each other and all other taxa to be placed on the tree as sister taxa. Hendy & Penny (1989), Steel (1989) and Zharkikh & Li (1993) found that this situation can occur even under equal rates of evolution and the phenomenon was generally called "long-branch attraction" (LBA).

There has been considerable controversy over the exact extent of inconsistency under parsimony. Steel *et al.* (1993a) reported that LBA was not a problem of the parsimony criterion *per se* but a result of its implementation. They showed that all models can be inconsistent when no adjustment is made for superimposed changes but that a non-linear transformation of the data can account for multiple substitutions. Alternatively, the addition of potential sister taxa of the divergent taxa might divide the long branches and therefore avoid inconsistency (Hendy & Penny, 1989; Swofford & Olsen, 1990). Eyre-Walker (1998) found that not only variations in the evolutionary rates but also biases in the base composition may cause parsimony to become inconsistent. If certain types of substitutions are rarer than others but parsimony assumes equal amount of change, it will incorrectly infer an excess of common to rare changes.

Lamboy (1994, 1996) pointed out that if the character (state) to taxon number ratio is very low, there are so many possible solutions that the chance of recovering the correct tree becomes very low. His approach has been criticised by Wiens & Hillis (1996) for the

inadequacy of the comparisons made in the 1994 paper but the inherent problem of low character (state) to taxon ratio remains a problem, especially with morphological data.

1.2.4.3.4 Likelihood

In contrast to parsimony, maximum likelihood methods (Cavalli-Sforza and Edwards, 1967; Felsenstein, 1981a, 1983) reconstruct each possible solution and assess the likelihood that the underlying model and the assumed evolutionary history (the tree) gave rise to the data. The model may be fully defined or its parameters are estimated from the data. The likelihood for each site is the sum of all probabilities of change. The likelihood for the whole data set is the product of all likelihoods over all characters.

Maximum likelihood estimation differs from parsimony in that maximum likelihood attempts to estimate the actual amount of change by taking into account all possible solutions (ancestral states) that are consistent with the tree and the branch lengths under an evolutionary model while parsimony seeks to find one single solution that minimises the amount of change. Under the likelihood criterion the cost of change is a function of branch lengths while under parsimony it only depends on the states of both daughter trees. Maximum likelihood can compensate for superimposed changes while parsimony will underestimate evolutionary change under all but very slow evolutionary rates.

This ability of maximum likelihood to account for branch length variation has an important consequence for the treatment of autapomorphies in the analysis. Parsimony ignores autapomorphies for the topology reconstruction because they are uninformative; their inclusion only affects the terminal branch lengths. On the other hand, likelihood considers that changes are more likely to occur on long branches than on short ones because long branches can account for more character changes including autapomorphies (Swofford *et al.*, 1996). In the case of character conflict, likelihood can use the information from autapomorphies to determine which one of the potentially conflicting characters is the homoplasious one (Wiens, 2003). This makes maximum likelihood consistent in many situations where parsimony is inconsistent.

Maximum likelihood methods also have a lower variance, i.e. they are less affected by sampling error (see Section 1.2.6), and they are relatively robust against model violations. The phylogenetic estimate will still be close to the true tree even if the parameters are assumed to be equal over characters or taxa when in fact they differ slightly.

Maximum likelihood therefore tends to out-perform parsimony and distance methods even for very small data sets. However, the disadvantage of likelihood methods is their computational complexity. Although computer programs can implement most evolutionary models (see Section 1.2.4.3.6), the actual calculation times can be beyond practical means and therefore other methods, such as parsimony and purely algorithmic methods, are still widely used for larger data sets.

1.2.4.3.5 Bayesian inference

Bayesian inference (Li, 1996; Mau, 1996; Rannala & Yang, 1996) has only very recently been introduced as a phylogenetic analysis technique. Although this is a new technique for phylogenetic analysis, the mathematical theory was developed 240 years ago by Thomas Bayes (posthumously published in 1763). Felsenstein (1968) already considered the use of Bayesian analysis in phylogenetic inference in his Ph.D. Thesis, but computer technology was not advanced enough at the time to provide practical implementations. Only recent advances in computing power have made it possible to put his ideas into practise (e.g. programs BAMBE, Simon & Larget, 1998; MrBayes, Huelsenbeck & Ronquist, 2001).

Bayesian inference is tightly linked to likelihood analysis as it uses likelihoods to assess the quality of the tree under consideration. The main difference between the two methods is that likelihood analysis calculates the probability of the data given the tree, while Bayesian analysis calculates posterior probabilities of the tree given the data. The posterior probability is calculated using the likelihood function multiplied by a prior, which is the *a priori* estimate of the parameters. Because Bayesian analysis makes use of likelihoods, the same underlying evolutionary models as in likelihood analysis can be used and the calculation inherits the statistical properties of maximum likelihood. Whilst the model parameters (e.g. base frequencies, gamma shape parameter α) have to be fixed for each individual likelihood analysis, they are usually left free in a Bayesian analysis and are approximated during the run across the likelihood surface.

The Markov Chain Monte Carlo (MCMC) approximation (Metropolis *et al.*, 1953; Hastings, 1970; Green, 1995) is implemented into Bayesian inference because the full implementation of pure Bayesian inference is too complicated for practical implementation. The MCMC allows Bayesian analyses to be run on a modern high-power desktop computer. During the MCMC the search for optimal trees wanders through the likelihood surface of tree space similar to a pure likelihood search. The difference lies in the fact, that the chain is given the chance, under certain conditions, to "cross valleys" and

find alternative "hilltops". Each time, the chain has found a new state it will be accepted without question if the likelihood is higher than the current state. However, if the likelihood is worse, a random number between 0 and 1 is chosen and if this number is smaller than the likelihood ratio of the current and new state, the new state will be accepted despite being worse. The advantage is that accepting a worse state temporarily will give the chain the chance of finding higher "hilltops" that are out of reach under a pure likelihood search.

There is no worry that the chain might be trapped in a valley for the rest of the search because the chain will always accept a better likelihood; but the ability to accept temporarily worse states increases the chance of finding the global optimum of the likelihood surface. Additionally, a sample of all the trees found during the search will be saved. The ability to retain states found throughout the search provides the great advantage of Bayesian inference over the pure likelihood criterion. The frequency with which any particular tree is sampled during the MCMC almost exactly corresponds to the frequency that the tree is found in the full likelihood distribution. The posterior probability of that tree therefore corresponds to the probability of that tree being the best tree amongst all possible trees. The sum of the posterior probabilities of all trees is one. Furthermore, the posterior probability of any given clade (the sum of the posterior probabilities of all trees containing that particular clade) can be interpreted as support for its presence in the best tree. The ability to calculate posterior probabilities that correspond to the confidence in a particular node eliminates the need to perform an additional analysis to calculate nodal support, such as the bootstrap. Therefore, Bayesian analysis provides a great time saver compared to performing a likelihood analysis plus a bootstrap. Bayesian inferences can be implemented even for large data sets for which robust likelihood searches are practically impossible.

Each MCMC will run for several thousands or millions of generations and the number of retained trees can be rather large. The results from a MCMC can be summarised by the maximum *a posteriori* probability (MAP) estimate of the phylogeny, the 95% credibility set or the majority rule consensus tree. The MAP is the single, most probable tree found by the search (Rannala & Yang, 1996). The 95% credibility interval contains the MAP tree plus all trees with decreasing probabilities to obtain a cumulative probability of 95% of all trees (Felsenstein, 1968). The majority rule consensus tree is the same as the one commonly used to summarise a bootstrap analysis but the support numbers obtained correspond to the posterior probabilities of the nodes (Larget & Simon, 1999).

After the initial euphoria about the implementation of Bayesian inference in phylogenetic analysis, several problems have been encountered. For some of these, solutions have been found, whilst others have yet to be fully understood. Huelsenbeck *et al.* (2002) gave a detailed account of potential pitfalls and how to handle such situations. In theory, the choice of priors of all parameters can have an influence on the posteriors of the analysis. However, in practice, if the chain was run long enough and the data sets are large enough, the initial settings of the priors have very little influence on the posteriors of the analysis. Stating the choice of priors even if they were kept as the default values (e.g. 1) of the program remains good practice.

The second, and probably most difficult problem is to make sure that the chain has found a stable likelihood and therefore that the sample of trees retained from the analysis truly corresponds to a sample of all possible trees. This is a problem of convergence and mixing. A quick and dirty test is to plot the distribution of likelihood per generation or per sample taken for the whole run of the chain. Initially, the likelihood should rise as the chain continues to find better and better trees. After a while, the expectation is that the chain converges onto a set of trees within a range of stable likelihood values. All trees retained during the initial period can then be discarded from any further analysis; this is called the 'burn-in period'. However, plotting the log-likelihood values is notoriously unreliable (Gilks *et al.*, 1996). The chain might seem to have converged but then suddenly, thousands of generations later, finds a significantly better tree and the likelihood is raised again to a new plateau.

Alternatively, the behaviour of all parameters in the model can be monitored, but some of them might not be expected to converge on any particular value but simply vary throughout the analysis (e.g. the proportion of invariant sites, Peter Foster, pers. comm.). Another option is to divide the sample trees into sections across time of the run and to construct the majority rule consensus tree of each of these subsets of tree samples. If the posterior probability of individual clades is relatively similar it indicates that the chain converged on a stable likelihood.

A good practice is also to run several independent chains and to compare the likelihoods and parameter values of the different chains. If they converge on the values there is clear indication that the chains converged and mixing was appropriate. It is also possible to compare the topological variation among and between chains by using tree-to-tree distances (Huelsenbeck *et al.*, 2001). Before convergence, the tree-to-tree distance within a chain is expected to be smaller than the distances between independent chains.

After convergence the tree-to tree distances within and between chains should be approximately the same. The posterior clade probability can also be compared between the independent chains and the precision can be given as the observed Monte Carlo standard variation around the mean.

A further improvement of the mixing can be made using the Metropolis-coupled MCMC (MCMCMC, Geyer, 1991). In this case, several chains are run at the same time, one of which is the 'cold chain' and the others are 'heated'. The heating of a chain improves its ability to "jump valleys" to "higher hilltops". At regular intervals, one of the heated chains changes states with the cold chain, enabling it to find new, better solutions. Only trees found by the cold chain are sampled, the function of the heated chains is purely to assist the cold chain to "jump to new hilltops" more easily. The best practical solution to ensure appropriate convergence and mixing is to run several long and independent chains using the MCMCMC and to monitor all parameters throughout the analysis.

Another common observation is that the posterior probabilities of a Bayesian analysis are generally higher than the corresponding bootstrap proportions of a likelihood or parsimony analysis. To date, there is no clear-cut explanation of this phenomenon but Huelsenbeck *et al.* (2002) suggested that it might be because of the inherent statistical bias of the nonparametric bootstrap (see Section 1.2.6.2). Alternatively, the two methods actually measure slightly different things because the bootstrap is a measure of uncertainty based on the resampling of the original data and has no relation to the evolutionary model in itself while Bayesian inference measures uncertainty based on the specific evolutionary model and the data.

The last unsolved problem has to do with model selection. There is so far no clear method of determining which model is the best and how well this model fits the data. Bayesian methods such as the Bayes factor or posterior predictive P values have been suggested but are not yet implemented in computer programs. The best practical solution is to use a likelihood ratio test and to stick to that model which gives the apparently best fit for the data.

Despite these problems with the practical implementations of Bayesian analysis, it is a good alternative to existing methods, especially if large data sets are under investigation. Other methods have their own flaws and as long as none of the alternatives are perfect, comparing results from analyses under different criteria can only provide more information about robust and weak relationships in the phylogeny under reconstruction. The intuitive appeal of Bayesian analysis lies in the clear-cut interpretation of the posterior probabilities

compared to, for example, bootstrap proportions (Gelman *et al.*, 1995). Posterior probabilities of a clade correspond to the probability that the clade is true given the model, the priors and the data.

1.2.4.3.6 Evolutionary models

All evolutionary models used in current likelihood programs are based on homogenous Markov models. All substitutions are fully reversible, i.e. it is assumed that it is just as likely that A changes into T as T into A. The model does not take into account any previous states of the character. The model assumes that each character has a constant rate of evolution, i.e. it does not account for different substitution rates in different parts of the tree, although there are ways of implementing different substitution rates across sites (see below).

All models account for changes in base frequency and substitution rates. They are expressed as a 4 x 4 instantaneous rate matrix Q. Each element of the matrix is the product of the instantaneous substitution rate factor μ, a base frequency parameter and the relative rate parameter for the substitution of one base into another. The elements of the diagonal (changes from one base to itself, i.e. unobservable changes) are chosen so that each row sums to zero.

Most current models are time-reversible, i.e. the relative rate parameter for changing one base into another is the same as the reverse change. It also means that the trees inferred are unrooted. The general time-reversible (GTR) model allows for four different base frequencies and six different substitution rates. All other models are special cases of the GTR model and can be nested into each other and into the GTR model. The other models are simplified by assuming equal base frequencies and/or reducing the number of substitution rate classes to two (transitions and transversions only) or three (one for transitions and two for transversions).

In all these models, branch length is the product of time passed between two evolutionary events and the number of substitutions that have occurred during this time interval, $t\mu$. This means, that a long branch can either express a long time difference between evolutionary events or many substitutions per unit time. These two components can only be separated properly if substitution rates are equal across lineages, i.e. under the assumptions of a perfect molecular clock. The branch lengths become estimates of substitution rates and likelihoods can be estimated from branching times rather than branch lengths (Bishop & Friday, 1985; Felsenstein, 1993). Although the molecular clock

assumption requires the use of rooted trees, there are still fewer parameters to be estimated making analysis more efficient. If no molecular clock can be assumed, the mean substitution rate μ is set to 1 and the branch lengths then represent the expected number of substitutions per site on a branch per arbitrary unit of time.

Faced with a variety of models it is not always obvious which one is the best to use. The model has to fit the data and should also fit to other reliable sources, i.e. show congruence with other data sets. Commonly, likelihood ratio tests (e.g. Sokal & Rohlf, 1981) are used to estimate the fit of a model to the data relative to another model. The likelihood values obtained under each model can be compared by goodness of fit tests (such as the χ^2 test) as long as the models are nested, i.e. one model is a special case of the other one. The most appropriate model is the one that has a significantly better likelihood value than the next simplest model. One could always choose the GTR model because it would seem to make the best use of all assumptions. However, as more parameters have to be estimated from the data the variance becomes larger. Therefore, if the addition of more parameters does not significantly improve the fit of the model to the data the simpler model should be chosen in favour of a lower variance.

All models based on the GTR model assume equal rates over sites and lineages. If substitution rates vary, likelihood inferences can also become positively misleading (see Section 1.2.4.3.3). Accounting for differential rates at different sites can prevent a likelihood or distance analysis under an explicit evolutionary model from becoming inconsistent. In a perfect world, the substitution rates for each character would be known and could be assigned accordingly. Since this is usually not the case, approximations to the rate heterogeneity over the characters have to be made.

The simplest model for accounting for different rates is the invariable-site model (Hasegawa *et al.*, 1985; Churchill *et al.*, 1992; Reeves, 1992; Sidow *et al.*, 1992). It assumes that a proportion of sites do not vary at all. The assumption that sites either have a constant substitution rate or do not vary at all is very unlikely. The alternative is to assign a continuous distribution for the rate heterogeneity over the data. The most common model is the gamma (Γ) model (Steel *et al.*, 1993b; Yang, 1993). It has two parameters, the shape parameter α, and the scale parameter β. This continuous model distribution is computationally very expensive. Yang (1994a) reduced the continuous model to the "discrete Γ model" in which the Γ distribution is divided into several rate categories of equal probability. He found that the discrete model gives a good approximation to the rate variation. The use of the discrete model saves a lot of computational time because only the

shape parameter α is added to the model independently of the number of rate categories (β is set to $1/\alpha$ for distributions of a mean rate of 1). A useful practical implementation is the combination of the invariant sites model and the discrete Γ distribution (Gu *et al.*, 1995; Waddell & Penny, 1996) in which a proportion of sites are assumed to be constant and the remaining sites are assumed to vary according to a set number of different rates.

It is one thing to choose the appropriate model and another to choose its parameters. As already mentioned, the parameters have to be estimated from the data. Simultaneous estimation of all model parameters, branch lengths and tree topologies is generally impossible because of the computational complexity. The most commonly used method of parameter estimation is by successive approximation. Initial parameter estimates are either calculated from the data itself (empirical) or a 'best guess' value is chosen arbitrarily, e.g. the base frequencies are chosen as those observed in the data or they are set as equal (0.25 each). A tree is constructed and branch swapping is used to find the best tree(s) under the initial parameter settings. New model parameters are estimated from the best tree(s) and the tree search starts over again under the new model settings. This process is repeated until the model parameters converge on a constant value.

1.2.4.3.7 Pairwise distance methods

All distance methods are based on the idea that evolutionary distance is a tree metric, that it defines a tree. If all evolutionary distances between all taxa in a given data set were known, it would be easy to reconstruct their relationships as a tree. However, the true evolutionary distances are usually not known, primarily because of superimposed changes. Pairwise distance methods aim to adjust the observed distances into corrected distances which act as approximations to the true evolutionary distances. Pairwise distance methods can make use of underlying evolutionary models to estimate evolutionary change and account for superimposed changes similar to those used in likelihood estimates. The difference between distance and likelihood approaches is that distance methods do not use the actual, discrete nucleotide data but convert it into similarities between taxa.

The drawback of using distances instead of discrete characters is the loss of information (Penny, 1982). Distance analysis does not allow the investigator to go back to the original data set and deduce which characters are informative or misleading. Another disadvantage is that distance methods cannot be used to combine different data sets in the same analysis (Miyamoto, 1985). These conceptual disadvantages and simulation studies have led to the opinion that distance methods are less powerful and usually outperformed by direct likelihood estimations under the assumptions of the same model (Kuhner &

Felsenstein, 1994; Yang, 1994b; Huelsenbeck, 1995). However, in practice, distance and likelihood estimates can perform equally well. For large data sets the use of distance methods are often of great advantage because they reduce computing time.

Pairwise distance methods are based on additive distances, that is, under the absence of stochastic (random) error and a perfect evolutionary model, the branch lengths connecting two taxa would be exactly equal to the evolutionary distance of these taxa. In reality, random error distorts the additivity of the tree even if the underlying model perfectly represents the actual evolutionary processes.

Pairwise distance methods evaluate trees according to an optimality criterion: (weighted) least-squares-fit or minimum evolution. The *least-squares-fit criterion* (Cavalli-Sforza & Edwards, 1967; Fitch & Margoliash, 1967) compares the observed distance between two taxa to its distance on a reconstructed topology. If none of the assumptions are violated, the distances are identical and the tree is truly additive. In the event of multiple substitutions however, some tree distances will be larger than those that have been observed. The best solution is the one for which the least-square-fit of all differences is optimal according to a goodness of fit statistic.

In the *minimum evolution method* (Kidd & Sgaramella-Zonta, 1971; Rzhetsky & Nei, 1992) branch lengths are fitted to the tree topology by unweighted least-squares similar to the above method. Instead of evaluating the branch lengths by a goodness of fit measure, all branch lengths are added to give the total tree length. The optimal solution is the one for which the total length of the tree is minimised. The minimum evolution criterion is therefore similar to parsimony with the difference that the tree length is calculated from the pairwise distances rather than the cost of fitting the individual nucleotides to the tree. The minimum evolution criterion seems to perform slightly better than the least-squares-fit criterion, but the exact reasons are unclear. With increasing sequence length and rejection of trees containing negative branch lengths the differences between the two methods decrease (Kidd *et al.*, 1974; Kuhner & Felsenstein, 1994).

Gaps in the alignment have to be treated with care in distance estimations because they are signs of indels (see Section 1.2.3.1) and not simple substitutions. Positions that contain gaps for one or more taxa can either be excluded entirely prior to the analysis or distances for pairs in which one or both sequences have a gap are omitted. The former option is usually used for long stretches containing many gaps, whilst the latter option is appropriate for single gaps in the alignment. If positional homology in regions of an

alignment containing many gaps cannot be ensured these regions should be excluded entirely from the analysis as should terminal end variations.

Observed dissimilarities between sequences initially increase rapidly as one base changes to another base. Over time, it becomes more likely that a position that has already seen one change will undergo a further substitution. This second substitution at the same position cannot increase dissimilarity any further, thus dissimilarity increases much slower with increasing evolutionary time. Uncorrected dissimilarity measures (such as the Hamming distance or p-distances [Kumar *et al.*, 1993]) cannot account for these superimposed changes but the implementation of evolutionary models can take care of this. The same evolutionary models as those that apply to maximum likelihood methods can be implemented in distance estimations. Instead of using a rate matrix, unequal base frequencies and substitution rates can be incorporated in the distance equation. Among-site rate variations can also be accommodated in the distance equation.

The most effective way of implementing the above models is by actually estimating maximum likelihood distances. In this method, the tree is treated as a single branch and the branch length that maximises the probability of sequence evolution is estimated. Here, the branch lengths equal the expected number of substitutions per site. The model parameters have to be estimated from the data in much the same way as in pure likelihood analysis. The advantage of using maximum likelihood distance estimates to full maximum likelihood analysis is again one of computational speed.

The *LogDet* (Lockhard *et al.*, 1994; Steel *et al.*, 1994) or *paralinear distance* (Lake, 1994) is a transformation of distances that can account for biases in the base frequencies amongst different lineages as long as sites evolve independently and among-site substitution rates are equal. Homogeneity of base composition can be tested using χ^2 statistics (Rzhetsky & Nei, 1995). The LogDet or paralinear model suffers when among-site rate variation is high. Under such circumstances it might be appropriate to include an estimate for invariant sites. Simply excluding all constant or parsimoniously-uninformative sites tends to be a very severe tactic (Waddell, 1995).

As with maximum likelihood models, the choice of the best model for the analysis of a given data set is fundamentally difficult. Accounting for all evolutionary processes and choosing the most general model (GTR+I+Γ or LogDet) could lead to overly high variance in the distance estimates. There is no one strategy of testing which model to use, but if two estimates give similar values, the simpler model should be used because it will have a lower variance (Kumar *et al.*, 1993).

1.2.5 Rooting and character polarity

One aim of phylogenetic reconstruction is to assess the direction of evolutionary change of alternative character states. For this to be achieved, the ingroup network has to be rooted. Rooting enables the investigator to make an informed decision about the polarity of the character state change, i.e. if a state is plesiomorphic (ancient with respect to the ingroup) or apomorphic (derived with respect to the most recent common ancestor of the ingroup). Furthermore, synapomorphic (shared and derived) character states provide evidence of a common origin of a group of taxa (Hennig, 1966).

The possible rooting methods for morphological data are outgroup comparison (Watrous & Wheeler, 1981; Farris, 1982; Maddison *et al.*, 1984; Nixon & Carpenter, 1993), the ontogenetic method (Mayr, 1969; Nelson, 1978, 1985; Patterson, 1982, 1983; Brooks & Wiley, 1985; de Queiroz, 1985; Kluge, 1985; Bryant, 1992) and the palaeontological method (Nelson, 1978; Patterson, 1982; de Queiroz, 1985; Eldredge & Novacek, 1985; Gauthier *et al.*, 1988; Donoghue *et al.*, 1989; Bryant, 1992), whilst for molecular sequencing data the only rooting method is outgroup comparison (Arnold, 1981; Smith, 1994). Only rooting via outgroup analysis will be discussed in the following section because no morphological characters were collected for the outgroups in this study.

Ingroup and outgroup taxa should be analysed simultaneously and the resulting cladogram rooted between the ingroup and outgroup (Farris, 1970, 1972; Nixon & Carpenter, 1993; Kitching *et al.*, 1998) rather than constraining the outgroup to monophyly prior to analysis (Maddison *et al.*, 1984). Unconstrained outgroup analysis allows testing for ingroup monophyly as well as assessing the ingroup polarity after ingroup ancestral states have been assigned. Therefore, character polarity does not have to be defined *a priori* but can be read off the resulting cladogram (Farris, 1970, 1972; Nixon & Carpenter, 1993; Kitching *et al.*, 1998). In fact, Nixon & Carpenter (1993) showed that the use of a single outgroup, rooted in the correct position, will achieve polarisation.

It is still controversial as to which, and how many, outgroup taxa are necessary for a robust outgroup rooting. Wiley (1981), Donoghue (1983), Donoghue & Cantino (1984), Stuessy (1990) and Mayden & Wiley (1992) insisted that the nearest sister taxon to the ingroup is necessary for outgroup rooting. Watrous & Wheeler (1981), Donoghue & Cantino (1984) and Smith (1994) suggested that outgroup taxa should be monophyletic with the ingroup in a wider phylogenetic aspect and that one should use a sister group taxon if at all possible. Furthermore, Donoghue & Cantino (1984) also recommended that only those ingroup nodes should be accepted that are robust to outgroup choice. Nixon &

Carpenter (1993) showed that the outgroup does not necessarily have to contain the sister group but that any outgroup can be used as long as it does not introduce homoplasies that cannot be dealt with. More recently, Lyons-Weiler *et al.* (1998) tested several outgroup approaches and suggested using two outgroup taxa, one of which should be the sister group to the ingroup and the other a more distantly related outgroup. Their argument was that the phylogenetic signal decreases if more than one representative from one outgroup taxon is included, and that the sister taxon would have a reduced plesiomorphic content relative to a more distantly related taxon if the sister taxon exhibits a high rate of nucleotide changes.

No matter what type of data is used, outgroup choice has to be carried out with great care. If the outgroups are too distantly related, several problems will occur. The more distantly related one molecular sequence is from another, the higher the probability that the observed nucleotide changes represent only random data as a result of multiple substitutions at the same homologous sites (Wheeler, 1990; Smith, 1994).

It is generally advisable to analyse the ingroup first and then to add outgroups to the resulting network (Lundberg, 1972; Nixon & Carpenter, 1993). For cases in which the outgroup relationships are poorly known, Donoghue & Cantino (1984) suggested that all plausible outgroups should be investigated, alone and in various combinations. They called this method the "outgroup substitution approach" and added that this method would be most useful for establishing an outgroup hypothesis prior to a full cladistic analysis.

There are three alternatives to outgroup comparison for cases where no outgroups are available: the use of a hypothetical ancestor, Lundberg rooting (Lundberg, 1972) or midpoint rooting (Farris, 1972). However, there are problems with all three methods. In the first two cases, there are high chances that the globally optimal root is not found, and midpoint rooting assumes a molecular clock. Therefore, outgroup comparison should be used whenever possible.

A major problem with rooting by outgroup comparison is the problem of long-branch-attraction (Felsenstein, 1978; Hendy & Penny, 1989; Wheeler, 1990; Tarrio *et al.*, 2000). If the outgroups are distantly related and therefore contain many homoplasies or if they have different evolutionary rates and hence contain only random data, the outgroups tend to root on the longest internal branch through random association rather than true phylogenetic signal (Wheeler, 1990; Smith *et al.*, 1992, Smith, 1994; Tarrio *et al.*, 2000). Wheeler (1990) suggested that if the outgroup(s) root on the longest internal branch one should avoid outgroup rooting altogether and instead use the asymmetric character transformation matrix for rooting. Smith (1994) suggested the use of several sister taxa as

outgroups because it increases the stemminess of the tree: The use of one taxon from each possible outgroup leads to an increase of long branches and random association of homoplasious data, while using several sequences from the same sister group will break up long branches and give balance to the tree. Maddison *et al.* (1984) had shown for morphological data that the use of a single outgroup was more likely to identify the wrong root because of the high potential for many homoplasies.

Tarrio *et al.* (2000) incorporated a model containing assumptions that deal with rate heterogeneity. They ran simulations under a variety of criteria and models using midpoint rooting instead of outgroup comparison. They found that more complicated but apparently biologically more accurate models under the maximum likelihood criterion often failed to find the true tree, but simple models were able to reconstruct the correct phylogeny. These findings were attributed to the presence of base compositional bias amongst the sequences. Accounting for rate variation reduces the phylogenetic signal of the GC content under base compositional bias (hence, long-branch attraction not only influences parsimony but also likelihood inference). The only advanced model that could cope with the base compositional bias was the LogDet method because it accounts for base compositional bias. If both, heterogeneity and base compositional bias, are present in the data, the investigator has to find a balance between accounting for one or the other, as there are no implementations to date to account for both at the same time.

Even if evolutionary rates are homologous and base composition is unbiased, other details of the cladogram, such as topology and convexity (i.e. the placement of the root), still vary with the choice and input order of the outgroups in the data matrix (Milinkovich *et al.*, 1996; Barriel & Tassy, 1998; Milinkovich & Lyons-Weiler, 1998). The "prime outgroup", that is the first outgroup in the matrix, has the greatest effect on the ingroup topology. It is therefore advisable to permute the outgroup order and then draw a consensus tree of all hypotheses. Alternatively, the input order of all taxa in the data matrix can be permuted in several independent analyses avoiding any effect of a "prime outgroup" (see Section 1.2.4.2).

Although the problem of alignment is dealt with elsewhere, a note should be made about the treatment of character selection when adding outgroups (Smith, 1994). After potential exclusion of ambiguous regions in the ingroup alignment, the outgroup alignment needs to be scrutinised. In a parsimony analysis, all ambiguous positions within the outgroup alignment should be replaced with '?' for missing data, while in a likelihood analysis these positions should be entirely excluded from the analysis. Furthermore, under

the parsimony criterion, it is possible to check the variance of individual sites within outgroups and only include those sites that are stable within the outgroups or weigh variant positions according to the consistency or retention index (see Section 1.2.6.1) of these sites after a first run through the data. In conclusion, Smith (1994) found that maximum likelihood had a higher degree of success of retrieving the correct ingroup topology when including ambiguous positions than maximum parsimony, however, both methods performed better if ambiguous characters where excluded. There was no improvement in the accuracy of several distance estimates when ambiguous characters were excluded; in fact, all distance methods tested performed quite poorly overall.

1.2.6 Support for and confidence in the data

As seen above, there is hardly any situation where the inferred phylogeny is the true reconstruction of evolutionary history. Incomplete sampling, homoplasy and the need to use heuristic approaches introduces errors to phylogenetic reconstruction. There are two types of error: Random (or sampling) error (Type I error) is the deviation of the observed parameter from the true parameter because of incomplete sampling. It will completely disappear if all data are collected but this is impossible because not every step of an evolutionary history can be sampled. Random error can be reduced by adding more observations, providing there is no bias in the data (Hillis *et al.*, 1994, Swofford *et al.*, 1996).

Systematic (or statistical) error (Type II error) is the deviation of the observed from the true value because incorrect assumptions are made about the statistical model. This type of error only increases with additional data. The most common example of systematic error is the phenomenon of long-branch attraction (see Section 1.2.4.3.3). Only using a better model for the analysis can reduce systematic error. Knowing which model is the best fit to the data and the hypothesis is difficult and has been dealt with in Section 1.2.4.3.6.

Increasing the complexity of the model to account for systematic error increases the amount of random error because more parameters have to be estimated from the data; it also increases the time requirements and might make the best combination computationally infeasible. Computational time also increases with the amount of data that is added to reduce random error. There is a trade-off between accounting for both types of error in a given data set. To find the best solution to the problem, different models can be tested and associated errors estimated. The best model is the one for which phylogenetic hypothesis does not change significantly by adding more data or free parameters. For likelihood

models, the likelihood ratio tests can be applied (Goldman, 1993a, 1993b; Rzhetsky & Nei, 1995; Posada & Crandall, 1998 [MODELTEST program]).

Many different methods have been suggested to test the support of the data for the resulting phylogeny. Only two methods, the consistency and retention index and the bootstrap, are applied in the present study and they are discussed further in this section. Alternatives methods include: whole randomisation tests for phylogenetic signal (Faith, 1991; Faith & Cranston, 1991; Goloboff, 1991; Hillis, 1991; Huelsenbeck, 1991), the Bremer support index for branch support and its variations (Farris *et al.*, 1982; Bremer, 1988; Källersjö *et al.*, 1992; Davis, 1993; Bremer, 1994) and the jackknife (Lanyon, 1985; Siddall, 1996).

1.2.6.1 Consistency and retention indices

One of the earliest methods for estimating random error is to test the support of the data for the cladogram, i.e. how much of the data are informative and how much homoplasy is contained in it. The fit of a character on a particular cladogram can be measured either as total branch length or by indices such as the consistency index (Kluge & Farris, 1969; Farris, 1989) and retention index (Farris, 1989). Total branch length is simply the total amount of character change over the whole cladogram but in the presence of homoplasy it is a biased estimate.

The *consistency index* (CI) is the ratio between the minimum amount of change that a character may show on a tree to explain the data (m) and the amount of change in the character on the actual cladogram (s). It is essentially a measure of homoplasy. High values of the CI indicate a low level of homoplasy; low values indicate high amounts of homoplasy. If no homoplasies exist the consistency index has a value of 1 (and m and s also equal 1). The consistency index can be calculated for each individual character or the values of m and s for all characters are added up to calculate the fit of all characters on the cladogram giving the ensemble consistency index. Since the consistency index of individual characters is dependent on the topology, they can have different values for all equally parsimonious solutions of the same data set; the ensemble consistency index for all characters will be identical for all most parsimonious trees (Carpenter, 1988). The *homoplasy index* (HI) is the complement of the consistency index and equals one if all similarities result from homoplasy and zero in the absence of homoplasy.

Inclusion of uninformative characters (apomorphies and invariant characters) in the analysis overestimates the CI. This inflation is particularly crucial when different data sets

are compared or if character weighting is employed as well (Mickevich & Farris, 1981; see also section 1.2.2). Therefore, uninformative characters should be excluded, especially if a weighting scheme is employed (Colless, 1983; Carpenter, 1988; Farris, 1989). The CI is inversely correlated to the number of taxa in the data set but independent of the number of characters and the distribution of the character states among the taxa. Effectively, the addition of taxa reduces the homoplasy of the data set (Kitching *et al.*, 1998). In order to account for these correlations, Farris (1989) introduced the retention index.

The *retention index* (RI) (Farris, 1989) is a measure of character similarity that can be interpreted as synapomorphy. The ensemble retention index (including all characters) places greater influence on those characters that support larger monophyletic groups and down-weights those that occur more distantly on the cladogram (Kitching *et al.*, 1998). The RI also equals one in the absence of any homoplasy and it actually takes on a value of zero if all similarities among the terminal taxa result from homoplasies. Neither the number of characters in the data set nor the type of characters influences the RI. The *rescaled consistency index* (RC) (Farris, 1989) is the product of the consistency and retention indices (*CI*RI*). It reaches a value of zero if all similarities in the characters are attributed to homoplasy and takes on a value of one in the absence of homoplasy.

1.2.6.2 The Bootstrap

The non-parametric bootstrap (Efron, 1979, 1982, 1987; Efron & Gong, 1983) was first introduced in phylogenetic analysis by Felsenstein (1985a) as a measure of confidence in the phylogeny or a monophyletic clade therein by repeated sampling of characters from the underlying pool of characters. The data are permutated by randomly sampling data columns irrespective of their original position in the data matrix until the original number of data points is reached; this is called the 'sampling with replacement' method. Because columns are chosen at random, one column can occur more than once in one pseudosample while another column may not get sampled at all for a given replicate. This resampling is essentially a weighting method, where omitted characters are weighted zero and all others have a weight according to the number of times they are sampled (Bremer, 1994). The number of times a given internal branch is found amongst the pseudosamples is counted. This is the bootstrap proportion for this particular branch. Generally, a high number of replicates (>100, better several thousand) are need for a good estimate.

There are four assumptions that have to be met for the bootstrap proportions to represent confidence limits correctly and without bias: (1) the underlying algorithm for the reconstruction of the phylogenetic hypothesis has to be statistically consistent, (2), all

characters in the data matrix have to be independent from each other, (3) the bootstrap procedure needs to be unbiased with respect to the underlying algorithm, and (4) the sample of the data has to be representative of the underlying universe of the character data and more importantly, the universe has to exist in the first place (Sanderson, 1989).

However, when using real data, all of the above assumptions are most likely to be violated. In many cases, the phylogenetic estimation method is statistically inconsistent (see Section 1.2.4.3.3), the characters are not independent of each other (e.g. secondary structure constraints of ribosomal RNAs), bias is inherent in the sample (Hillis & Bull, 1993), and there is no way of proving that there is an underlying universe of the character distribution that is represented by the data at hand. Character correlation causes an overestimate of the confidence levels for the internal branches (Felsenstein, 1985b; Bermingham & Avise, 1986) and inherent bias in the pseudosamples leads to larger variance of the bootstrap proportions than that of independent sampling of data (Zharkikh & Li, 1992a; Hillis & Bull, 1993). Furthermore, Carpenter (1996) showed that the inclusion of uninformative characters can have a substantial effect on the bootstrap results, a consequence of the limit of available data that had previously been disputed (Harshman, 1994).

Nevertheless, even if not all assumptions are met, the bootstrap is still "valuable because it provides a systematic method of assessing the robustness of a data set to perturbation" (Sanderson, 1989). For example, the bootstrap can reveal the confidence limit of an apparently monophyletic group when only a few more characters support the monophyly of the group than the number of characters that do not support the monophyly. However, it is important to note that even if the above assumptions apply, bootstrap proportions cannot be translated into probabilities because congruence could also result from systematic errors that would override the phylogenetic signal (Milinkovich & Lyons-Weiler, 1998).

There are three possible interpretations for bootstrap proportions with respect to the underlying phylogenetic hypothesis that is tested: (1) *precision*, "the degree to which bootstrap proportions based on a finite set of pseudosamples are expected to match the values that would be obtained from an infinite set of pseudosamples", (2) *accuracy*, "the probability that a specific group is contained in the true phylogeny", and (3) *repeatability*, "the probability that a specific group will be found in an analysis of an independent sample of characters" (all quotations from Hillis & Bull, 1993).

There has been considerable controversy over the exact meaning of bootstrap proportions. Felsenstein (1985a) and Sanderson (1989) went with the interpretation as a measure of repeatability, while Felsenstein & Kishino (1993) saw it as a measure of accuracy and interpreted the probability of a type I error, that is of falsely favouring the null hypothesis that the group is not there, as $1 - P$, where P is the proportion of pseudosamples. Hillis & Bull (1993) compared the two interpretations using simulated and real data, and showed that the bootstrap proportions at best "provide an indication only of the degree of support of a particular technique for a particular clade". They concluded that with an infinitive number of iterations, one can get a perfectly precise but highly inaccurate estimate of repeatability and that the bootstrap proportions provide a biased estimate of accuracy (also Zharkikh & Li, 1992a, 1992b).

Bootstrap estimation has the same problems as phylogenetic inference in that if evolutionary rate varies greatly across taxa, estimates can be "positively misleading" (Felsenstein, 1978). For high rates of change with ratios of short: long branches of less than 1:5, bootstrap proportions of more than 50% are consistent conservative measures; for higher ratios parsimony becomes positively misleading (Felsenstein, 1978), but bootstrap proportions are usually lower as well (<85%) (Hillis & Bull, 1993).

With increasing sample size (Felsenstein & Kishino, 1993), higher character to taxon ratio (Felsenstein, 1985b; Sokal & Shao, 1985; Bremer *et al.*, 1999) and a higher number of informative characters (Hillis & Bull, 1993; Bremer *et al.*, 1999) the bootstrap proportions and therefore the confidence in the recovery of the true distribution increases. However, increasing the number of taxa but leaving the number of characters constant decreases the bootstrap support (Bremer *et al.*, 1999), unless the number of characters was already very high (high character to taxon ratio) (Graybeal, 1998). The level of support is also dependent on the taxon sampling itself rather than the absolute character to taxon ratio (Bremer *et al.*, 1999). Cummings *et al.* (1995, for whole genome analysis) and Poe & Swofford (1999) found that it was usually better to add characters, even if long-branch attraction was a problem, but that adding slow-evolving taxa to break up long branches reduced accuracy (Smith, 1994). The increase or decrease of the bootstrap proportions is also dependent on the addition sequence of the taxa (Hillis, 1998; Kim, 1998). The extent of the bias further depends on the location of the internal branch (Hillis & Bull, 1993; Zharkikh & Li, 1995).

Whether high bootstrap proportions over- or underestimate the probability that a group is correct depends on whether the data are highly informative or not. Felsenstein &

Kishino (1993) showed that if the data are resolved, the bootstrap proportions are usually high and underestimate the true probability. They also concluded that confidence intervals are not biased and that the probability that the confidence interval includes the true mean is 0.95 for a normal distribution and slightly larger under the bootstrap unless the data set is quite large. However, Zharkikh & Li (1995) estimated that under the 95% critical value criterion, the bootstrap proportions underestimate the true confidence level and that this underestimation increases with the number of alternative topologies.

To get 95% bootstrap support, a clade has to be supported by at least three synapomorphies in the absence of any homoplasies, with even distribution of character change on the tree and with binary characters (Felsenstein, 1985b; Bremer *et al.*, 1999). In the presence of homoplasy and with multistate characters the number of synapomorphies must be greater. Hillis & Bull (1993) concluded that bootstrap proportions (BP) above 85% indicate strong support for a clade and BP between 65 – 85% moderate support. BP below 65% indicate insignificant support and therefore monophyly of clades with such low support is doubtful.

1.2.7 Taxonomic versus character congruence – consensus versus combined analysis

Accuracy of phylogenetic reconstruction can be estimated as the confidence in the resulting hypothesis or the amount of phylogenetic signal contained in the data. However, accuracy can also be measured by congruence amongst data sets derived from different suites of characters because they produce different sets of hypotheses (Penny & Hendy, 1986; Swofford, 1991). Congruence between the fundamental cladograms from the different analyses can be assessed with three different methods: consensus analysis (taxon congruence), combined analysis (character congruence or total evidence) and character weighting. All three methods have their advantages and disadvantages and are used in slightly different situations. Character weighting has already been dealt with in Section 1.2.2 above.

1.2.7.1 Taxonomic congruence: partitioned analysis and consensus trees

Phylogenetic analysis of one or more data sets can result in fundamentally different topologies because of homoplasy, i.e. incorrect coding, different evolutionary histories of the characters (character conflict), or incomplete sampling resulting in only partial signal of evolution. It is often practically impossible to display all trees from a given analysis. Consensus methods combine the common evidence of many alternative and conflicting

topologies from all trees recovered in the original phylogenetic analysis, i.e. all fundamental trees, into one single tree. Thereby, taxonomic congruence is a measure of the stability of the relationships of the taxonomic groups in the data set (Mickevich, 1978; Nelsen, 1979). Consensus analyses can be used to test for conflict and congruence amongst cladograms from different analyses of the same taxa using different sets of characters (its original function, Adams, 1972; Carpenter, 1988) or multiple equally best trees from a single analysis of one data set.

The strict consensus method (Schuh & Polhemus, 1981; Sokal and Rohlf, 1981) is the most conservative consensus method (Hillis, 1987; de Queiroz, 1993). It only retains those clades or components that are present in all fundamental trees. In the worst-case scenario when no common components are found amongst the fundamental trees the strict consensus tree is the unresolved bush. The strict consensus method is most widely used for combining the evidence from fundamental trees of a single analysis of one data set.

A slight relaxation of the strict consensus rule is the combinable component or semi-strict consensus method (Nelson, 1979; Bremer, 1990). A semi-strict consensus tree accepts components that are resolved in some fundamental trees but polytomous in others. The reasoning behind accepting a possible resolution of a polytomy is that it does not contradict the polytomy itself. If all fundament trees are fully resolved, the semi-strict consensus tree is identical to the strict consensus tree.

A majority-rule consensus tree (Margush & McMorris, 1981) retains not only those components that are found in all fundamental trees, but also those that occur in more than a pre-specified percentage (typically >50%) of the underlying fundamental trees. The majority-rule consensus method is useful if many different topologies are to be combined and the strict consensus tree is very unresolved. Majority-rule consensus trees are typically used in combining trees from a bootstrap analysis and Bayesian inference.

Adam consensus trees (Adams, 1972) contain nested taxa common to all fundamental cladograms and place those taxa of conflicting positions at the nearest common node. Adam consensus trees thus contain components not observed in any of the fundamental cladograms and have to be treated with caution. They are useful, however, for determining the degree of preserved structure in the fundamental cladograms, especially if there are 'rogue' taxa that appear in widely different positions in the different fundamental trees (Funk, 1985; Hillis, 1987).

While consensus methods are widely used they are also controversial (Penny *et al.*, 1982; Farris, 1983; Miyamoto, 1983, 1985; Hillis, 1987; Patterson, 1987; Cracraft &

Mindell, 1989; Kluge, 1989; Barrett *et al.*, 1991; Donoghue & Sanderson, 1992; Bull *et al.*, 1993; de Queiroz, 1993; Chippindale & Wiens, 1994; Huelsenbeck *et al.*, 1994; Wiens & Chippindale, 1994; Miyamoto & Fitch, 1995; Wiens & Reeder, 1995; Bremer, 1996; Huelsenbeck & Bull, 1996; Nixon & Carpenter, 1996; Baker *et al.*, 1998; Ballard *et al.*, 1998). Opponents of consensus analysis (e.g. Hillis, 1987; Cracraft & Mindell, 1989; Kluge, 1989; Barrett *et al.*, 1991; Pesole *et al.*, 1991; Shaffer *et al.*, 1991; Swofford, 1991; Marshall, 1992; Chippindale & Wiens, 1994; Wiens & Chippindale, 1994; Bremer, 1996; Nixon & Carpenter, 1996) claim that the underlying data need to be tested for character congruence and that consensus analyses are not repeatable in a strict statistical sense (Kluge, 1989). They prefer the use of combined data analysis. Some authors who disagree with the use of consensus trees for the comparison of hypotheses from different data sets nevertheless agree on their use for the comparison of all equally parsimonious or likely trees from the analysis of one data set (e.g. Barrett *et al.*, 1991; Kluge & Wolf, 1993; Chippindale & Wiens, 1994).

1.2.7.2 Character congruence: total evidence and simultaneous analysis

While consensus trees form a compromise between all fundamental cladograms, total evidence analysis (Carnap, 1950; Hempel, 1965) seeks to find the best-fitting phylogenetic hypothesis based on all available evidence. Miyamoto (1985) and Kluge (1989) provided major contributions to our understanding of the concept and to the correct implementation in phylogenetic analyses. Proponents of strict cladistics claim that only simultaneous analysis of all available information can give the most parsimonious solution that is the only hypothesis acceptable under the principle of cladistics. This point, however, is strongly debated because there is no proof that the hypothesis from the combined analysis is closer to the true tree than any of the hypotheses from separate analyses (Barrett *et al.*, 1991; Bull *et al.*, 1993; Huelsenbeck *et al.*, 1994).

The advantages and disadvantages of combined analysis are best understood in comparison to consensus analysis. For each argument against one method, there is also a counter argument against the alternative method. Which method to choose depends on the type of question that is asked. If the choice seems arbitrary, character congruence can be tested.

The main argument for consensus analysis lies in its ability to test for alternative hypotheses of the evolutionary history of separate data sets (Bull *et al.*, 1993; Huelsenbeck *et al.*, 1994). It is argued, that if separate analyses result in fundamentally different topologies, the different types of characters must have evolved under different evolutionary

processes, providing there is no sampling error. If different evolutionary processes are the cause of the observed phylogenies, then combination of the data sets can lead to positively misleading results (Felsenstein, 1978; Section 1.2.4.3.3) because at least one of the data sets will be violating the assumptions of the model. Proponents of combined analysis argue that accounting for different evolutionary models for the different sets of data will lead to the use of maximum likelihood methods, a criterion that is in strong opposition to the principles of cladistics.

Separate analyses can determine the degree of support for the optimal phylogenetic reconstruction of the individual data sets by recovering information about the amount of phylogenetic signal in the data (Huelsenbeck *et al.*, 1994). If the next best solution only requires a few more steps there is not much support for the optimal solution. If, on the other hand, the suboptimal solution requires many additional steps there is a lot of support for the optimal solution and it carries significant phylogenetic signal (Kitching *et al.*, 1998). Proponents of separate analysis also claim that a data set with a large number of characters (e.g. molecular data) might swamp a data set with fewer characters (e.g. morphological data) in a combined analysis (Hillis, 1987; Patterson *et al.*, 1993). In practice this argument does not seem to hold true. Morphological characters often contain strong signals that filter through in the combined analysis (e.g. Baker *et al.*, 1998).

On the other hand, combined analyses can tackle the presence of different evolutionary histories and phylogenetic signal by differential weighting of the data sets (Hillis, 1987; Barrett *et al.*, 1991; Chippindale & Wiens, 1994). Many studies found that, especially under character dependence, downweighting fast evolving, misleading characters and/or adding more slowly evolving ones can lead to a better estimate of phylogeny (Wheeler & Honeycutt 1988; Swofford & Olsen, 1990; Donoghue & Sanderson 1992; Doyle, 1992; Bull *et al.*, 1993; Dixon & Hillis, 1993; Chippindale & Wiens, 1994; Huelsenbeck *et al.*, 1994). Alternatively, non-linear transformation (Steel *et al.*, 1993c) can be used for inconsistent data to account for superimposed substitutions (Chippindale & Wiens, 1994). Simply ignoring fast-evolving characters may result in the unnecessary loss of phylogenetic signal.

However, there are a number of fundamental problems to the application of partitioning data sets and differential weighting. Firstly, it is inherently difficult to decide which partitions are natural (Chippindale & Wiens, 1994), or if partitions exist at all in nature (Kluge & Wolf, 1993). Secondly, even if each data set is accepted as an individual partition there is no proof that characters from different partitions are more likely to be

independent from each other than those within the same partition (Kluge & Wolf, 1993). Thirdly, it is almost impossible to detect which characters cause the potentially misleading property of a partition (Huelsenbeck *et al.*, 1994) because for this the true phylogeny has to be known which is usually the unknown quantity in a phylogenetic analysis.

On the other hand, consensus analysis ignores other weighting problems (Miyamoto, 1985) because consensus trees often imply more character state changes than are found in the individual trees. This can lead to incorrect estimates of the pattern of evolution (Miyamoto, 1985; Hillis, 1987). Simply combining the fundamental trees of the separate analyses does not even imply equal weighting. If the number of characters in the different underlying data sets is different, the data sets as a whole will be weighted equally but the constituent characters will be weighted unequally (Miyamoto, 1985; Cracraft & Mindell, 1989).

The consensus tree is probably the more conservative estimate of the phylogeny because it takes into account all evidence that the fundamental trees have in common. It is therefore less prone to lead to a wrong hypothesis about the evolutionary history (Barrett *et al.*, 1991; Swofford, 1991). On the other hand, the consensus tree lacks any information that is found only in a selection of fundamental trees while the combined analysis can resolve the conflict according to the evidence and give a much more resolved solution (Hillis, 1987; Barrett *et al.*, 1991; Bull *et al.*, 1993; de Queiroz, 1993). However, a single solution from the combined analysis can fail to provide evidence of how many components of the combined analysis can also be found in the separate analyses and vice versa (de Queiroz, 1993). While there is still no indication as to which hypothesis is closest to the true tree (Barrett *et al.*, 1991; Chippindale & Wiens, 1994; Huelsenbeck *et al.*, 1994) the choice is fundamentally one between accuracy and precision, between choosing an ambiguous solution (the consensus tree) that most likely contains the truth and an unambiguous solution (combined hypothesis) that may contain components that are not found in any of the fundamental trees. Under true cladistics the most parsimonious reconstruction is as precise as possible with the least number of *ad hoc* hypotheses, a condition only achieved with combined analysis (Miyamoto, 1985; Kluge, 1989; Kluge & Wolf, 1993; Chippindale & Wiens, 1994; Bremer, 1996; Nixon & Carpenter, 1996).

1.2.7.3 Conditional data combination or the prior agreement approach

Under the 'conditional data combination' (Huelsenbeck *et al.*, 1996) or 'prior agreement approach' (Chippindale & Wiens, 1994) partition heterogeneity has to be tested prior to combined analysis and separate analyses have to be carried out unless partition

homogeneity can be proven (Bull *et al.*, 1993; de Queiroz, 1993; Huelsenbeck *et al.*, 1994; Huelsenbeck & Bull, 1996; Baker *et al.*, 1998; Ballard *et al.*, 1998). Separate analyses result in conflicting hypotheses either because of different evolutionary processes (e.g. rate variations, functional constraints, selection pressure, horizontal gene transfer, recombination in bacteria, ancient polymorphism) or sampling error (Bull *et al.*, 1993). If the evolutionary process differs between the different partitions the conflict cannot be reduced with statistical transformations and combined analysis would violate the assumptions for at least one of the partitions causing the analysis to be positively misleading (Felsenstein, 1978, Section 1.2.4.3.3). However, if the cause for the differential evolutionary rates or transformational probabilities can be determined, a combined analysis under differential weighting (non-uniform weighting, Felsenstein, 1981b; Steel *et al.*, 1993c) is allowed.

Further complications arise when characters are not independent (see Section 1.2.1.2). If independence cannot be proven consensus analysis should be carried out (de Queiroz, 1993). If all characters are independent combined analysis is probably appropriate because potential conflict is among the characters and not the separate data sets (Kluge, 1984; Shaffer *et al.*, 1991; Swofford, 1991; Nixon & Carpenter, 1996). However, if separate analyses reveal more than two data sets with the same phylogeny but with low support then high levels of homoplasy have to be suspected. Under such circumstances a combined approach is considered acceptable because it increases the number of characters and thereby potentially the phylogenetic signal (de Queiroz, 1993). If character dependence within data sets is suspected and separate analyses result in conflicting but weakly supported hypotheses, neither of the two approaches is more appropriate and both should be used and the results compared (Shaffer *et al.*, 1991; Swofford, 1991; de Queiroz, 1993). If heterogeneity is the result of a single 'rogue' taxon a simultaneous analysis can be carried out (Littlewood & Smith, 1995). Barrett *et al.* (1991) suggested that if data sets are combined some weighting scheme should be applied because using equal weights is as much a biological assumption as applying a specific weighting scheme (Sober, 1988).

If separate analyses result in strongly supported, fully resolved trees, conflict is likely to be among the data sets, not the characters. Under such circumstances, at least one of the data sets would violate the assumptions of the model in combined analysis and consensus analysis should be used (de Queiroz, 1993). However, if the time between branching events is large, ancestral polymorphism can lead to discordance between gene and species

trees (Bull *et al.*, 1993) and combined analysis might be appropriate; if no timing data are available, consensus analysis should be employed.

Many proponents of combined analysis agree that performing both, consensus and combined analysis and comparing both results is a good practical solution to the various problems mentioned above (Miyamoto, 1985; Cracraft & Mindell 1989; Kluge 1989; Barrett *et al.*, 1991; Donoghue & Sanderson 1992; Chippindale & Wiens, 1994; Wiens & Chippindale, 1994; Wiens & Reeder, 1995). Several authors (Miyamoto, 1985; Kluge & Wolf, 1993; Bremer, 1996) pointed out that the consensus tree of separate analyses is useful for highlighting the (in-) stability and ambiguity of the clades of the fundamental trees; they show the difference in the explanatory power between the total evidence and taxonomic congruence. Separate analyses can also be used to re-evaluate homologies (Bremer, 1996). However, all these authors favour the use of combined analysis for establishing a phylogenetic hypothesis, especially in contrast to some recent approaches that would favour a hypotheses from a molecular data set over those from a morphological one (Bremer, 1996).

1.2.7.4 Test for two trees

Numerous methods have been proposed for testing character congruence: Mickevich-Farris index (Mickevich & Farris, 1981); the non-parametric bootstrap (Felsenstein, 1985a); the Bremer support/decay index (Templeton, 1983; Bremer, 1988); the Kishino-Hasegawa Test [KH-Test] (Kishino & Hasegawa, 1989); the consistency index (Kluge, 1989); the T-PTP Test (Faith, 1991); the partition homogeneity test or incongruence length difference test (Rodrigo *et al.*, 1993; Farris *et al.*, 1994, 1995); the likelihood ratio test (Huelsenbeck & Bull, 1996; Waddell *et al.*, 2000); the partitioned Bremer support (Baker & DeSalle, 1997; Baker *et al.*, 1998); the Shimodaira-Hasegawa Test [SH-Test] (Shimodaira & Hasegawa, 1999). However, all these tests have major drawbacks. Many of these methods are biased and potentially introduce rather more ambiguity in the result than they recover support for individual trees (Wiens & Chippindale, 1994).

Tests for character heterogeneity under the parsimony criterion lack an underlying model that can account for known biases; they might confuse true topology incongruence with systematic error (Huelsenbeck & Bull, 1996; Dolphin *et al.*, 2000; Goldman *et al.*, 2000; Baker *et al.*, 2001; deBry, 2001; Sota & Vogler, 2001; Yoder *et al.*, 2001; Buckley *et al.*, 2002; Dowton & Austin, 2002) and bootstrap estimates are biased (Zharkikh & Li, 1992a, 1992b; Hillis & Bull, 1993; Steel *et al*, 1993a; Section 1.2.6.2).

Although some tests have been developed that use the likelihood criterion they have severe practical limitations. Goldman *et al.* (2000) showed that the non-parametric KH-Test requires the *a priori* selection of the two topologies to be included, an assumption violated in most practical applications because one of the topologies is usually the one found to be optimal after the phylogenetic analysis. Using an *a posteriori* topology either reduces the power of the test or increases its type I error. They proposed the use of the SH-Test (Shimodaira & Hasegawa, 1999) or the SWOH-Test (Hillis, 1996; Swofford *et al.*, 1996) instead when comparing *a posteriori* topologies. The problem with these tests is that they become practically infeasible with more than 5 or 6 taxa because all possible tree reconstructions have to be considered in the calculation of the test statistic. The problems Goldman *et al.* (2000) highlight also apply to most other congruence tests (Wilkinson, pers. comm.). Huelsenbeck & Bull (1996) introduced a parametric likelihood ratio test but it is only applicable to molecular sequencing data, not to morphological data, and it is computationally expensive and therefore again not useful for large data sets (Buckley *et al.*, 2002).

Buckley *et al.* (2002) introduced a congruence test based on Bayesian inference. While there are advantages to using a Bayesian approach (e.g. speed combined with confidence limits) there are still practical and theoretical limitations: programs to test congruence are still under development and if different parts of the same gene evolve under different models it is difficult to decide which model is appropriate (Buckley *et al.*, 2002; see also Section 1.2.4.3.6). If computational difficulties can be overcome, there are some advantages to using Bayesian techniques for congruence tests: Firstly, Bayesian methods are much faster than likelihood methods because no separate bootstrap run needs to be conducted to assess nodal support and no *a priori* decisions have to be made about topologies and parameter settings (Huelsenbeck & Bull 1996; Swofford *et al.*, 1996; Goldman *et al.*, 2000). Secondly, posterior probabilities are a much more intuitive measure of uncertainty (Gelman *et al.*, 1995). And thirdly, the strength of evidence of one model relative to another can be measured by comparing the phylogenetic signal among the partitions (Kass & Raftery, 1995).

Until a computationally feasible method for *a posteriori* tree comparison has been developed or computer technology has advance to make the current tests applicable for large data sets, 'simpler' tests have to be used. The violation of some of their assumptions has to be taken into account and the results have to be interpreted very carefully.

1.2.8 Conclusions

During the last 40 years empirical phylogenetic analysis techniques have come a long way from cladistic parsimony analyses conducted by hand (Hennig, 1966) to sophisticated tests of Bayesian inference (Li, 1996; Mau, 1996; Rannala & Yang, 1996). During all this time the theory was always one step ahead of what was practically achievable. The applicability and limitations of new algorithms could, and still can, only be tested when computer technology has caught up with the theory. And even when such tests commence it often takes a long time until the true restrictions are known.

This chapter has shown that amongst the multitude of analysis techniques there is not a single test whose assumptions are not likely to be violated by any given data set. Therefore, any new data set should be analysed with as many different methods as possible and only after all options are exhausted an informed but critical decision should be made of the potential phylogeny. Every systematist should remember that:

"Phylogeny cannot be observed. It is necessarily an inference from observations that bear on it, sometimes rather distantly, and that can usually be interpreted in more than one way." (Simpson, 1945)

1.3 Aims and Objectives

This thesis aims to integrate phylogenetic information from all types of nematodes. Whilst molecular sequences from a variety of animal and plant parasitic and soil nematodes were publicly available, few sequences had yet been obtained from marine nematodes. New sequences from marine nematode species were added to the existing database of several hundred nematode species. Additionally, all morphological characters used in previous classification attempts were examined and a suite of characters was chosen that was applicable in some form to all nematodes species under investigation.

Modern phylogenetic analysis methods also allow for a more detailed analysis of the effects of applying a variety of mathematical algorithms. The reconstruction of the systematics of nematodes was therefore undertaken using a range of tools and the aim was to find an overall consensus amongst the mathematical models and the two data sets that were analysed both individually and in combination.

Using a suite of molecular and morphological characters and a range of analysis tools, an attempt is made to revise the phylogenetic relationships of the whole phylum. The revised classification is compared to the classical systems.

2. Materials and Methods

2.1 Sampling and identification

The samples for this study were taken from five different sources:

i) *Southampton Water:* Three different sites were sampled by the author in autumn 2001: Warsash Mud Flat (22/10/01, 1° 18' W, 50° 51' 30" N), Weston Shore (24/10/01, 1° 23' W, 50° 53' 30" N) and Hythe Salt Marsh (7/11/01, 1° 23' W, 50° 52' N). Altogether, 47 samples were taken from different microhabitats on the three shores. Figure A1 (Appendix I) shows the three sampling sites in Southampton Water, Table 2.1 contains a list of all samples and the microhabitats they were taken from and Table 2.2 contains physical data from the three sites.

The sampling procedure was identical at all three sites. Sediment cores were taken using a Perspex hand corer (5.7 cm diameter) that was pushed into the sediment, twisted and pulled out. The sediment was pushed out to the top of the corer with a plunger and two horizons were cut off, 0 - 2 cm and 2 - 5 cm. The sediment was transferred into a sample container and fixed in 99.7% molecular grade ethanol. At the high shore at Weston the sediment was too difficult to sample with the corer. Instead, for samples 29, 30 and 31, sand was put into a 2-litre measuring cylinder. The cylinder was filled up with filtered seawater and inverted several times to mix the sample. The heavy sediment was allowed to settle for a minute and the suspension containing the nematodes was decanted over a 45 μm sieve. The cylinder was refilled and the procedure repeated until the flow was clear. The sediment retained on the sieve was washed into a bottle and fixed in 99.7% ethanol as with the other samples. For samples 15 and 16 (Warsash) and 32 and 33 (Weston), a cluster of algae from mid shore was washed in a jug filled with filtered seawater (45 μm mesh). The water was poured through a sieve and the sediment and animals retained in the sieve fixed in ethanol as before.

Air and water temperatures were measured with a standard mercury thermometer and salinity with a hand-held refractometer. The redox potential depth was measured as the depth before the sediment turned black, using a ruler.

Table 2.1: Sample locations and dates, including position on shore where each samples was taken.

Sample No.	Location	Date	Shore Level	Replicate	Horizon
1	Warsash	22/10/01	Low Water	2A	0-2 cm
2	Warsash	22/10/01	Low Water	2B	2-5 cm
3	Warsash	22/10/01	Mid Water	1A	0-2 cm
4	Warsash	22/10/01	Mid Water	1B	2-5 cm
5	Warsash	22/10/01	Mid Water	2A	0-2 cm
6	Warsash	22/10/01	Mid Water	2B	2-5 cm
7	Warsash	22/10/01	Mid Water	3A	0-2 cm
8	Warsash	22/10/01	Mid Water	3B	2-5 cm
9	Warsash	22/10/01	High Water	1A	0-2 cm
10	Warsash	22/10/01	High Water	1B	2-5 cm
11	Warsash	22/10/01	High Water	2A	0-2 cm
12	Warsash	22/10/01	High Water	2B	2-5 cm
13	Warsash	22/10/01	High Water	3A	0-2 cm
14	Warsash	22/10/01	High Water	3B	2-5 cm
15	Warsash	22/10/01	Mid Water	Algae 1	-
16	Warsash	22/10/01	High Water	Algae 2	-
17	Weston Shore	24/10/01	Low Water	1A	0-2 cm
18	Weston Shore	24/10/01	Low Water	1B	2-5 cm
19	Weston Shore	24/10/01	Low Water	2A	0-2 cm
20	Weston Shore	24/10/01	Low Water	2B	2-5 cm
21	Weston Shore	24/10/01	Low Water	3A	0-2 cm
22	Weston Shore	24/10/01	Low Water	3B	2-5 cm
23	Weston Shore	24/10/01	Mid Water	1A	0-2 cm
24	Weston Shore	24/10/01	Mid Water	1B	2-5 cm
25	Weston Shore	24/10/01	Mid Water	2A	0-2 cm
26	Weston Shore	24/10/01	Mid Water	2B	2-5 cm
27	Weston Shore	24/10/01	Mid Water	3A	0-2 cm
28	Weston Shore	24/10/01	Mid Water	3B	2-5 cm
29	Weston Shore	24/10/01	High Water	Spades 1	surface
30	Weston Shore	24/10/01	High Water	Spades 2	surface
31	Weston Shore	24/10/01	High Water	Spades 3	surface
32	Weston Shore	24/10/01	Mid Water	Algae 1	uncovered
33	Weston Shore	24/10/01	Mid Water	Algae 2	waterline
34	Hythe	07/11/01	Creek	1A	0-2 cm
35	Hythe	07/11/01	Creek	1B	2-5 cm
36	Hythe	07/11/01	Creek	2A	0-2 cm
37	Hythe	07/11/01	Creek	2B	2-5 cm
38	Hythe	07/11/01	Creek	3A	0-2 cm
39	Hythe	07/11/01	Creek	3B	2-5 cm
40	Hythe	07/11/01	Marsh Top	Knifes 1	-
41	Hythe	07/11/01	Marsh Top	Knifes 2	-
42	Hythe	07/11/01	Mud Flat	1A	0-2 cm
43	Hythe	07/11/01	Mud Flat	1B	2-5 cm
44	Hythe	07/11/01	Mud Flat	2A	0-2 cm
45	Hythe	07/11/01	Mud Flat	2B	2-5 cm
46	Hythe	07/11/01	Mud Flat	3A	0-2 cm
47	Hythe	07/11/01	Mud Flat	3B	2-5 cm

Table 2.2: Physical properties of the three sampling sites in the Solent: Salinity, Temperature and Redox Potential Depth (RPD).

Property	Site		
	Warsash	**Weston**	**Hythe**
Salinity	34‰	22‰	31‰ Creek 33‰ Mud flat
Temperature	14°C	13°C Air 14.5°C Water	11.5°C Puddles
RPD	few mm	≥ 5 cm mid and high shore 4 cm low shore	1 cm Creek Surface on Mudflat and Marsh Top

18 core samples were water decanted in the laboratory: the sediment was poured into a 2-litre measuring cylinder and the bottle rinsed out with filtered tap water. The cylinder was filled with filtered tap water and inverted several times. The sediment was allowed to settle for 2 minutes and the water decanted through a sieve. The water decantations were repeated up to 5 times or until the flow was clear. The sediment retained on the sieve was again preserved in ethanol. The top 2 cm horizon of six cores and one algae sample were used for the final work: Mid-water sediment from Warsash (Sample 3), low (S17) and high shore (S29) sediment and algae (S32) from Weston Shore and marsh top (S40) and mud flat (S46) sediment from Hythe.

ii) *Foodbancs/Antarctica*: In June 2000, Mega-corer (10 cm diameter) sediment cores were taken on the FOODBANCS Cruise NBP 00-04 off the Antarctic Peninsula on board the Antarctic research and supply vessel, *Nathaniel B Palmer*. The top 2 cm of the core were sliced off with a metal slicer and the sediment put into a 2-litre measuring cylinder. The cylinder was filled with filtered seawater and inverted 5 times to mix the sample. The heavy sediment was left to settle for a few minutes and the water containing the nematodes decanted over a 45 μm sieve. The cylinder was re-filled and the procedure repeated until the flow was clear. The sediment retained on the sieve was rinsed with 99.7% molecular grade ethanol and washed into a small glass bottle topped up with ethanol for preservation. The only sample from this cruise that was used for the current work came from site A, deployment #597, core 11 (65° 10' S, 64° 47 W).

iii) *Whale falls off California/USA*: In October 1999, Eckman cores (20cm x 20cm) of sediments from the vicinity of a whale fall off the coast of California (33° 29.6' N, 119° 22.0' W) were taken by the submersible *Alvin* during the Atlantis II Cruise 3/42. Cores were taken along several transects running from the side of the whale carcass away from it, up to a distance of 30 m. Each core was divided into four 10 x 10 cm sub-cores and one sub-core was preserved for molecular analysis of nematodes. The top 2 cm were treated in the same way as the cores from Antarctica. Two samples from dive AD 3483 were

examined; they were taken 0 m (E6 B) and 1m (E11 D) away from the cadaver of a 1.5 year old whale carcass.

iv) *The model nematode Caenorhabditis elegans* (strain Bristol N2). This terrestrial species can easily be cultured on agar in Petri dishes and animals were kindly supplied by Dr. Candida Rogers, School of Biological Sciences, University of Southampton. Animals were picked alive from the agar and either DNA was extracted straight away if they were used as PCR controls, or they were fixed in alcohol and desiccated in molecular grade glycerol for the time series experiment (see Section 2.2.3).

v) *Gilkicker Lagoon, Southeast Hampshire/UK*: On 19th October 2000, the author took sediment samples from the southern side of Gilkicker Lagoon (1° 08' W, 50° 46' 30" N). Surface sediment was scooped up with a little spoon and put into small (75 ml) sampling bottles. Five samples were fixed in 95% molecular grade ethanol (Aldrich) and three samples left in lagoonal water. Using a stereomicroscope, four nematodes were successfully picked out of untreated sediment, placed into individual 0.5 ml PCR tube containing Tris-EDTA buffer and stored in the fridge.

After the initial water decantation animals from sources i), ii) and iii) above were treated in the same way. For all stages (including the above), sieves with a mesh size of 45 μm were used and any water that came into contact with any sample was sieved beforehand.

After the water decantation the sediment was poured into a sieve and the bottle and sample were washed with water. The sieve with the retained sediment was rinsed thoroughly with LUDOX™ 50 to get rid of all the water. The LUDOX™ 50 was mixed with sieved tap water to a specific gravity 1.15 which is similar to that of nematodes. The sediment was washed into a bottle that was subsequently filled up with LUDOX™ 50, inverted to mix the sample and left settling for a minimum of 4 hours. During this time particles of higher gravity sink to the bottom of the bottle but the nematodes (and other lighter particles) remain in suspension. After 4 hours the liquid containing the nematodes was decanted through a sieve, the LUDOX™ 50 rinsed off with water and the retained sediment washed into a petri dish. This technique is a modification of the LUDOX flotation technique of Platt & Warwick (1983). The number of nematodes retained was checked under a stereomicroscope and the LUDOX™ 50 extraction repeated if a higher number of nematodes was required. When enough nematodes were extracted, the remaining sediment was washed through a sieve, rinsed with water and again fixed in

ethanol. During the extraction procedure nematodes were kept in LUDOX™ 50 or water for no more than nine hours (two decantations).

Individual nematodes were picked out of the petri dish and put into a cavity block containing molecular grade glycerol. The cavity block was covered with a glass lid leaving a small gap for evaporation and kept in an airtight desiccator over night. Anhydrous copper sulphate crystals were used as the desiccant. The desiccated nematodes were mounted individually onto slides (75 x 25 mm with a round cover slip of 19 mm diameter) resting in a drop of anhydrous glycerol surrounded by a paraffin wax ring. The slide was placed on a hot plate until the wax ring was melted sealing the cover slip onto the slide. Some specimens from sources ii) and iii) were desiccated for several days or weeks.

The nematodes were identified to the highest possible taxonomic level by the staff of The Natural History Museum London: Ms Nicola Debenham, Dr John Lambshead and Dr Tim Ferrero, using taxonomic keys (Bastian, 1865; Riemann, 1966; Gerlach & Riemann, 1973/74; Lorenzen, 1977; Platt, 1983; Platt & Warwick, 1983, 1988; Warwick *et al.*, 1998).

2.2 Molecular analysis

2.2.1 DNA extractions

Three methods of DNA extraction were compared. Initially, nematodes were placed into 0.5 ml PCR tubes containing 15 μl of Microzone Microlysis and run through the following thermal cycle: 5 min at 65°C, 2 min at 96°C, 4 min at 65°C, 1 min at 96°C, 1 min at 65°C, 30 sec at 96°C and held at 20°C. Amplification of DNA that was extracted with this method was very good. Secondly, the microlysis protocol from Thomas *et al.* (1997) was tried out but not followed up for the main part of the study because of contamination problems. The extraction method that was finally applied came from Stanton *et al.* (1998) and was adjusted accordingly. In all three methods extraction mixtures were used directly for DNA amplification. The method by Stanton *et al.* (1998) was favoured because it was the cheapest and easiest method and gave high quality DNA amplifications and sequences. DNA of single worms were extracted in 0.5 ml PCR tubes containing 20 μl 0.25 M NaOH and frozen at -20°C for a minimum of one hour. Specimens from sources ii) and iii) and time series 1 animals that were incubated for up to 36 hours were incubated at 25°C straight away without an initial freezing step. After freezing, extracts were incubated at 25°C overnight. The next morning, the lysate was heated for 3 min at 99°C. When the tubes had cooled to room temperature, 4 μl 1 M HCl,

10 µl 0.5 M Tris/HCl (pH 8.0) and 5 µl 2% Triton X-100 were added. The mix was stirred briefly and spun down in a tabletop-centrifuge at 3000g for a few seconds using the touch-spin function. The lysate was heated again for 3 min at 99°C and the extract used either directly for PCR amplification or stored at -20°C. In cases where the extraction digest became rather solid an additional 40 µl Tris/HCl buffer were added to re-dissolve the sample.

2.2.2 PCR optimisations

Five different genes were examined for their suitability for sequence analysis of marine nematodes: The large and small subunits of the mitochondrial ribosomal DNA genes (16S and 12S), the small subunit of the nuclear ribosomal DNA gene (18S) and the internal transcribed spacers of the nuclear ribosomal genes (ITS-I and ITS-II). Table 2.3 contains a list of all primers tested and used in this study, including their sequence, the position at which they align to the respective gene in *C. elegans* and the reference where applicable. The base pair positions are given for 12S and 16S in respect of the whole mitochondrial genome (Okimoto *et al.*, 1992; GENBANK Acc. No.X54252) and for 18S and the two ITS regions in respect of the nuclear repetitive region containing all ribosomal genes and external and internal transcribed spacers (Ellis *et al.*, 1986; GENBANK Acc. No. X03680). Primers 16SarMOD and 16SbrMOD are modified versions of the universal 16S primers (16Sar and 16Sbr). The former two 16S primers and primers 300F and 1055R were previously developed by a research assistant (Lange) at the Southampton Oceanography Centre and were the starting point of this investigation. Known sequences of nematodes retrieved from GENBANK were used to design all other 16S and 12S primers. A denotation of "CE" in the name of the 16S primers stands for the exact *C. elegans* sequence, while "D" stands for a primer sequence with ambiguous bases accounting for all available nematode sequences. The remaining primer sequences were taken from publications. The position for primers rDNA1.58S and rDNA2.144 (Powers *et al.*, 1997) could not be determined. All primers and PCR reaction conditions were tested on 3 unidentified nematodes from Gilkicker Lagoon (source v), specimens from Antarctica (source ii) and on *C. elegans* (source iv).

For each primer pair, annealing temperatures where varied around the melting temperatures of the primers to find the optimum temperature. Concentrations of $MgCl_2$, primers and template DNA were also varied to find the optimum conditions. None of the primers for mitochondrial genes resulted in amplified products for marine nematodes while the use of 18S primers from Blaxter *et al.* (1998) and ITS primers from Chilton & Gasser

(1999) resulted in successful PCR products from the same individual worms. The ITS primers from Power *et al.* (1997) also failed to amplify marine nematode DNA.

Table 2.3: Description of primers: region of amplification, sequence, base pair position on *C. elegans* gene and reference (Ref.). a = Lange (unpubl.), b = this study, c = Blaxter *et al.*, 1998, d = Powers *et al.*, 1997, e = Chilton & Gasser, 1999.

Primer Name	Amplified Region	Sequence (5'-3')	Alignment onto *C. elegans* gene	Ref.
16SarMOD	16S	CGCCTGTTTATCAAAAACAT	10779 – 10798	a
16SbrMOD	16S	CCGGTCTGAACTCAGATCACGT	11253 – 11232	a
300F	18S	AGGGTTCGACTCCGGAG	1289 – 1305	a
1055R	18S	CGGCCATGCACCACC	2179 – 2165	a
$16SF_1CE$	16S	AATGGCAGTCTTAGCGTGAG	10839 – 10858	b
$16SF_1D$	16S	AAWRGCASYYTTAGCGTGAK	10839 – 10858	b
$16SR_1CE$	16S	AATTTCCGAAGACTTATCTTTG	11007 – 10986	b
$16SR_1D$	16S	AATTTCYRAAGACTTWTCTTWG	11007 – 10986	b
$16SF_2CE$	16S	ACAAAGATAAGTCTTCGG	10985 – 11002	b
$16SF_2D$	16S	ACWAAGAWAAGTCTTYRG	10985 – 11002	b
$16SR_2CE$	16S	GAATTAAACTAATATCACG	11251 – 11233	b
$16SR_2D$	16S	GAAYTAAACTAATWTCAMG	11251 – 11233	b
$12SF_1$	12S	GTTCCAGAATAATCGGCTAG	1027 – 1046	b
$12SF_2$	12S	GACTCGTGTATGATCG	1311 – 1326	b
$12SR_1$	12S	CAACTTACTCCCCTTTGGG	1556 – 1538	b
G18S4	18S	GCTTGTCTCAAAGATTAAGCC	963 – 983	c
22R	18S	GCCTGCTGCCTTCCTTGGA	1350 – 1332	c
22F	18S	TCCAAGGAAGGCAGCAGGC	1332 – 1350	c
26R	18S	CATTCTTGGCAAATGCTTTCG	1861 – 1841	c
24F	18S	AGRGGTGAAATYCGTGGACC	1802 – 1821	c
$24F_1$	18S	AGAGGTGAAATTCTTGGATC	1802 – 1821	c
13R	18S	GGGCATCACAGACCTGTTA	2336 – 2318	c
18P	18S	TGATCCWKCYGCAGGTTCAC	2790 – 2771	c
rDNA2	ITS	TTGATTACGTCCCTGCCCTTT	2523 – 2503	d
rDNA1.58S	ITS	ACGAGCCGAGTGATCCACCG		d
rDNA2.144	ITS	GTAGGTGAACCTGCAGATGGAT		d
NC1	ITS-2	ACGTCTGGTTCAGGGTTGTT	3291 – 3310	e
NC2	ITS-2	TTAGTTTCTTTTCCTCCGCT	3764 – 3745	e
NC5	ITS-1	GTAGGTGAACCTGCGGAAGGATCATT	2667 – 2692	e
NC13R	ITS-1	GCTGCGTTCTTCATCGAT	3207 – 3190	e

The initially aim of this thesis was to investigate the structures of several different nematode populations but the failure of amplification of the mitochondrial genes prevented this. However, amplification of the 18S gene proved successful and the presence of three studies into the phylogenetic relationships of nematodes using this gene was a good starting point to a further investigation of this phylum. This provided the chance of adding many, previously under-represented, marine species to the existing database of nematode 18S sequences. The final protocol for the amplification of the partial 18S gene is described in Section 2.2.4.

ITS genes have been used successfully in strain detection of parasitic nematodes (e.g. Powers *et al.*, 1997; Chilton & Gasser, 1999). In trials with *C. elegans* and unidentified marine nematodes the ITS genes showed high length variations in the amplification

products suggesting a high mutation rate. Additionally, the ITS regions also potentially exhibit high levels of intragenomic variations. If such variations are present DNA amplification may result in multiple products from the same individual and products from different individuals may not be from homologous regions. Therefore, the ITS genes were not used for phylogenetic reconstructions.

2.2.3 Time series

Very low or no DNA yields following PCR were observed for animals that had been desiccated in glycerol for several days or weeks or had been on slides for several weeks compared to specimens that had been extracted fresh and had not been desiccated in glycerol. In order to test the long-term effect of glycerol on the DNA yield, a time series using *Caenorhabditis elegans* was set up:

Series 1: Live animals were picked into distilled water containing a few drops of molecular grade ethanol. Then glycerol was added to the water. At each time interval, three worms were taken out of the glycerol and each worm was placed into a 0.5 ml PCR tube containing 20 μl of 0.25 molar NaOH. Worms were taken out after one hour, at two hour intervals between two and eight hours, 16 and 32 hours and 40 and 56 hours. All specimens between one and 32 hours were fully extracted straight away without freezing. All later samples were stored in NaOH at -20°C for a minimum of 24 hours before the DNA was fully extracted. Ideally, animals should have been picked straight into glycerol, but it proved difficult to place live specimens into glycerol as they curled tightly around the needle and could not be scraped off in one piece. When they were picked into water, they let go of the needle immediately. Ethanol (99.7%) was used to sedate the worms more quickly.

Series 2: In this series, worms were desiccated in glycerol for up to 14 days. They were either placed directly into glycerol as in Series 1 or first fixed in 99.7% molecular grade ethanol for two days and then transferred into glycerol. Three animals were taken out of the glycerol and placed into NaOH once a day for 14 days and after 28 and 54 days in the unfixed and after 38 and 60 days in the fixed experiment. The differences in times for the last two points were the results of a miscalculation. As a control, animals were also fixed in ethanol for two days and then kept in distilled water instead of glycerol to assess the effect of glycerol on DNA degradation as opposed to the effect of a lack of fixing. Due to a lack of animals the control was only carried out for up to 14 days and at some time intervals (2, 4, 5, 8 and 10 days) only two animals were taken out. All samples were frozen

for at least 24 hours before they were fully extracted. The worms were again initially picked into distilled water as in Series 1.

Series 3: The agar of one petri dish full of worms was cut into three segments, two segments were transferred onto clean petri dishes and the agar in the three dishes covered with 70%, 95% and 99.7% molecular grade ethanol respectively. After two and seven days respectively, worms were transferred from ethanol into glycerol. Again, three worms were taken out of the glycerol at each time interval and put into PCR tubes containing NaOH. The following time intervals were used: one day, four days, one week, two weeks and four weeks. Two sets of controls were taken: Before the animals were picked into glycerol, six worms were picked out of the ethanol, three of those were put straight into NaOH and a further three worms were quickly washed in distilled water before they were put into NaOH. Controls were taken separately for all three ethanol concentrations. All samples were frozen for a minimum of 24 hours before they were fully extracted.

The results were compared by PCR amplification of the first 400 bp of the 18S gene (primers G18S4 and 22R). Ideally, the increase in DNA concentration before and after PCR would have been measured. However, it is not possible to quantify the DNA concentration of single-worm-extraction and although care was taken that all worms were relatively the same size there were differences that may also have led to differences in the of DNA yield of the extraction. Therefore, the results were only scored as ranks of amplification success. A negative score (-) was given for unsuccessful amplifications, a positive score (+) was given for yields sufficient for direct DNA sequencing and a neutral score (0) was given for positive yields that were insufficient for direct sequencing.

No regression analysis for changes in DNA yield over time can be performed because only ranks were scored. However, changes between treatments in Series 2 and 3 were analysed using the Spearman Rank Correlation Coefficient. Initially, all three replicates per time were ranked individually. Then the three ranks for each time interval were averaged and the average ranks were re-ranked to provide a single rank for each time interval and treatment. The Spearman Rank Correlation Coefficient for ties was applied because ties occurred frequently in the data.

The correlation coefficient, r_s, is calculated as:

$$r_s = \frac{\sum\chi^2 + \sum\gamma^2 - \sum d^2}{2\sqrt{\sum\chi^2 * \sum\gamma^2}}, \text{ where}$$

$$\sum\chi^2 = \frac{N^3 - N}{12} - \sum T_\chi \quad \text{and} \quad \sum\gamma^2 = \frac{N^3 - N}{12} - \sum T_\gamma, \text{ where } T = \frac{t^3 - t}{12}$$

and $\sum d^2$ is the sum of squares of the differences between the ranks for each pair of observations, N is the number of values in each sample and t is the number of observations tied at a given rank.

The significance of the correlation coefficient for samples with less than ten observations (Series 3) is taken straight from the one-tailed distribution, for samples with more than ten observations it is taken for the Student's t one-tailed distribution where

$$t = r_s \sqrt{\frac{N - 2}{1 - r_s^2}} \text{ with degrees of freedom} = N - 2.$$

2.2.4 Final protocol for PCR and DNA sequencing

The partial 18S gene of 49 specimens was amplified in 3 segments using the following primer pairs (all Blaxter *et al.*, 1998): G18S4 and 22R, 22F and 26R and 24F and 13R; for a few animals the modified 24F1 primer was used. PCR reactions were performed using the following mix: 50 mM KCl (from 10x PCR buffer, Qiagen), 10 mM Tris/HCl (from 10x PCR buffer, Qiagen), 2.75 mM $MgCl_2$ (Qiagen), 0.8 mM dNTP Mix (Perkin Elmer), 10 pmol of each primer (10pm/μl) (Sigma Genosys), I U Taq Polymerase (Qiagen), 2 μl Q-Solution (Sigma) and 1.0 - 2.0 μl DNA digest filled to a total of 10 μl with ddH_2O (Sigma). PCR reactions were run on a Perkin Elmer DNA Thermal Cycler 480 or a Hybaid PCR Express under the following conditions: initial denaturation for 5 min at 95°C, followed by 35 cycles of 45 sec at 94°C, 30 sec at 60 - 65°C (dependent on primer pair) and 1 min at 72°C, followed by a final extension phase of 10 min at 72°C. The ramping time was 1°C per sec. The products were held at 4°C and then stored at - 20°C. Products of low DNA yield were re-amplified using 1 μl of the PCR product as template for the new PCR. For some re-amplifications the annealing temperature was lowered by a few degrees Celsius but all other conditions were left the same. 1 μl of the PCR product was visualised and photographed under UV light on a 1 % agarose (Sigma) gel stained with 1 % ethidium bromide (Sigma). The remaining PCR product was purified using the QIAquick PCR Purification Kit (Qiagen) and the DNA concentration calculated using the

Uvidoc system (Uvidoc 008-XD documentation system with Uvisoft Version 98 gel quantification software). 10 - 20 ng of purified DNA were cycle sequenced using BigDye Terminators (Perkin Elmer). 25 cycles of 10 sec at 96°C, 5 sec at 50°C and 4 min at 60°C were run on the MWG-Biotech Primus 96 ^{plus}. The ramping time was 1°C per sec. The products were held at 4°C. Cycle sequencing products were cleaned up using the Dye-Ex Spin Kit (Qiagen) and dried on a hotblock at 70°C. The clean DNA was re-suspended in 2 μl of red Microstop denaturing buffer (Microzone). For all animals, all 3 segments were sequenced in both directions on the ABI Prism 377 automated DNA sequencer (Perkin Elmer). 0.7 - 1.0 μl of denatured DNA were loaded onto the sequencing gel.

2.3 Analyses

2.3.1 Sequencing analysis

2.3.1.1 Alignment

For all animals, the forward and reverse sequences of each segment were compared using Chromas version 1.45 (McCarthy, 1997); ambiguous bases were compared to its complementary strand and adjusted accordingly. The three segments were nested and the complete sequence of the partial 18S rDNA gene edited into Fasta format.

Additional sequences were obtained from three sources:

a) The European Ribosomal RNA Database, an online server that contains all published rRNA sequences in aligned format based on the estimated secondary structure (Wuyts, *et al.*, 2002, http://rrna.uia.ac.be/ssu/list/Eucaryota.html).

b) Dr. Mark Blaxter, University of Edinburgh, gave access to a large database of unpublished nematode rRNA sequences which he kindly supplied for analysis. The sequences were all treated in the same way in the following alignment steps.

c) Peter Mullin, University of Nebraska, kindly provided rRNA sequences of six species of the Dorylaimia. The sequences were treated in the same way as those from b).

All available nematode sequences from the European RNA Database were downloaded. This alignment was used as first profile in the profile alignment mode of CLUSTAL_X version 1.81 (Thompson *et al.*, 1997) and the remaining unaligned sequences added as second profile. The default parameter settings were used. The resulting alignment contained many positions in which CLUSTAL misaligned an obviously conservative region of nucleotides following a very variable section. Such areas were

edited by hand in BioEdit version 5.0.9 (Hall, 2001) preserving the secondary structure as indicated in the sequences from the RNA database. The first 191 characters and all characters after position 2479 constitute variable terminal end positions and were excluded, as well as all-gap positions resulting from the above manipulations. Therefore, 1769 characters remained for the NEXUS file that was used for all subsequent analyses. The final alignment can be found in the NEXUS file in Appendix VI (file <molecular.nex> on CD).

2.3.1.2 Phylogenetic analysis tools

Parsimony and distance analyses were performed with PAUP* version 4.0beta10 (Phylogenetics Analysis Using Parsimony; Swofford, 2002). Bayesian inferences were calculated with MrBayes version 3.0B (Huelsenbeck & Ronquist, 2001). Likelihood setting parameters were calculated with Modeltest version 3.06 (Posada & Crandall, 1998). Phylogenetic trees were visualised using Treeview version 1.6.6 (Page, 1996). All PAUP* and MrBayes calculations were conducted on Solaris Sun Blade 2000 UNIX workstations with an UltraSPARK III 900MHz processor and 1024 Mb RAM.

2.3.1.3 Outgroup analysis

The outgroup analysis was performed with a smaller dataset to reduce computational time. 40 representatives of all major nematode clades were selected on the basis of an initial parsimony reconstruction of 214 nematode sequences (final data set, see below). Where possible, sequences of shorter branch length were used to reduce the problem of long-branch attraction. Table A2.1 in Appendix II (file <TableA2_1.rtf> on CD) gives the accession numbers for all sequences used in this study and Table A2.2 (file <TableA2_2.rtf> on CD) shows which sequences were included in the outgroup analysis.

Outgroups were selected according to Peterson & Eernisses's (2002) phylogenetic reconstruction of the metazoa. The European RNA Database was searched for all sequences of the Chaetognatha (3), Gastrotricha (3), Kinorhyncha (1), Nematomorpha (5), Priapulida (3) and Tardigrada (7) as these phyla appeared in the same clade as the Nematoda in at least one of the analyses of Peterson & Eernisse (2002) (numbers in parentheses correspond to the number of sequences found).

The parameter settings for all analyses were kept constant: parsimony was chosen as the optimality criterion and a heuristic search was conducted using ten repetitions of random sequence additions with TBR branch swapping. All other parameters were kept as default in PAUP*. All positions with gaps were excluded from all outgroup analyses to

reduce the chances of long-branch attraction because of homoplasy in ambiguously aligned regions.

At first the phylogenies of all 40 ingroup taxa alone and all 40 ingroup taxa plus 22 outgroup sequences were constructed. Then, separate analyses were carried out with all representatives of one phylum each plus the 40 nematode sequences. The three sequences of *Priapulus caudatus* Lamarck, 1816, were virtually identical and the four *Gordius* species were also very similar. From both of these genera only one sequence each was selected for any further analyses. A maximum of two representatives each were selected from all six outgroup phyla to give a total of ten outgroup sequences whose relationships to the nematodes were reconstructed (see Table A2.2, file <TableA2_2.rtf> on CD). In the following analyses representatives of two to five of the six phyla were selected and analysed with the ingroup in all possible combinations (Table A3 in Appendix III).

No arthropod sequences were included as outgroups because the choice is so large that there was the danger of effectively choosing a random sequence and introducing long-branch attraction without adding any value to the analysis. In order to test the effects of a distantly related outgroup, the sequence of *Sagitta crassa* Tokioka, 1938, was shifted by two positions in the alignment making it a random outgroup. The resulting topology was not markedly different to any of the other topologies found (Figure not shown).

All topologies were compared to (a) the ingroup topology analysed on its own, and (b) the outgroup topology using all outgroup taxa. A large number of trees were reconstructed (see Table A3), and therefore only an overview of the variation in the recovered topologies was taken for both in- and outgroup taxa. Instead of comparing the recovery of all individual nodes, the number of most parsimonious trees reconstructed in different analyses was analysed (Table A3). First, the number of trees recovered under the inclusion of a certain phylum were counted and compared for the six phyla. Secondly, comparisons were made between including a taxon in a certain combination of outgroups to simply excluding this taxon and leaving all other taxa the same. In this second analysis, the number of instances of an increase or decrease in the number of recovered MP trees was recorded. If the exclusion of a taxon increased the number of trees this taxon was likely to be critical to the reconstruction of the phylogeny. If it reduced the number of trees it was likely to be a problematic taxon introducing many homoplasies (Kitching *et al.*, 1998).

For the actual reconstruction of the phylogeny of the phylum Nematoda all possible outgroup phyla were used in order to reduce the effects of long-branch attraction.

However, those sequences that were virtually the same were omitted apart from one representative leaving 16 outgroup sequences for all further analyses.

2.3.1.4 Phylogenetic analysis of molecular sequences

Table A2.1 in Appendix II (file <TableA2_1.rtf> on CD) gives the source for each sequence used in this study by either GENBANK accession number and/or name of the Principal Investigator of any sequences used in this study.

An initial analysis was conducted with 278 sequences including 16 outgroup taxa. The remaining 6 outgroup sequences were so similar to other outgroups that they did not include any more information. No characters were excluded for these analyses. The heuristic parsimony search was run with simple addition sequence of taxa and the tree limit was set to 500. A second heuristic search was conducted under the LogDet distance criterion with the neighbour-joining tree as starting tree. Only one shortest tree was found. In both analyses tree-bisecting-and-regrafting (TBR) was used as the branch-swapping algorithm. The positions of species of the same genus were analysed and if they formed one clade in all trees only one sequence per genus (the one with the shortest branch length) was retained in further analyses. Table A2.2 (file <TableA2_2.rtf> on CD) lists all sequences used with information on which sequences were included in each analysis. 48 nematode sequences were excluded from the initial analysis for all subsequent analyses. Table 2.4 summarises all information on all analyses run under the parsimony, LogDet distance and maximum likelihood distance criteria.

Four different character sets were used for parsimony analyses: all 1769 characters, excluding all positions with gaps (690 characters remaining), excluding all uninformative characters (constant characters and apomorphies, 989 characters remaining) and excluding all positions with gaps and uninformative characters (393 characters remaining). In all cases, the sequences were added by stepwise addition, and 10 replicates of random sequence addition were performed. Branch swapping was performed under TBR. The number of rearrangements per replicate was limited to 2×10^9 because a search without any limitations proved too time-consuming. Multistate taxa were interpreted as uncertainties. The limit was hit in each replicate and in the analysis with all characters the shortest tree was found only in the tenth replicate; therefore, a new analysis was run with 20 random addition replicates and a shorter tree was found in the 15^{th} replicate.

All four parsimony analyses were repeated with characters weighted under the rescaled consistency index. The most parsimonious trees from the unweighted analysis

were used as first reference for the weighting scheme. After a first heuristic search under the same settings as before the characters were weighted on the resulting most parsimonious trees of the weighted search. The weighting process was repeated until no shorter trees were found. In all four cases the second weighting (on the first set of weighted most parsimonious trees) was the best result. A third weighting run never improved the score. Maximum weights were used when different trees gave varying weights for one character (default option in PAUP*).

The consistency (CI), retention (RI), rescaled consistency (RC) and homoplasy indices (HI) were calculated for all trees under both, unweighted and weighted parsimony using the `describetrees` command in PAUP*. The Goloboff-Fit (G-Fit) was calculated for all parsimony trees using the `pscores` command in PAUP*.

A χ^2 test of base frequency homogeneity with 687 degrees of freedom resulted in insignificant heterogeneity when all characters were included ($\chi^2 = 708.74$, $p = 0.27$) and when positions with gaps were excluded ($\chi^2 = 176.93$, $p = 1.00$) and when both, positions with gaps and parsimony-uninformative characters were excluded ($\chi^2 = 296.31$, $p = 1.00$), but it resulted in significant base frequency heterogeneity when only parsimoniously uninformative sites were excluded ($\chi^2 = 1013.97$, $p = 0.00$). A heuristic search under the LogDet distance criterion was performed with all four data sets as under the parsimony criterion (see above). Although simply excluding uninformative characters is thought to be a drastic measure under the LogDet criterion there was no alternative because of the limitation in the option settings in PAUP*. Again, ten random replicates of stepwise addition of sequences and TBR branch swapping were chosen. The limit of rearrangements per replicate was set to 5×10^6 and hit each time but no shortest tree was found in the tenth replicate. The limit was set higher initially but the score never improved after the first few 10,000 rearrangements.

A full maximum likelihood search was too time-consuming because the data set is very large. Therefore, two alternatives were used. Firstly, a heuristic search under the likelihood distances criterion was performed. The likelihood parameter settings (base frequencies, number of substitution types, substitution rate matrix, Γ rate distribution and shape and proportion of invariant sites) were determined using Modeltest and no further adjustment was made as the search would have taken too long with a data set of this size. When all characters were included the general time reversible model with a Γ rate distribution and proportion of invariant sites was chosen. Under the exclusion of positions with gaps the Tamura-Nei Model (Tamura & Nei, 1993) with a Γ rate distribution and

proportion of invariant sites was chosen. The heuristic search parameter settings were the same as above, but with different limits to the number of rearrangements. When including all characters, the limit was set to 1 x 10^9 and was never hit. Only one shortest tree was found, in replicates 1 and 10 respectively. However, using the same rearrangement settings excluding positions with gaps brought no improvement as under the LogDet criterion and the limit was again reduced to 5 x 10^6. The limit was hit each time and the shortest tree found in replicates 1, 6 and 9. Uninformative characters were not excluded in either likelihood approach because they have considerable effect on the recovery of the correct topology under the likelihood criterion.

The second alternative to a full likelihood search that was performed was a Bayesian analysis. The model was chosen based on the Modeltest results (see above), i.e. using 6 substitution rates (equating to the GTR model) and assuming a proportion of invariant sites and a Γ distribution for the remaining sites. The priors were set to the default uniform values, i.e. the substitution rate matrix parameters between 0.00 and 50.00, dirichlet state frequencies, a Γ shape parameter between 0.00 and 100.00, the proportion of invariable sites between 0.00 and 1.00, unconstrained branch lengths varying between 0.0 and 10.0 and all topologies were assumed equally probable. Four rate categories were used for the Γ distribution, and five categories for the β distribution. The analysis was performed again including all characters and excluding positions with gaps. An initial analysis was conducted for both character sets running 4 million generations and sampling trees every 1000 generations. All parameters (log likelihood, tree length, substitution rates, base frequencies, the proportion of invariant sites and the α parameter of the Γ distribution) were plotted against the number of generations to ensure the burn-in was set appropriately. The burn-in value for the full analyses was set to 250 for the inclusion of all characters and 150 for the exclusion of positions with gaps which equates to the nearest 50 trees above the level at which the log likelihood reaches a stable value. The total number of generations to be run was set ten times higher than the burn-in value, i.e. 2.5 million generations when all characters were included and 1.5 million generations when positions with gaps were excluded. Four parallel chains were used in all analyses. Each of the two data sets were analysed five times. After all five replicate runs were completed the behaviour of all parameters were plotted against the number of generation to assess the smoothness of the runs. The run(s) with the best parameter behaviour was selected for the interpretation.

Table 2.4: Details of phylogenetic reconstruction analyses for the molecular characters alone; ntax = no. of taxa included, pars = parsimony, wpars = weighted parsimony, LogDet distances, ML Dist = maximum likelihood distances, bst = bootstrap analysis, excl = character set excluded (gapped = positions with gaps, uninf = uninformative sites), nchar = no. of characters remaining, nreps = no. of random addition sequence replicates, rearrlimit = rearrangement limit per replicate, maxtrees = maximum tree limit, no. trees = no. of shortest trees retained, tree length: no. of steps for parsimony, genetic distance for LogDet or ML distances, limit hit: either rearrangement or tree limit, time: CPU time required for analysis in days:hours:minutes

ntax	Criterion	excl	nchar	nreps	rearrlimit	maxtrees	No. Trees	Tree Length	Shortest tree in replicate	Limit hit	Time (dd:hh:mm)
278	pars	none	1769	1	none	500	500	17662	-	yes	00:00:49
278	LogDet	none	1769	1	none	500	1	11.96152	-	no	00:00:47
230	pars	none	1769	20	$2x10^9$	none	2825	16563	15	in 19	01:01:32
230	pars	gapped	690	10	$2x10^9$	none	7925	4148	1, 5	all 10	00:09:16
230	pars	uninf	989	10	$2x10^9$	none	5391	15993	4, 8	all 10	00:10:59
230	pars	gapped uninf	393	10	$2x10^9$	none	17472	4017	5, 9	all 10	00:07:37
230	wpars	none	1769	10	none	none	5	1552.68198	6	no	00:00:32
230	wpars	gapped	690	10	$2x10^8$	none	1546	606.18331	10	all 10	00:00:40
230	wpars	uninf	989	10	$2x10^8$	none	3	1278.63349	6	no	00:00:14
230	wpars	gapped uninf	393	10	$2x10^8$	none	976	476.68411	2	all 10	00:00:38
230	LogDet	none	1769	10	$5x10^6$	none	1	11.03058	7	no	00:02:51
230	LogDet	gapped	690	10	$5x10^6$	none	278	5.68628	2	all 10	00:13:05
230	LogDet	uninf	989	10	$5x10^6$	none	1	16.81324	1,2,4,8,10	no	00:00:58
230	LogDet	gapped uninf	393	10	$5x10^6$	none	471	10.05045	4, 9	all 10	00:13:04
230	ML Dist	none	1769	10	$1x10^9$	none	1	14.14439	1, 10	no	00:03:32
230	ML Dist	gapped	690	10	$1x10^9$	none	720	6.59668	12, 6, 9	all 10	00:13:12
230	pars bst	none	1769	10	none	100	97477			yes	11:05:20
230	pars bst	gapped	690	10	none	100	100000			yes	05:21:13
230	pars bst	uninf	989	10	none	100	97573			yes	09:10:12
230	pars bst	gapped uninf	393	10	none	100	100000			yes	05:13:40
230	wpars bst	none	1769	1	none	100	53133			yes	02:18:40
230	wpars bst	gapped	690	1	none	100	100000			yes	05:18:02
230	wpars bst	uninf	989	1	none	100	42862			yes	07:20:19
230	wpars bst	gapped uninf	393	1	none	100	100000			yes	06:13:20
230	LogDet bst	none	1769	1	none	100	1128			no	03:16:46
230	LogDet bst	gapped	690	1	500000	100	99723			yes	06:01:41
230	LogDet bst	uninf	989	1	none	100	1119			no	03:15:43
230	LogDet bst	gapped uninf	393	1	500000	100	99950			yes	06:03:30
230	ML Dist bst	none	1769	1	none	100	1245			no	04:06:36
230	ML Dist bst	gapped	690	1	500000	100	99272			yes	06:04:53

Additionally, bootstrap analyses were performed for all four data sets under the parsimony, weighted parsimony and LogDet criteria and under the likelihood distance criterion including all characters and excluding positions with gaps. For the weighted parsimony criterion, characters were weighted on the best trees from the original heuristic search under weighted parsimony. In all bootstrap analyses 1000 replicates were performed under heuristic search criteria. Ten random replicates of stepwise sequence additions were performed under the unweighted parsimony criterion and one random replicate for all other bootstrap analyses to save time. TBR branch swapping was chosen again for all bootstrap analyses. 100 trees were saved from each bootstrap replicate. The bootstrap proportions were then added to the consensus trees of the full heuristic analysis.

For all straightforward heuristic searches strict consensus trees of all saved trees were constructed. Strict consensus was chosen because it only contains relationships that are found in every fundamental tree. PAUP* and MrBayes construct majority-rule consensus trees from the fundamental trees found in bootstrap and Bayesian analyses respectively.

2.3.2 Analysis of morphological characters

2.3.2.1 Data gathering

Morphological character analysis was conducted for 167 nematode species for which 18S sequences were available and the specimen was identified to the species level. Morphological character data were obtained from specimens in collections and descriptions in the literature.

Drawings were made of all species that could be found in the Freeliving Nematodes Collection of the Natural History Museum (NHM), London, and from a selection of specimens available in the nematode collections from the Parasitic Worms Collection at the NHM and at Rothamstead Research, Harpenden, Hertfordshire. Not all species available in the collections at Rothamstead Research and the Parasitic Worm Collection (NHM) could be drawn because of time restrictions. Some species investigated were not available in UK collections and therefore no drawing was made. Character acquisition for those species without drawings was based solely on data in the literature. For all other species, literature data were used in addition to the drawings of collection specimens. Information from the literature was especially valuable for those species that exhibit habitat related morphological variations and for those characters that are not preserved on slides but visible on fresh specimens.

References for the literature search were mostly obtained from 'The Bremerhaven checklist of aquatic nematodes' (Gerlach & Riemann, 1973 & 1974), 'Nomenclatorial compilation of plant and soil nematodes' (Tarjan & Hopper, 1974) and the Host-Parasite Catalogue of the Parasitic Worm Section (NHM). Additional references have been found on the Web of Science and in general internet searches. The full list of references that were used in the data acquisition for each species is found in Table A4.3 in Appendix IV (file <TableA4_3.rtf> on CD) and the full bibliography in Table A4.4 (file <TableA4_4.rtf> on CD).

2.3.2.2 Justification for the coding of each character

The author would like to emphasise that the current set of characters and their states is based entirely on her own interpretations and that she believes that there is a considerable amount of scope for discussion. My criticism about the interpretation by previous authors in the Introduction may, therefore, seem patronising in the light of my own subjective homology assessment. However, there is one major difference between previous approaches using classic cladistics and the current study. In the classic studies, the author had to make a conscious decision of the evolutionary history of each character whilst in the current analysis this step is left to a clearly defined mathematical algorithm. The computer model, rather than the investigator, defines the plesiomorphic and apomorphic states of each character. This removes an important source of subjectivity.

The characters for this analysis were selected from those used by Lorenzen (1981, 1994), Andrássy (1976) and other classic papers on nematode systematics and taxonomy. However, no morphometric information was used, as this would have exceeded the timeframe of this work.

Characters relating to the buccal cavity morphology have been excluded from this study because homology assessments for the different parts of the buccal cavity of the different groups of nematodes are in their infancy. The assessment is made even more difficult by the high potential of convergent evolution as a result of similar feeding mechanisms in distantly related groups of nematodes. Therefore, buccal cavity characters are highly susceptible to homoplasy. However, a discussion of the current knowledge of the buccal cavity homologies is presented in Section 9.2, at the end of this thesis.

Since morphological character coding without a defined genetic basis for a character is intrinsically subjective, special care has to be taken in the definition of homologous characters. All potential characters were subjected to a rigorous analysis of homology. A

suite of potentially homologous character states was compared across the whole phylum. If it seemed very likely that a variety of states were derived from a common, ancestral state, the character was coded as a multistate character. If there was any doubt over the homology of different states, the suite of character states was reduced to a set of binary characters. This way, many traditional multistate characters were reduced to binary characters. As explained in the Introduction, binary character coding is more likely to avoid problems with homoplasy than complex multistate characters. A full list of all morphological characters is given in Appendix IV, Table A4.1. The character matrix can be found in Table A4.2 (file <TableA4_2.xls> on CD) and the NEXUS file in <morphology.nex> (on CD). The following section provides a discussion of each character included in the morphological analysis. The numbers in square brackets refer to the figures in the plates in the back of this thesis and depict schematic drawings of many characters.

2.3.2.3 Descriptions of morphological characters

1 – 4, 7, *cephalic sensilla* [1-4, 18, 37]: Cephalic sensory organs can either be papilliform (#1A, 2A, 3A) or setiform (#1B, 2B, 3B) and are arranged in a particular pattern; if the full set of sensilla is present there are six labial, six inner cephalic and four outer cephalic sensilla. Each circle has two subdorsal and two subventral sensilla and the inner two circles have also one lateral sensillum on each side. Alternatively, the three circles are simply called the first (R1), second (R2) and third circle (R3) (counting from the mouth opening to the neck) or the inner and outer labial sensilla and cephalic sensilla. The three circles are either at different levels (#4A) or the six inner and four outer cephalic sensilla are found at the same level (#4C). In the latter case they usually differ in size and are therefore distinguishable (#7). In some cases the inner two circles are found at the same level (#4B). In the Rhabditida often not all papillae are observed for a species or only an opening in the prestome (anterior mouth end) can be seen instead of labial sensilla (Tylenchomorpha partim) (#1C). If the sensory organs are setiform the sensilla in the two cephalic circles often vary in length (#7). In few exceptions the lateral sensilla of the inner cephalic sensilla are shorter than the submedian sensilla (#8). In some species only the nerve ends of the sensilla but not the organs themselves can be observed (#2C, 3C).

5 – 6, *jointed nature of cephalic sensilla* [2]: In many aquatic species the cephalic sensilla are composed of two to four segments.

9 – 11, *additional cephalic setae and subcephalic and cervical setae* [4]: Many aquatic and some freeliving terrestrial species have additional setae along the anterior end of the body that are distinct from the remaining body (somatic) setae (#161). If these setae

are found at the same level as the outer sensilla circle and appear similar in appearance to the cephalic setae, they are accounted for as *additional cephalic setae* (#9). If the additional setae are found distinctly behind the cephalic setae (often near the posteriorly placed amphids), they are called *subcephalic setae* (#10). If setae are found all along the oesophageal region they are termed *cervical setae* (#11).

12, *dereids* [5]: Many Rhabditida and Plectida possess a pair of lateral papillae at around the level of the nerve ring. They are either named *cervical papillae* or *dereids* depending on the author.

13, 122/123, 154, *alae* [5, 53-54, 67]: Lateral alae along parts of the body are found in many Rhabditida. These can be either restricted to the oesophageal region (*cervical alae*, #13), they may run along most of the length of the body (*lateral field with alae*, #154) or they are found in the anal region of the male where it is homologous to the *bursa* (#122/123). In all cases, lateral alae are extensions of the cuticle.

14 – 17, *amphids* [1-3, 6, 11, 17, 29, 37]: The amphids are anterio-lateral sensory organs. All Rhabditida have small, *pore-shaped* (#14B) amphids situated very close to the mouth opening (*labial*, #15A). In the other orders, the amphids are larger and found either *within* or *behind the outer cephalic sensilla circle* (#15B-D). The amphids can have a round or slit-like opening with a large inner pocket (*pocket-shaped*, #14A) or they are essentially *round* with no large internal innervations. In the latter case, they can form a simple circle (#14D) or a *ventral or dorsal spiral* (#14C, D). The spiral can be simple or multispiral. In other forms the spiral has been extended into an *elongated loop* (#14E) or reduced to a simple *slit* (#14F). In some species the amphids are slightly offset to the dorsal side of the body (#16) and in many aquatic species the amphids of the male are larger than those of the female (relative to the body diameter) (#17).

18, 19, *cephalic helmet* [7-8]: Some aquatic nematodes possess a cephalic helmet; however, this is derived from different organs in the Enoplea and Chromadorea. In the Enoplea the helmet is formed by the inner cuticle (#18) while in the Desmodoridae (Chromadorea) it is made out of a specific material giving the head the appearance of a blocked structure (#19).

20, *cephalic framework* [9, 31]: The Tylenchomorpha are missing lips and a profound buccal cavity, instead they possess a stomatostyle (hollow spear made of the buccal cavity walls) and a strong cuticularisation of the inner anterior end. This cephalic framework is hexaradiate in accordance with the hexagonal bodyplan of all nematodes.

The six arms or blades are situated one each mid-dorsally and mid-ventrally and two each sub-dorsally and sub-ventrally.

21, *cephalic cuticular inflations* [10]: Some Plectida have cuticular inflations laterally in the head region that are different to the alae described for #13.

22, *cephalic organs*: Some Enoploidea have additional organs in the head region whose functions are yet unknown.

23, *lips* [11-22, 25, 28-29, 38]: Nematodes exhibit a variety of lip shapes and replacements for well-formed lips. The basic structure is thought to be six clearly distinguishable lips (#23A, B) [14, 21, 28, 38]. Three (#23 D, E, G, K, M, P, Q) [12-13, 18] and two (#23H) [16] lips probably arose by merging of two or three adjacent lips. If all lips become merged together, they can form a variety of low lip regions and labial discs (#23 F, I, J, L) [20, 22, 29]. Finally, there are a number of genera in which no lips are observed at all. The amalgamation and reduction of the lips has probably occurred several times as there are different types of merged structures; therefore, different types of merged structures are coded as separate states. The complete loss of lips has probably also taken place several times and is therefore coded as missing data (-).

24, *cephalic probolae* [23]: Cephalobidae have exquisite tentacular structures in the head region. Some authors have homologised them with the lips but they are probably additional structures. They are not homologous to the cephalic sensory organs (#1-8) because at least the papillae of the inner and outer cephalic circles are present.

25, *striations on lips* [24-25]: Some species have longitudinal or transverse markings on the inside of the lips that are an extension of the buccal cavity structures.

26, *longitudinal striations in vestibule (rugae)* [15]: The cheilorhabdia (anteriormost buccal cavity cuticularisations) of many aquatic Chromadoria often form longitudinal pleated structures that look like a crown, called rugae.

27, *transverse striations in vestibule* [14]: The anterior end of the buccal cavity contains transverse striations in some Mononchida.

28, 29, *cephalic sutures* [9, 11, 16, 21b, 26]: Many species possess distinct sutures on the head, either right behind the lips (#28) or some distance further back roughly level with the posterior end of the buccal cavity. In Enoplea with large pocket-shaped amphids the opening of the amphids can be mistaken for a suture if the worm lies in dorso-ventral orientation.

30, *position of the mouth* [26-27]: In most taxa the buccal cavity and mouth opening lie terminally at the anterior end. However, in some taxa the head region is tilted to one side because of the ecology of the animals.

31, *sexual dimorphism of stylet development*: In many Tylenchomorpha only the females have a well-developed stylet while that of the males and juveniles is reduced or even absent.

32, *sexual dimorphism in buccal morphology* [32]: Female *Aduncospiculum halicti* Giblin & Kaya, 1984, possess a buccal cavity dimorphism dependent on the population. They have either stenostomatous or eurostomatous buccal cavities whilst all males have stenostomatous buccal cavities. The stenostomatous buccal cavity has got one tooth and clearly distinguishable rhabdia while there are three teeth in the eurostomatous buccal cavity and the rhabdia are less distinguishable.

33, *oesophageal musculature* [9, 14-15, 21a, 23-32, 37]: The oesophageal musculature is organised in different ways depending on the nature of the buccal cavity structure. Since the homology of the fine structure of most buccal cavities is poorly understood only rough distinctions can be made. The Tylenchomorpha all have a stomatostyle instead of a buccal cavity and the oesophageal tissue commences at the posterior end of the spear (#33G) [31]. A similar situation is found in some Dorylaimia where the actual oesophageal tissue commences at the posterior end of the odontostyle and a different tissue (muscle) surrounds the spear to its anterior end (#33B) [30]. The Diphtherophoroidea also possess a spear but they lack distinct oesophageal tissue in the anterior region altogether. The spear and buccal cavity are free (#33E) [29]. In many Rhabditina and all Cephalobidae the oesophageal tissue reaches up alongside the buccal cavity. This is called the *pharyngeal sleeve*, but investigations into its ultrastructure have shown that the marginal cells are not homologous between the two groups (#33 D & F) [21a, 23, 26]. In the remaining taxa the buccal cavity carries a variety of dentate armature and the oesophageal tissue reaches to varying levels anteriorly. Few ultrastructural investigations have so far been undertaken on these groups and therefore no detailed homology assessment is possible. However, all these taxa have in common that a distinct buccal cavity is present and that the oesophageal tissue reached at least its posterior end and might surround it up to its anterior end (#33A) [14-15, 24-26, 32, 37].

34, *cuticularisation of the oesophageal lumen*: In most taxa the wall of the oesophageal lumen is cuticularised throughout its lengths, but there are a few exceptions where there is no distinct cuticularisation in parts of the oesophageal tube. For many taxa

the level of cuticularisation could not be determined without investigating actual specimens or proper photographs.

35 – 37, *cross-section of the oesophagus* [35]: In cross-section the lumen of the oesophagus can either be simply round or triradiate or triangular (both *tripartite*). In some species the shape might change from the anterior to the posterior end of the oesophagus. Therefore this character is split into three characters.

NB: Nomenclature of the oesophagus (esophagus): There are many different terms for the different parts of the oesophagus because of variation in its structures. The definition used in this thesis is as following: The *procorpus* starts at the bottom of the buccal cavity and reaches to the median bulb (if this is present). The median bulb or the equivalent region is the *metacorpus* and the *postcorpus* reaches from the end of metacorpus to the end or the oesophagus including the posterior bulb if it is present. Many species lack a median bulb but display a distinct constriction near the middle of the oesophagus. This region is called the *isthmus* (#38) and distinguishes the meta- and postcorpus. According to Chitwood & Chitwood (1950) the radial and marginal cells also change with these regions and therefore it is possible to distinguish the three regions even in those taxa that have a uniform oesophagus.

39, 40, 43, *oesophageal musculature*: Each of the three parts of the oesophagus are either surrounded by muscle cells (#39A, 40A, 43A) or cells from the oesophageal glands (#40B, 43B) (see #53). In some taxa there is little tissue around the inner oesophageal tube at all (#39B, 40C).

41 – 42, 44 – 48, *oesophageal bulbs* [31-34, 38, 40]: If a median (#41) or posterior bulb (#44) is present, this may contain a valvular apparatus (#42, 45) [31-32] whose sucking action aids the movement of food from the buccal cavity to the intestine and avoids regurgitation. The valvular apparatus is present either in the median or the posterior bulb; it is present in the posterior bulb if both bulbs are present (in some Rhabditina). Many marine Chromadorea possess plasmic interruptions in the posterior bulb that can lead to the duplication of the inner, cuticularised chamber (#46) [34]. In some Plectidae and Cephalobidae the inner chamber of the posterior bulb contains strong cuticularisations that resemble teeth (#47). Many Rhabditidae possess a specifically cuticularised valve in the posterior bulb that is termed the *double haustrulum* (#48) [33]. In some species no posterior bulb but a distinct swelling can be observed (#44C) [38].

49, *oesophageal tubes* [35]: Oesophageal tubes are found if the distal ends of the radii of a triradiate oesophagus open up to form another, small lumen. For many taxa no data were available because oesophageal tubes can only be observed in cross-section.

50, *stichosome* [36]: The stichosome is a collection of cells (stichocytes) on the outside of the oesophagus without any visible connection to the oesophageal lumen. They are thought to play a role in osmotic regulation and nutrition and may have evolved from the oesophageal glands. In Mermithida the stichosome and oesophagus are embedded in the trophosome (#69).

51, *sexual dimorphism of the oesophagus*: In many Tylenchina only the females have a well-developed median bulb while that of the males and juveniles is reduced or even absent.

52, *sexual dimorphism of the buccal cavity* [37]: In *Calyptronema maxweberi* de Man, 1922, the buccal cavity of the female is well developed, contains teeth and is heavily cuticularised whilst the buccal cavity of the male is reduced to a small, unarmed chamber.

53, *number of oesophageal glands* [38]: All nematodes probably have oesophageal glands even if they have not yet been observed in every species. Most commonly there are three glands, one dorsal and two subventral ones, their nuclei are found in the posterior end of the oesophagus, in the bulb if one is present (#53A). In a number of Dorylaimia and some other taxa a second pair of subventral glands has been observed (#53B) and in the Oncholaimidae the subventral glands are dinuclear (#53C). There is the suggestion that in fact there are always five oesophageal glands but that the second pair of subventral glands has not yet been found in many taxa. In the Spiruromorpha the oesophageal glands are polynuclear giving the whole oesophagus a glandular nature (#53D).

54, *shape of oesophageal glands* [31]: In many Tylenchomorpha the three glands are not contained in the oesophageal tissue but extend freely in the body cavity over the intestine.

55, 56, *position of oesophageal gland openings* [31, 38]: The opening positions of the oesophageal glands vary extensively between different taxa. The dorsal gland always opens further forward than the subventral glands. All glands may open between the posterior end of the oesophagus and the bottom of the buccal cavity or even through the teeth apart from the second pair of subventral glands, which always opens furthest posteriorly in the oesophagus. Homology assessment for the opening position is very

difficult and has been examined in relation to the arrangement of the radial and marginal cells (Chitwood & Chitwood, 1950).

57, *excretory system* [32, 37, 39-40]: There are two basic types of excretory systems in nematodes; all aquatic nematodes have a single, usually ventrally situated gland cell with a more or less long excretory duct leading to a ventrally placed excretory pore (unless no gland is observed at all) (#57A). In many Enoplea and in particularly in most Dorylaimia no excretory system has yet been discovered. In all Rhabditida where an excretory system has been seen it consists of lateral canals that join together before leading to the ventrally situated excretory pore (#57B). In contrast to the aquatic system the excretory duct is cuticularised. This form can be reduced to a single lateral canal with potentially a small gland cell near the junction with the excretory duct (#57E). In other Rhabditida two gland cells are joined to the bridging duct between the lateral canals and the excretory duct (#57D). In the Plectida there are usually one ventral and two subdorsal glands with a cuticularised excretory duct but no lateral canals (#57C).

58, *position of single excretory gland* [37, 40]: If there is a single gland, this may be found either within the oesophageal region (#58A), near the cardia (#58B) or extend well into the intestinal region (#58C).

59, *lateral or central position of excretory glands* [37, 40]: Whilst the single gland is situated ventrally in most taxa (#59A), in some it is found sub-ventrally (#59B) or displaced entirely to the right (#59C) or left side of the intestine (#59D). The latter situation may be found if the anterior gonad reaches far anterior almost to the posterior end of the oesophagus.

60, 61, *position of the opening of the excretory pore* [37, 40]: The excretory pore is most commonly found in the oesophageal region, and here either anterior to (#60B), posterior to (#60D) or near the level of the nerve ring (#60C). In some taxa the pore is situated in the head region between the lips (#60A) or in the intestinal region (not if only a single gland) (#60E, F). In some species it is not certain whether the position anterior or posterior to the nerve ring is constant. Therefore, an alternative coding with just three states, opening in head region (#61A), oesophageal region (#61B) or intestinal region (#61C), has also been tested.

62, *hemizonoid* [31]: The hemizonoid is only found in the Rhabditida and some Dorylaimida. It is usually situated near the nerve ring and the excretory pore but often it is not observed at all. The hemizonion is observed even more rarely than the hemizonoid and therefore the hemizonion is excluded from this analysis.

63 – 67, *cardia* [32-33, 38, 40, 42]: The cardia (or oesophageal-intestinal junction) is usually found protruding from the oesophagus into the intestine. It can be so small that it is not observable (#63B). Similar to the oesophagus the lumen of the cardia can have different shapes in cross-section (#64). In some Ascaridoidea a long appendage is observed posterior to the cardia (#65A). In a number of taxa where no valvular apparatus is found in the posterior part of the oesophagus there might be one in the cardia (#66). In some taxa distinct glands have been observed adjacent to the cardia (#67). There is a possibility that these glands have been overlooked in many species.

68, *symbiotic gut bacteria*: Some nematodes contain symbiotic bacteria in their intestine. In this data set the two species with bacteria (*Steinernema carpocapsae* Weiser, 1955, *Heterorhabditis bacteriophora* Poinar, 1975) are both entomopathogenic nematodes (endoparasitic parasites of insects) and possess the bacteria to digest their prey.

69, *trophosome* [36]: The Mermithida lack a distinct intestine and instead possess a specialised organ, the *trophosome*. The trophosome engulfs the oesophagus and stichosome (see #50) but there is no obvious connection between these organs.

70, *anterior intestinal caecum* [42]: In a number of Ascaridoidea the intestine continues anterior to the cardia forming an *intestinal caecum*. The caecum can reach alongside the oesophagus as far anterior as the nerve ring.

71, *posterior intestinal caecum* [41]: In some Tylenchomorpha the intestine continues posterior to the anus/cloaca into the tail region.

72, *anus*: Some parasitic species have non-feeding life stages in which the anus is non-functional and may be absent altogether.

73, *pre-rectum* [43]: The pre-rectum is only present in the Dorylaimia. It is separated from the posterior end of the intestine by a constriction in the lumen of the intestine and at its distal end leads into the rectum proper and anus. Some Trichodoridae also have a pre-rectum but it is distinct from that of the Dorylaimia.

74, 75, *lateral position of gonads*: Lorenzen (1981) observed that the position of the gonads relative to the intestine is constant in most species, genera and even families. They may lie either both on the same side or on opposite sides of the intestine; in some taxa they lie ventrally. However, the position is only clearly distinguishable in fresh specimens and in the current analysis it was difficult to find data from older species for which it had not originally been recorded.

76, *number of ovaries* [44-48]: In most species there are either two fully developed ovaries or only one anterior or one posterior ovary. However, in some taxa one ovary may have become reduced to form simply a sac on the opposite side of the vulva. This post- or pre-vulval sac may act as a *receptulum seminis*.

77, *structure of ovaries* [44-48]: Most ovaries are either entirely outstretched (#77A) or reflexed somewhere along its lengths. If the geminal zone of the ovary lies entirely in the reflexed part this is called *antidromously reflexed* (#77B), if it continues into the other part it is called *homodromously reflexed* (#77C). In the Cephalobidae the reflex point is found at the commencement of the oviduct and the spermathecae point forward from this point (#77D). In most Spirurina, which have two posterior ovaries, and in the Tylenchomorpha where the females become obese the ovaries are coiled in the entire body region and may even reach the oesophageal region (for posteriorly situated vulva) (#77E).

78, *structure of tip of ovaries* [44-48]: The end of the ovary can be folded over in some taxa. However, this character is only recorded if the fold is obvious and stable for the species.

79, 80, *position of vulva*: The vulva lies somewhere around the midbody- region to the posterior quarter in most taxa. However, in some taxa it may be found well posterior beyond 80% body length or even be terminal with anterior facing ovaries. On the other extreme it may be found in the oesophageal region with posterior facing ovaries. Two alternative coding strategies have been tested, one that only distinguishes between oesophageal, intestinal and terminal position (#80) and the other one distinguished between more detailed positions in the intestinal region (#79).

81, *demanian organ* [56]: In some Oncholaimidae a special organ, the *demanian organ*, is found either side of the ovary. There is usually a connection between this organ and the ovaries and the intestine. Its function is uncertain. This organ was discovered by de Man and therefore named after him.

82, 83, *reproduction*: Most nematodes reproduce by laying eggs into the environment (soil/sediment or host tissue), the larvae hatch there and grow into adults via four larval states and as many moults (#82A, 83A). However, in a considerable number of taxa the eggs are retained in the mother's body. They may develop until the first larval stage in the uterus, the female still lays the eggs and the larvae hatch just after egg deposition (#82B, 83B). In a further development the larvae hatch in the uterus and the mother releases first-stage juveniles (#82C, 83B). In some plant parasitic taxa the mother never releases the eggs but its body forms a protective cyst in which the young hatch and

develop until the infective juvenile stage (usually first or third). The latter case is not represented in this data set. It is not certain how the two different ovoviviparous states are related and therefore two alternative character codes are tested, once dividing all three alternatives (#82) and then only distinguishing between oviparous and ovoviviparous reproduction (#83).

84, *spermathecae* [45-46]: In a considerable number of Dorylaimia and Tylenchina and some other taxa a cavity between the oviduct and uterus, the *spermathecum*, is specialised as a *receptulum seminis*. It is separated by constrictions and there may even be a sphincter muscle between the oviduct and the spermathecum. If a spermathecum is present the sperm is retained in this cavity and fertilisation of the eggs takes place here. In those taxa that lack a specialised spermathecum fertilisation occurs in the uterus.

85, *ovijector* [48]: In a few taxa a specialised ovijector is present between the uterus and the vulva.

86, *Z-organ*: The *Z-organ* is only found in the uterus of the genus *Xiphinema* Cobb, 1913. Its function is uncertain.

87, *vulval papillae*: In some taxa distinct papillae are found around the vulva. They probably aid in copulation by preventing the two animals to slide apart.

88, *vulval flap* [48]: In some Trichostrongylidae the vulva is covered with a flap that may vary considerably in size within a species.

89, *vulval glands*: In some taxa glands have been observed in the vulva region. These probably secrete a sticky substance that aids in copulation. It is possible that such vulval glands are overlooked in many species.

#90, *number and position of testes* [49, 55]: The testes are always found anterior to the cloaca because the tail commences immediately behind it. However, a testis is effectively facing posterior if it bends backwards at the testis - vas deferens junction.

91, *reflexion of testes* [49, 55]: In most cases the testes lie straight anteriorly or posteriorly, but there are many exceptions where one or both testes are reflexed due to the space available in the body cavity (#91B). In some taxa with two opposed testes the posterior testis lies anteriorly to the testis - vas deferens junction at first before it bends backwards (#91C).

92 – 94, *spicules* [49-55, 57-65]: Almost all nematodes have two, equally sized and shaped spicules. However, in a few taxa only one spicule is present (#92A) [51] or the

spicules are entirely absent (#92C) [65]. In a number of taxa the spicules are unequal (#93B) [74, 76] and/or dissimilar (#94B) [74, 76].

NB: The actual shape of the spicules could not be included because it is so variable and often difficult to describe. Chitwood & Chitwood (1950) already discussed this problem and stated that in most cases instead of a wordy description a detailed drawing is much more informative. This concept has not changed and all nematode descriptions show a figure of the spicules (if a male specimen has been found). Potentially it would have been possible to summarise similar spicule shapes and to code them arbitrarily. However, there was a great chance of introducing homoplasies if the exact nature of the spicules was not known. Therefore, only the relative size and shape differences are included as characters in this study.

#95, *fusion of spicules* [49-55, 57-65]: In most nematodes the two spicules are separate (#95A) but in some taxa they are either fully fused (#95B) [52] or partially fused distally (#95C) [53b]. The latter is commonly observed in the Rhaditidae.

96, *spicular alae* [50]: In some taxa a fine ala (wing) stretches in the spicule curvature from its distal to its proximal end.

97, *spicule sheath* [51]: In *Trichuris muris* Schrank, 1788, the spicule is engulfed within an eversible sheath that is covered in spines whose arrangement is species specific. The spines on the sheath ensure that copulation is not interrupted prematurely.

98 – 100, *gubernaculum* [49, 52-53, 55, 57, 60-61, 63, 65, 74]: Most nematodes have a gubernaculum which acts as a guiding piece for the spicules during copulation. The gubernacula can take a variety of forms and have a number or different pieces. Since it has not yet been established which structures are derived from each other, only the presence or absence of the gubernacula has been coded for this study (#98). In the Dorylaimida, the gubernaculum has been replaced with two lateral accessory pieces (#99) [58] and in the Mononchidae both, gubernacula and two lateral accessory pieces are present [61]. Many nematodes have one or two distinct dorsal apophyses on the gubernaculum. The number of apophyses depends on the type of gubernaculum present and therefore here only the presence or absence has been recorded (#100) [49].

101, *rectal glands* [56]: In many nematodes, and especially in the Rhabditida, which are lacking caudal glands (#176), another set of glands opens into the rectum. Sometimes they are only observed in one sex and they may have been overlooked in some species.

NB: Many species of nematodes possess *supplementary copulatory organs* around the male cloaca. They all aid the attachment of the male and female individual to each other during copulation. They are mainly papilliform or setiform. Whilst there are great differences in the type, position and number of the papillae or setae it is likely that the underlying genetic information comes from the same gene for papillae or setae respectively in the same region because many species have these organs only in a particular part around the anus.

102 – 111, 121, *copulatory papillae* [52-55, 57-62, 74, 76]: Copulatory papillae are found pre-anally (#102-106) [53-55, 57-61, 74, 76], ad-anally (#107) [53-55, 59, 74] or post-anally (#108-110) [53-54, 59, 62, 74, 76]. Most commonly they occur either in one row on the midventral line (#102-104, 108) [55, 57-58, 60-61] or in two parallel subventral rows (#105, 107, 109) [52-54, 58-59, 74]. In some Spirurina the papillae are not arranged in distinct rows but in groups pre- and post-anally (#121) [76]. In the Chromadoridae they form distinct cups that are most certainly homologous (#103) [57]. In a few species these papillae are not restricted to the posterior body but reach into the oesophageal region (#104). In *Nudora bipapillata* Platt, 1973, these papillae are modified to distinct cones (#106) [60]. In some taxa the papillae are also found on the dorsal side (#110). Most of the times the papillae sit in the body cuticle and only protrude a little, however in some taxa they sit on stalks (#111).

112, *copulatory tubules* [61]: In some taxa, tubules reach deep into the underlying tissue and open to the surface through the cuticle. These are usually numerous and quite small but may be rather large and less in number.

113, *shape of the copulatory tubules* [62]: In the Enoploidea the tubules may become very well developed resembling a trumpet in shape.

114, *glands of supplements* [60]: It is generally thought that the papillae and tubules may secrete sticky substances but actual gland openings have only been observed in selected taxa. In some taxa, a hint of a canal has been observed in the underlying tissue and in other taxa it is thought that a nerve may be present.

115 – 118, *copulatory setae* [62-65]: Either additionally or instead of papillae or tubules many nematodes have specialised setae in their preanal (#115-116) or postanal region (#118) or around the anal opening (#117). Pre-anally, these setae may again be found in one ventral row (#115) [62-63, 65] or in two subventral rows (#116) [64]. These setae are distinguishable from simple somatic setae because they either differ in number

(less on the rest of the body or in the female) or shape (stronger or longer than somatic setae).

119, 120, *copulatory thorns* [65]: In *Monoposthia costata* (Bastian, 1865) and *Molgolaimus demani* (de Man 1922) Jensen 1978 the setae are very much enhanced and form strong thorns, both pre-anally (#119) and post-anally (#120).

122, 123, *bursa* [52-55, 59]: As already mentioned under #13, many male nematodes have specialised posterior alae that form a distinct bursa. The bursa varies from a simple cuticular flap either side of the anal region (#122 E, G or H) [55, 59] to distinct vela either side (#122 B, C, or D) [53-54] to the large extravagant loops and rays of the Strongyloidea (#122F) [52]. If a distinct velum is present, this may reach to the tip of the tail (*peloderan bursa*, #122B) [53], or the tail may remain free (*leptoderan bursa*, #122 C or D) [54]. Most peroderan bursae are open at the anterior end (#122D) [53b] but in some taxa the vela from either side join closing the bursa anteriorly (#122C) [53a]. Many bursae contain papillae along its edges [52-54]. It is most likely that these are homologous to the subventral rows of papillae in taxa that lack a bursa and the presence of such papillae has been included in those characters (#105, 107, 109). Since the homology of the different bursae is not entirely certain character 123 provides an alternative test for the simple presence and absence of a bursa.

124, *genital cone*: In some Strongylidae a distinct genital cone is present within the centre of the bursa.

125, *Cuticular patches*: A number of taxa from many groups of nematodes have roughened cuticular patches (*area rugosa*) either anterior or posterior to the cloaca or in both positions. They increase traction between the two bodies during copulation. These patches can be of many different types and homologies may not be certain.

126, *denticular plates*: Some taxa contain specific denticular plates in the caudal region that also increase traction between the two bodies during copulation (see #125).

127, *body shape*: Most nematodes have simple elongated, rod-shaped bodies but the females of some plant-parasitic Tylenchomorpha become saccate or spindle-shaped when mature.

128, *crystalline bodies*: Some nematodes contain crystalline bodies in their body cavity. The exact function of these crystals is unknown.

129 – 133, *transverse cuticular striations* [11, 22, 66a, b, 67]: The cuticle can have a variety of transverse striations. These can be simple annules (#129A, #130A), strong

articulations that resemble segments but are entirely cuticular (#129B, #131A), they can be very faint and frequently overlooked (#129C, #132A) or the cuticle may be entirely devoid of transverse markings (#129D, #133A). It is uncertain if there is a common genetic cause for the different types of transverse markings and therefore this suite of characters has been tested as multistate character (#129) and as a set of binary characters (#130-133).

134, *longitudinal cuticular striations* [66c, 67]: Additionally to the transverse markings the cuticle may also show longitudinal markings. These are quite constant in structure across the phylum but may form well-marked elevations in the lateral fields (see #154).

135 – 137, *cuticular patterns*: The cuticular pattern is often not observed along the entire length of the body. It may be absent on the head (#135B) or tail (#136B) or absent from the mid-body region backwards (#137B).

138, 149, *cuticular inflations* [70]: Some Trichostrongylidae have particular cuticular inflations around the head and neck region (#149A). The inflations are found around all sides the body rather than being restricted to the lateral fields in the case of alae. In some of these species the cuticular patterns are only found on these inflations (#138A).

139, *cuticular spines and scales* [16]: In some species the annules are especially adorned with backwards facing scales and spines.

140 – 141, *sub-cuticular striations*: In some species the sub-cuticle has transverse (#141A) or longitudinal (#140A) striations while the surface appears smooth.

142 – 143, *cuticular or sub-cuticular punctations* [66d]: The cuticle may also be adorned with punctations found either on the surface (#142A) or in the sub-cuticle (#143A).

144 – 146, *cuticular bosses* [68-69]: Some vertebrate-parasitic species have specific cuticular bosses along their body that aid in the attachment to their host tissue or to their partner during copulation. These are found near the vulva (#144A), along the main body (#145A) or on the tail of the female (#146A).

147, *cuticular warts* [27]: The Bunonematidae are characterised by the presence of warts and complicated 'Kriechleisten' on the right side of the body. The 'Kriechleisten' consist of a net-like structure of cuticular material embedded into the cuticle.

148, *Ectosymbionts*: A number of different freeliving nematodes form ecto-symbiotic relationships with bacteria or blue-green algae (Cyanobacteria).

150, *bacillary band* [79]: In the genus *Trichuris* a bacillary band is found along the ventral side of the anterior region of the body. It consists of punctiform projections penetrating the cuticle; they are the ends of small rod-shaped structures originating in the cells of the sub-cuticle. The bacillary band usually reaches the vulva and behind it hypodermal gland pores continue in its place. The band may be homologous to the lateral alae found in many other nematodes.

151, *cuticle with V-like projections* [65]: In the Monoposthiidae the annules may form V-shaped projections in regular, longitudinal rows.

152, *annules with spiral thickenings* [71]: In the genus *Onchocerca* Diesing, 1841, the cuticle is adorned with spiral thickenings instead of regular transverse markings.

153 – 157, *lateral field* [41, 66d, 67]: The lateral field may have a variety of adornments. It may be innervated by longitudinal incisures that are stronger than those found around the rest of the body (#153A). The cuticle of the lateral field may be elevated to form lateral alae (#154A). If alae are present and the transverse striations continue across the lateral field the field is areolated (#155A). The lateral field may also be marked off from the rest of the body by a differentiation in the punctation (#156A) [66d]. In many species the lateral field is not marked off specifically.

158, *hypodermal glands*: Many nematodes have hypodermal glands either along the whole body around all sides, restricted to the lateral field or restricted to particular regions only. However, it has been suggested that all nematodes may have some form of hypodermal glands and that these may simply be under-reported.

159 – 160, *ocelli and pigment spots* [72-73]: Some freeliving nematodes have some form of light sensitive pigments in the oesophageal region. These can be sophisticated ocelli (#159A) with or without lenses or simple accumulations of pigments (#160A).

161, *somatic setae*: Many freeliving nematodes have setae along the main part of their body either scattered around or in regularly arranged rows. However, the Rhabditida are entirely devoid of such organs.

162, *glandular setae*: In some genera of the Desmodoridae the somatic setae are accompanied by secretory gland cells.

163, *metanemes:* Some Enoplida contain special structures in the cuticle, called *metanemes*. They were discovered by Lorenzen (1978) and have so far only been observed in selected species.

164, *phasmids* [41]: *Phasmids* are found in most Rhabditida and Plectida either side of the tail. They are thought to be some form of sensory organs. In some species they are found anterior to the anus or vulva.

165, *prophasmids*: In the Anguinidae a second pair of phasmids, the *prophasmids*, are found in the vulva region.

166, *postdereids*: In a few selected species, structures similar to the dereids are found in the posterior part of the body. They are called the *postdereids*.

167, *caudal pores* [58]: Many Dorylaimia species have a number of pores in the tail region that resemble hypodermal glands.

168, *terminal caudal papillae* [74]: In some Spirurina terminal papillae are found on the tail, which are differentiated from the copulatory papillae.

169 – 170, 172, *tail* [75, 76]: The tail for both sexes can exhibit a variety of forms. In most species the tails of both sexes are very similar but especially in the case where the mail has a bursa the tails shapes between the sexes varies. Variation can also occur in the relative length and thickness of the tail and in some species the form is entirely different. Therefore, the tail shape was coded separately for each sex and an additional character added to include information whether the tail shape is the same or not (#172).

171, *terminal setae* [63]: Many aquatic nematodes have two to four terminal setae on the tip of the tail.

173 – 174, *caudal setae* [62, 65, 78]: In many aquatic species the setae on the tail are distinguishable from the somatic setae and often there are more setae on the tail of the male than that of the female. As it is uncertain whether the difference between the number of setae on the male and female tail was genetically significant two characters were created and tested against each other, one accounting for the difference between the sexes (#173) and the other one just accounting for their presence or absence (#174).

175, *terminal cuticular flaps* [77]: In the genus *Trichinella* Railliet, 1895, terminal cuticular flaps are found on the end of the tail, forming a pseudobursa.

176 – 179, *caudal glands* [61, 63-65]: Aquatic nematodes do not possess phasmids but instead they possess caudal glands in the tail. These glands produce a secretion that aids in the attachment of the tip of the tail to the substratum preventing the animals from being swept away with the currents. Either two or three glands are reported but sometimes the third may simply not be visible and therefore no distinction is made between the

numbers. In the Triplonchida caudal glands are often observed but sometimes only the opening can be distinguished (#176C). There may be one common opening for all glands (#177A) or a separate opening for each gland (#177B). The glands open through a specific cuticularised structure, the spinneret, which may be markedly setoff from the tip of the tail (#178A) [78]. The caudal glands are usually restricted to the tail (#179A) [61, 64, 65] but may extend into the precaudal region, at least in the male (#179B) [64].

180, *reproductive strategy*: Most nematodes are dioecious (amphimictic) reproducing by copulation of male and female individuals (#180A). However, males are often very rare in terrestrial nematodes. In these cases they reproduce either by (protandrous) hermaphroditism (#180B), which occurs in some Rhabditomorpha, or by parthenogenesis (#180C). In both these cases copulation is optional and may occur when a female encounters a male. In hermaphroditism the eggs and sperm are both produced in the female while parthenogenesis does not require sperm at all.

2.3.2.4 Phylogenetic analysis

All analyses of the morphological characters were performed with PAUP*. No outgroups were included in the analysis of morphological characters alone because most characters are nematode specific. All 180 characters were divided into four different sets; the first set of 145 characters was included in all analysis as the author had confidence in the homology assessment. Two other sets contain alternative coding of nine binary (# 61, 80, 83, 123, 130-133, 174) and six multistate characters (# 60, 79, 82, 122, 129, 173) respectively. Multistate character 129 replaces the four binary characters 130-133 coding the cuticular annulations. Initially, all nine binary characters were analysed together with the first set. Then, each binary character was replaced by its alternative multistate coding character in individual analyses. In order to assess which alternative coding was more homologous the alternative topologies of the strict consensus trees were compared for resolution as well as position of problematic taxa (esp. Trichinelloidea being placed with the Spirurina) and the consistency indices of the consensus trees were calculated (Table 2.5). Two binary characters, 83 and 123, were excluded in favour of their multistate partners, 82 and 122, resulting in 154 approved characters for any further analyses.

The final 20 characters were recognised as potentially homoplasious and were tested individually against the 145 characters of the first set and the 'binary' set (before the two characters were exchanged). As for the multistate characters, the consensus trees were compared for resolution and topology and the consistency indices calculated (Table 2.6a). Five 'dubious' characters, 14, 23, 33, 137, 169 showed improvement of the data and were

included in all following analyses. A further 11 characters were tested against the first set including the two multistate characters and five 'dubious' characters but did not show any improvement and were therefore excluded from the analysis (Table 2.6b). Therefore, 159 characters were included in the final character set that was analysed.

Table 2.5: Tree length (TL), consistency index (CI), CI without uninformative characters, retention index (RI), rescaled consistency index (RC), homoplasy index (HI), HI without uninformative characters and Goloboff-Fit (G-Fit) for the exchange of one binary character with its alternative multistate character.

	TL	CI	CI excl uninf	RI	RC	HI	HI excl uninf	G-Fit
Set 1	911	0.2162	0.1978	0.7210	0.1559	0.7838	0.8022	-68
Incl 60	942	0.2123	0.1944	0.7143	0.1517	0.7877	0.8056	-68
Incl 79	945	0.2116	0.1937	0.7155	0.1514	0.7884	0.8063	-67
Incl 82	913	0.2169	0.1984	0.7205	0.1562	0.7831	0.8016	-68
Incl 122	919	0.2209	0.2027	0.7195	0.1589	0.7791	0.7973	-68
Incl 129	885	0.2215	0.2025	0.7230	0.1601	0.7785	0.7975	-68
Incl 173	927	0.2136	0.1954	0.7150	0.1527	0.7864	0.8046	-68

Table 2.6: Tree length (TL), consistency index (CI), CI without uninformative characters, retention index (RI), rescaled consistency index (RC), homoplasy index (HI), HI without uninformative characters and Goloboff-Fit (G-Fit) for the inclusion of individual 'dubious' characters.

a) Tests of all 20 characters

	TL	CI	CI excl uninf	RI	RC	HI	HI excl uninf	G-Fit
Set 1	911	0.2162	0.1978	0.7210	0.1559	0.7838	0.8022	-68
Incl 14	929	0.2185	0.2004	0.7248	0.1584	0.7815	0.7996	-68
Incl 15	940	0.2128	0.1948	0.7205	0.1533	0.7872	0.8052	-68
Incl 23	943	0.2269	0.2093	0.7205	0.1635	0.7731	0.7907	-69
Incl 28	930	0.2129	0.1947	0.7163	0.1525	0.7871	0.8053	-68
Incl 29	936	0.2115	0.1934	0.7155	0.1514	0.7885	0.8066	-68
Incl 33	931	0.2180	0.2000	0.7231	0.1577	0.7820	0.8000	-69
Incl 49	919	0.2155	0.1971	0.7226	0.1557	0.7845	0.8029	-68
Incl 55	926	0.2203	0.2022	0.7265	0.1601	0.7797	0.7978	-68
Incl 56	925	0.2195	0.2013	0.7262	0.1594	0.7805	0.7987	-69
Incl 57	920	0.2185	0.2002	0.7241	0.1582	0.7815	0.7998	-69
Incl 74	932	0.2146	0.1965	0.7186	0.1542	0.7854	0.8035	-68
Incl 75	935	0.2139	0.1958	0.7172	0.1534	0.7861	0.8042	-68
Incl 78	921	0.2150	0.1967	0.7192	0.1546	0.7850	0.8033	-69
Incl 125	923	0.2145	0.1962	0.7180	0.1540	0.7855	0.8038	-68
Incl 135	945	0.2095	0.1916	0.7139	0.1496	0.7905	0.8084	-68
Incl 136	926	0.2138	0.1956	0.7174	0.1534	0.7862	0.8044	-69
Incl 137	912	0.2171	0.1978	0.7210	0.1565	0.7829	0.8022	-69
Incl 169	969	0.2157	0.1983	0.7111	0.1534	0.7843	0.8017	-68
Incl 170	964	0.2189	0.2015	0.7133	0.1561	0.7811	0.7985	-69
Incl 172	934	0.2120	0.1939	0.7198	0.1526	0.7880	0.8061	-69

b) Tests of 11 remaining characters

	TL	CI	CI excl uninf	RI	RC	HI	HI excl uninf	G-Fit
Set 1 +	1050	0.2343	0.2179	0.7151	0.1675	0.7657	0.7821	-69
Incl 29	1073	0.2302	0.2141	0.7109	0.1636	0.7698	0.7859	-69
Incl 49	1057	0.2337	0.2174	0.7170	0.1675	0.7663	0.7826	-69
Incl 55	1065	0.2376	0.2215	0.7203	0.1711	0.7624	0.7785	-69
Incl 56	1066	0.2364	0.2203	0.7193	0.1700	0.7636	0.7797	-69
Incl 57	1059	0.2361	0.2199	0.7180	0.1695	0.7639	0.7801	-70
Incl 74	1071	0.2325	0.2164	0.7130	0.1658	0.7675	0.7836	-70
Incl 75	1072	0.2323	0.2162	0.7124	0.1655	0.7677	0.7838	-69
Incl 78	1062	0.2326	0.2163	0.7128	0.1658	0.7674	0.7837	-69
Incl 125	1063	0.2324	0.2161	0.7121	0.1655	0.7676	0.7839	-70
Incl 136	1066	0.2317	0.2155	0.7115	0.1649	0.7683	0.7845	-69
Incl 170	1097	0.2370	0.2214	0.7103	0.1683	0.7630	0.7786	-69

The 159 remaining characters were analysed under both, unweighted and weighted parsimony. Characters for the weighted analysis were weighted on the resulting trees of the unweighted analysis using maximum values of the rescaled consistency index. After the first round of weighted parsimony analysis, characters were weighted on the resulting trees and the procedure was repeated until no improvement in tree length was recorded. The shortest trees were found in the first round of weighting. In a second analysis, 21 uninformative characters (6, 8, 30, 32, 52, 69, 97, 104, 106, 110, 126, 128, 137, 140, 144 – 147, 150, 152, 175) were excluded and the same analyses were performed as for all characters. The shortest trees were found after two rounds of weighting. Table 2.7 gives an overview of the parameters from each analysis.

Table 2.7: Details of phylogenetic reconstruction analyses for the morphological characters only; pars = parsimony, wpars = weighted parsimony, bst = bootstrap analysis, excl = character set excluded (uninf = uninformative sites), nchar = no. of characters remaining, no. trees = no. of shortest trees retained, tree length: no. of steps, limit hit: either rearrangement or tree limit, time: CPU time required for analysis in days:hours:minutes

Criterion	excl	nchar	No. trees	Tree length	Shortest tree in replicate	Limit hit	Time dd:hh:mm
pars	none	159	9858	1103	8	yes	0:03:23
pars	uninf	138	23707	1029	1, 10	yes	0:02:35
wpars	none	159	586	158.50610	5	yes	0:02:38
wpars	uninf	138	10634	142.69416	10	yes	0:02:48
pars bst	none	159	100000			yes	1:22:47
pars bst	uninf	138	100000			yes	1:22:54
wpars bst	none	159	100000			yes	1:20:30
wpars bst	uninf	138	100000			yes	1:21:47

All analyses were preformed using heuristic searches with 10 replicates of random addition of sequences under TBR branch swapping and 10^9 rearrangements. Multistate taxa were interpreted as uncertainties. Strict consensus trees were constructed for all four analyses and the consistency indices were estimated.

Bootstrap analyses were performed for all four conditions as described above. 1000 replicates were performed using heuristic searches with a single replicate of random addition of sequences under TBR branch swapping and a maximum of 100 trees was saved per replicate. Multistate taxa were interpreted as uncertainties. Majority-rule consensus trees were constructed for all analyses as it is the default in PAUP*.

2.3.3 Phylogenetic analysis of the combined data set

The combined analysis of molecular and morphological characters was performed in PAUP* under the parsimony criterion. Molecular and morphological characters were chosen for the 167 nematode species included the morphological analysis and molecular characters alone for the 16 outgroup taxa from the molecular analysis; gaps were introduced for the morphological characters of the outgroup taxa. The data set originally contained all 1769 molecular characters and the 159 morphological characters used in the final analyses of the morphological data alone. This data set was analysed including all characters, excluding all positions with gaps in the molecular data set, excluding all parsimony-uninformative characters and excluding all positions with gaps in the molecular data set and all parsimony-uninformative characters.

The potential character incongruence between the two data sets (partitions) was established with the 'Incongruence Length Difference Test' (Farris *et al.*, 1995) as it is implemented in PAUP*. 100 random replicates were analysed for each of the four character sets. The heuristic searches were conducted using one replicate of random additions of sequences each and setting the rearrangement limit to 10^9 for each replicate. The rearrangement limit was always hit but more liberal settings would have increased the run time unfeasibly. The searches took between 8 and 68 hours. For all analyses the total tree lengths of the original partitions were significantly shorter ($p = 0.01 < 0.05$) than any tree lengths for random partitions. Therefore, there is no significant incongruence between the characters from the two data sets and the performance of a combined analysis was justified.

Characters were first treated as unweighted and subsequently weighted on the maximum value of their rescaled consistency index using all fundamental trees from the original, unweighted analysis as staring point. Characters were re-weighted until no shorter trees could be found. The shortest trees were found after three rounds of re-weighting when uninformative characters were excluded and after two rounds for the three other characters sets. Table 2.8 gives an overview of the parameters from each analysis.

All analyses were performed using heuristic searches with 10 replicates of random addition of sequences under TBR branch swapping and 10^9 rearrangements. Multistate taxa were interpreted as uncertainties. Strict consensus trees were constructed for all eight analyses and the consistency indices were estimated.

Bootstrap analyses were performed for all eight conditions as described above. 1000 replicates were performed using heuristic searches with a single replicate of random addition of sequences under TBR branch swapping and a maximum of 100 trees was saved per replicate. Multistate taxa were interpreted as uncertainties. Majority-rule consensus trees were constructed for all analyses as it is the default in PAUP*.

Table 2.8: Details of phylogenetic reconstruction analyses for the combined analyses of molecular and morphological characters; ntax = no. of taxa included, pars = parsimony, wpars = weighted parsimony, bst = bootstrap analysis, excl = character set excluded (gapped = positions with gaps, uninf = uninformative sites), nchar = no. of characters remaining, no. trees = no. of shortest trees retained, tree length: no. of steps, limit hit: either rearrangement or tree limit, time: CPU time required for analysis in days:hours:minutes

Criterion	excl	nchar	No. trees	Tree length	Shortest tree in replicate	Limit hit	Time dd:hh:mm
pars	none	1928	1294	14331	2	8x, not in 9, 10	0:06:11
pars	gapped	753	8150	4130	1, 8	Yes	0:02:48
pars	uninf	1068	2224	14035	1	7x, not in 2, 6, 7	0:05:14
pars	gapped uninf	440	2770	3992	2	yes	0:02:34
wpars	none	1928	1	1624.80339	6	no	0:00:05
wpars	gapped	753	21	647.46516	1	no	0:00:06
wpars	uninf	1068	1	1326.53221	1	no	0:00:03
wpars	gapped uninf	440	21	506.75645	3 (island hit 4x)	no	0:00:05
pars bst	none	1928	77549			yes	6:14:46
pars bst	gapped	753	100000			yes	2:03:48
pars bst	uninf	1068	74068			yes	3:09:55
pars bst	gapped uninf	440	99949			yes	1:22:04
wpars bst	none	1928	7811			in 6 replicates	1:11:54
wpars bst	gapped	753	98838			yes	3:08:36
wpars bst	uninf	1068	8100			in 14 replicates	0:21:41
wpars bst	gapped uninf	440	98998			yes	2:08:40

2.3.4 Consistency analysis of all morphological characters

In order to establish the quality of the morphological character coding, all 180 characters were mapped onto two phylogenetic reconstructions under weighted parsimony; one based on the morphological characters alone and one based on all characters excluding positions with gaps in the DNA and uninformative characters. The potential evolutionary history of each character was analysed based on the most parsimonious character state

changes. Assumptions were made about the homology of the individual states based on the amount of change that was needed to explain the reconstructions. In some cases alternative coding strategies are suggested.

Additionally, the consistency indices (CIs) and character weights for each character were analysed using MINITAB version 12. The Anderson-Darling Test of Normality showed that the values of the consistency index and the weights were not normally distributed (all p = 0). However, Levene's Test of Homogeneity of Variance for non-normally distributed but continuous data showed that the shape of the distribution for the two sets of consistency indices and weights are equal (CIs: test statistic: 0.675, $p = 0.412 > 0.05$; weights: test statistic: 1.152, $p = 0.284 > 0.05$). Therefore, the Mann-Whitney U-Test of the median was performed to test whether there is a significant difference between the medians of the consistency indices and weights between the two sets of data. In both cases, the test statistics were greater than the critical value and therefore there was no significant difference between the medians of the consistency indices and weights between the two different treatments (CIs: W = 31738.0, critical value = 0.4430; weights: W = 31470.5, critical value = 0.2986). However, if the non-normality of the data was ignored, a Paired T-Test resulted in significantly higher means for the CIs and weights when the characters were mapped on the morphological tree (CIs: T-Value = -3.26, $p = 0.001 < 0.01$; weights: T-Value = -3.20, $p = 0.002 < 0.01$).

3. Time Series Experiments

In order to establish how quickly the DNA degrades while the animals are desiccated in glycerol and mounted on slides for identification, a time series experiment was performed. All experiments were conducted with the same strain of *Caenorhabditis elegans* (Bristol N2). In the first series animals were desiccated in glycerol for up to 56 hours without prior fixation in ethanol. The animals were placed alive straight from the agar plate into the glycerol. In a second series animals were either placed into glycerol straight from the agar plate or fixed in 99.7% ethanol for two days and then kept either in glycerol or water for up to 14 days. In a third series the effect of three different ethanol concentrations was tested. Animals were fixed in 70%, 95% and 99.7% ethanol respectively for either two days or one week and then desiccated in glycerol for up to four weeks. Since it was not possible to determine the DNA concentration in the PCR reaction prior to amplification, the results were only scored as ranks. Three levels were used, no amplification (-), positive amplification but not enough DNA for direct sequencing (0) and positive amplification resulting in sufficient DNA for direct sequencing (+). Tables 3.1 to 3.3.2 show the scores for each of the three series and all treatments.

Table 3.1: Series 1: DNA yields for individual *C. elegans* desiccated in glycerol for up to 56 hours. DNA yields are indicated as sufficient for direct sequencing (+), insufficient for direct sequencing (0) and no yields at all (-).

Time (hours)	DNA Yield
1	++ -
2	+++
4	+++
6	++0
8	++0
16	+00
18	+00
20	000
22	++ -
24	+++
26	000

Time (hours)	DNA Yield
28	+00
32	+00
40	+++
42	+00
44	++0
46	++ -
48	+++
50	++ -
52	+++
54	+++
56	+++

Table 3.1 shows the results for Series 1 when nematodes were not fixed in ethanol at all and kept in glycerol for up to 56 hours. The DNA amplification from specimens desiccated in glycerol for up to 32 hours was very variable. Specimens desiccated for 40 hours and longer produced good DNA yields for at least one of the three animals per time and in most cases DNA of two or all three animals was amplified very strongly.

Table 3.2: Series 2: DNA yields for nematodes desiccated in glycerol or kept in water for up to 60 days, either unfixed or previously fixed in ethanol for two days. Legend as in Table 3.1.

Time (days)	DNA Yield		
	in Glycerol		in Water
	unfixed	fixed	fixed
1	+++	000	+00
2	+++	000	00
3	++	+0 -	000
4	++0	000	00
5	+++	000	00
6	++0	000	000
7	+0 -	000	++0
8	+++	000	+0
9	+++	+00	+++
10	+++	+00	++
11	++0	+00	++0
12	+00	++0	+++
13	+++	000	+++
14	+++	+++	+++
28	+++		
38		000	
54	++0		
60		000	

Table 3.2 shows the results of Series 2 when nematodes were either placed into glycerol directly from the agar plate or fixed in ethanol for two days and then put into glycerol or water. Nematodes that were not fixed in ethanol prior to glycerol treatment showed very strong bands in most cases with only one negative result and 7 weak amplifications. As for material fixed in ethanol and then kept in glycerol, apart from 12 and 14 days, at least two of the three animals per time interval showed low DNA yields and often the bands were very faint for all three specimens. The control animals, kept in water after fixation, gave very weak results for the first six days and then the yields slightly increased. There are no data after 14 days because of a lack of animals.

The Spearman Rank Correlation was applied for a comparison between the three treatments for up to 14 days of desiccation. The DNA yields were significantly greater at both the $p < 0.05$ and 0.01 significance levels when animals were not previously fixed in

ethanol compared to fixation in ethanol and desiccation (r_s = - 0.083, t = -2.781) and compared to fixation in ethanol and storage in water (r_s = 0.062, t = 0.215). There was no significant difference between DNA yields from animals stored in ethanol and then either kept in glycerol or water (r_s = 0.635, t = 2.847).

Table 3.3: Series 3a: DNA yields for nematodes fixed in ethanol for two days and desiccated in glycerol for up to four weeks. Legend as in Table 3.1.

Time in Glycerol	Ethanol Concentrations		
(Days)	99.7%	95%	70%
Control	00 -	++0	+00
Control + Wash	+00	+00	++0
1	++0	+++	000
4	+++	++0	+0 -
7	++0	+++	000
14	++0	+00	000
28	000	000	00 -

Table 3.4: Series 3b: DNA yields for nematodes fixed in ethanol for one week and desiccated in glycerol for up to four weeks. Legend as in Table 3.1.

Time in Glycerol	Ethanol Concentrations		
(Days)	99.7%	95%	70%
Control	+++	+0 -	---
Control + Wash	++0	000	---
1	+++	+00	000
4	+++	000	000
7	+00	+00	+0 -
14	+++	+00	00 -
28	000	+++	00 -

Tables 3.3 and 3.4 show the results of Series 3 when nematodes were fixed in 70%, 95% and 99.7% ethanol for 2 and 7 days respectively and then kept in glycerol for 1, 4, 7, 14 and 28 days. Almost all animals that were fixed in 70% ethanol showed no DNA amplifications at all or only very faint bands were visible. Animals fixed in 95% ethanol for two days gave good amplifications for up to two weeks in glycerol, but if they were kept in ethanol for one week at least two of the three bands per time interval were weak. The best results were obtained for animals fixed in 99.7% ethanol: There was no significant difference between fixation periods of two days or one week. DNA amplifications for animals incubated in glycerol for up to two weeks were very good and for at least two of three animals per time interval the DNA yields were high. Amplifications were very low for animals desiccated for four weeks in glycerol. The control specimens gave very variable results: after two days of fixation, DNA yields from animals that were kept in 99.7% ethanol were lower than those from animals from 70% and 95% ethanol. However, when animals were fixed for one week, no PCR products

could be obtained from specimens kept in 70% ethanol and those that were fixed in 99.7% ethanol gave the highest yields. Washing the ethanol off the specimen with distilled water did not improve the results.

The Spearman Rank Correlation showed that there were significantly higher DNA yields at both the $p < 0.05$ and 0.01 significance levels when the animals were kept in 99.7% ethanol as compared to both other concentration under both incubation periods (2 days: cf. 95%: $r_s = 0.419$, cf. 70%: $r_s = -0.037$; 7 days: cf. 95%: $r_s = -0.444$, cf. 70%: $r_s = 0.154$). Fixation of animals in 95% ethanol compared to 70% ethanol did not give significantly higher DNA yields after a two-day incubation period ($r_s = 1.685$) but there was significantly more DNA after a seven-day incubation period ($r_s = 0.203$) at both the $p < 0.05$ and 0.01 significance levels. A longer fixation time also had significant effect when the DNA yields were compared for the same concentrations. The DNA yields were significantly higher at both the $p < 0.05$ and 0.01 significance levels when animals were fixed in 99.7% ethanol for one week ($r_s = -0.194$), and lower when fixed in 95% ethanol ($r_s = 0.344$) and 70% ethanol ($r_s = -0.157$) for one week compared to a two-day fixation period.

No regression analysis could be performed on these data and there are no obvious visual trends for a change in DNA yield over time. Only after fixation in 99.7% ethanol a slight decrease in the DNA yield can be noted after 14 days (Series 3). In all other series the change in DNA yield shows no noticeable patterns and yields even increase over time. This inconsistency is probably partially a result of the data being only presented in three different rank categories and partially related to the size of the animals. Although care was taken to use similarly-sized animals, larger animals may have yielded significantly more DNA whilst amplification of DNA extracts of small animals may have failed to produce sufficient DNA. In the future, the sizes of the animals should be taken into account and the same experiments should be performed with different size classes. DNA yields should also be measured quantitatively; although no actual increase can be measured before and after PCR, quantitative data will at least allow the use of parametric tests and regression analyses. It should also be noted that the DNA yields of the same treatments in Series 2 and 3 (fixed for two days and then desiccated) showed very variable results. This suggests that repeated experiments may have large inherent variability. In a future study it would be advisable to use at least 10 animals per sampling time to get a better average result for each treatment.

The Spearman Rank Correlation showed that fixation in 99.7% ethanol resulted in significantly higher DNA yields than fixation in 95% or 70% ethanol. However, there was no significant difference when animals were fixed in 95% or 70% ethanol for only a short period of time. Fixation for a longer period of time only showed an improvement of DNA yields when the highest ethanol concentration was used. These results suggest that the use of high ethanol concentrations is required to get high quality PCR results and that longer fixation periods may also be of an advantage. However, the current data are limited by the use of only three animals per sampling time and the non-parametric analysis of the data.

Significantly higher DNA yields were obtained from animals that were extracted alive without any treatment in ethanol. This suggests that already during fixation DNA may be lost from the animals. In most practical situations extracting live animals is not possible, especially if they need to be identified prior to DNA analysis. However, a longer fixation period in ethanol of the highest molecular grade showed no loss in DNA yields. When the real samples were used in PCR they had been stored in ethanol for several weeks but the time in glycerol was kept to a minimum of a week and PCR results were generally very good. Therefore, it should be safe to store the animals in ethanol for several weeks at least.

With regards to the effect of the glycerol, DNA yields were not significantly different whether the animals were desiccated in glycerol or transferred to distilled water after fixation in ethanol. This suggests that the glycerol did not have an adverse effect on the DNA but that the degradation was related to the animals not being preserved during the period of desiccation. It was unlikely that glycerol should cause a loss of DNA because it is also frequently used in PCR to improve reactions and to store PCR chemicals such as the Qiagen Tag Polymerase and Q-Solution, both reagents that were used in this study.

4. Outgroup Analysis

The reconstruction of evolutionary relationships of any group of taxa is enhanced if it can be rooted with an outgroup. The outgroup or outgroup taxa are best chosen from a group of closely related organisms that do not form a monophyletic clade with any of the ingroup taxa. The theoretical background to rooting with outgroups is discussed in Section 1.2.5. In order to determine the best choice of outgroups for the current study the phylogenetic relationships of forty nematode sequences were reconstructed using various combinations of a total of 22 outgroup taxa. The details of each outgroup combination are found in Table A3 of Appendix III and in section 2.3.1.3 (Materials and Methods).

4.1 Critical and problematic taxa

The smallest overall number of trees was recovered when the Nematomorpha were included in the outgroups, followed by the Chaetognatha and Kinorhyncha. Inclusion of Priapulida and Tardigrada recovered the second most trees and the most trees were found when including the Gastrotricha (Table 4.1). If individual analyses were compared, excluding the Tardigrada or Gastrotricha resulted in more instances of decrease than increase of recovered trees, while the exclusion of all other phyla resulted in more instances of increase than decrease of recovered trees. The largest increase was found for the exclusion of the Nematomorpha.

Table 4.1: Recovery of most parsimonious (MP) trees under inclusion or exclusion of individual phyla. C = Chaetognatha, G = Gastrotricha, K = Kinorhyncha, N= Nematomorpha, P =Priapulida, T = Tardigrada

		C	G	K	N	P	T
Including Phylum -	**No. MP trees recovered**	180	241	205	155	235	235
	Rank (lowest to highest)	2	6	3	1	4	4
Excluding Phylum - No. of instances ...	**of increase in no. MP trees recovered**	16	11	16	18	17	14
	of no change in no. MP Trees recovered	2	6	3	3	2	2
	of decrease in no. MP trees recovered	13	14	12	10	12	15
	Instances of increase – instances of decrease	3	-3	4	8	5	-1
	Rank (lowest to highest)	4	6	3	1	2	5

Comparison of these two analyses shows that the Gastrotricha are the most problematic taxon followed by the Tardigrada, and that the Nematomorpha are the most critical taxon. The Chaetognatha and Kinorhyncha are intermediate and the Priapulida are critical or problematic depending on the analysis.

4.2 Topologies

Figures 4.1 and 4.2 show the topologies of the ingroup (Nematoda) analysis alone and the simultaneous analysis of the 40 nematoda and all 22 outgroup taxa respectively. There is only one most parsimonious solution for the ingroup analysis, but there are four equally parsimonious trees for the simultaneous analysis. All four equally parsimonious trees from the simultaneous analysis differ from each other only in the relationships amongst the Monhysterida, Araeolaimida, *Metachromadora remanei* Gerlach, 1951, and *Isolaimium* sp. Cobb, 1920. The relationships of these groups are identical in Tree 4.2.4 and in the ingroup analysis. The topologies of the simultaneous analysis differ from the ingroup topology in the following points: the relative relationships within the Enoplea, Chromadorida, Microlaimoidea, Monhysterida and Araeolaimida, and the positions of *Trichinella spiralis* Owen, 1835, *Metachromadora remanei* and *Isolaimium* sp. Only the relationships of the Rhabditida remain exactly the same amongst all topologies.

The following ingroup variations occur between the ingroup analysis and the topologies of the four equally parsimonious trees of the simultaneous analysis and those of the analyses with different outgroup combinations.

Adding outgroups to the analysis always changes the relationships within the Enoplea. Virtually all possible topologies were found during the analyses and none was reconstructed particularly frequently. However, in most cases, the Enoplea formed a monophyletic clade within the clade of the Nematoda. In some cases, *Adoncholaimus fuscus* (Bastian, 1865) is found more basal on the tree than all other nematode taxa. One could argue that this is long-branch attraction, but an examination of the initial alignment and a comparison of the branch lengths shows that *Adoncholaimus fuscus* does not form a particularly long branch compared to the remaining sequences and it is not very divergent from other Enoplea. Therefore, the first conclusion from the outgroup analysis can be that the Enoplea are the most basal nematode group and probably contain the most ancient taxa. Secondly, outgroups seem to have the strongest effect on ingroup topology in the clade closest to them.

A more complicated situation unfolds when examining the position of *Trichinella spiralis*. *Trichinella spiralis* was grouped traditionally within the Enoplea and in the ingroup analysis alone and in almost all analyses of the whole data set (see Chapter 5) it is also found there. During outgroup analysis however, it tends to group within the Chromadorida or Microlaimoidea, more particularly as sister taxon to either *Nudora bipapillata* or in come cases *Metachromadora remanei*. In the analysis of the full data set

in Chapter 5, *Trichinella spiralis* is the sister taxon to *Trichuris muris*. Both species belong to the order Trichinellida. This relationship of *Trichinella spiralis* with the Chromadorida or Microlaimoidea may result from long-branch attraction, which would be more pronounced in a small data set like this one (see Section 1.2.4.3.3). A more detailed analysis of the position of the Trichinellida is given in Chapter 8.

Metachromadora remanei (Desmodoridae, Chromadorida) is found either as the sister taxon to the Monhysterida (as in ingroup analysis alone) or intermediate between the Chromadorida and the Monhysterida. On a very few occasions it is found in other positions altogether.

Isolaimium sp. is either found between the Monhysterida clade and *Axonolaimus spinosus* (Bütschli, 1874) and *Desmolaimus zeelandicus* de Man, 1880 (ingroup alone), in one clade with the latter two species or as sister taxon to the Monhysterida in their clade. Occasionally, the *Isolaimium-Axonolaimus-Desmolaimus* clade is found closer to the chromadorids, between *Metachromadora* and the Monhysterida clade.

In the ingroup analysis alone, the Monhysterida clade contains two sub-clades, *Sabatieria punctata* Kreis, 1924 plus *Terschellingia longicaudata* de Man, 1907, and *Daptonema normandicum* (de Man, 1890) plus *Sphaerolaimus hirsutus* Bastian, 1865 with *Geomonhystera disjuncta* (Bastian, 1865). In many alternative trees, the two-clade topology is broken off into simply staggered branches of subsequently more distantly related species (see simultaneous analysis Tree 4.2.4). The latter case occurs whenever *Isolaimium* sp. groups with this clade.

The Chromadorida and Microlaimoidea are either found in two separate clades, the Microlaimoidea being more distantly related to the Monhysterida than to the Chromadorida, or they are all grouped in one clade, usually also containing *Trichinella spiralis* (see above, simultaneous analysis).

Apart from the expected variation of the ingroup topology, the outgroup topology also varies with the different combinations. The closest sister taxon to the Nematoda are the Tardigrada. Only in two cases (3 trees, see below) where the Tardigrada and *Priapulus* (Priapulida) are included in the analysis, the latter forms the closest sister group to the Nematoda. The Chaetognatha always represent the most distantly related phylum to the Nematoda. The Gastrotricha do not form a monophyletic clade. In all most parsimonious trees of the simultaneous analysis and also in the analyses of the whole data set, *Chaetonotus* sp. Ehrenberg, 1830, and *Lepidodermella squammata* Dujardin, 1841, form one clade and *Turbanella cornuta* Remane, 1924, is more distantly related. In the

following summary of the variations within the outgroup relationships this paraphyletic relationship becomes more evident:

Pycnophyes kielensis Zelinka, 1928, (Kinorhyncha) shows an unstable position in the outgroup topology in the absence of *Priapulus caudatus* (Priapulida). It is normally found as the closest sister taxon to *Priapulus caudatus*, but in its absence it sometimes splits the two Gastrotricha species, *Lepidodermella squammata* and *Turbanella cornuta*, or in one case (Gastrotricha, Kinorhyncha, Nematomorpha, Tardigrada), it was found to split the Nematomorpha. In another case (Chaetognatha, Gastrotricha, Kinorhyncha, Nematomorpha), *Pycnophyes kielensis* split the Gastrotricha and *Lepidodermella squammata* was found within the Nematomorpha. These findings suggest that *Pycnophyes kielensis* only shares significant synapomorphies with *Priapulus caudatus* and that *Priapulus caudatus* thereby stabilises its position.

Priapulus caudatus is found to be more closely related to the ingroup (Nematoda) than *Pycnophyes kielensis* when the Nematomorpha, or in one case the Gastrotricha, are excluded. Similarly, *Priapulus caudatus* has been found to be the closest sister taxon to the Nematoda even in the presence of the Tardigrada, which in all other cases are the closest sister taxon; this situation occurred in the absence of *Pycnophyes kielensis* and the Nematomorpha for two combinations of outgroup phyla ([Chaetognatha, Priapulida, Tardigrada] and [Gastrotricha, Priapulida, Tardigrada]). It seems that the Nematomorpha keep *Priapulus caudatus* in a stable position within the outgroup as they are one of the closely related taxa.

In the absence of *Pycnophyes kielensis*, *Turbanella cornuta* is sometimes more closely related to the Nematoda than *Lepidodermella squammata* (both Gastrotricha). There was one combination (Chaetognatha, Gastrotricha, Nematomorpha) where the Nematomorpha were found to split the Gastrotricha.

Overall it shows that including more outgroups stabilises at least the outgroup relationships. This comes as no surprise as few distantly related taxa will increase the possibility of long-branch attraction. As the choice of outgroups does not seem to influence the ingroup topology particularly, it was most appropriate to include all possible outgroups to the analysis of the nematode relationships.

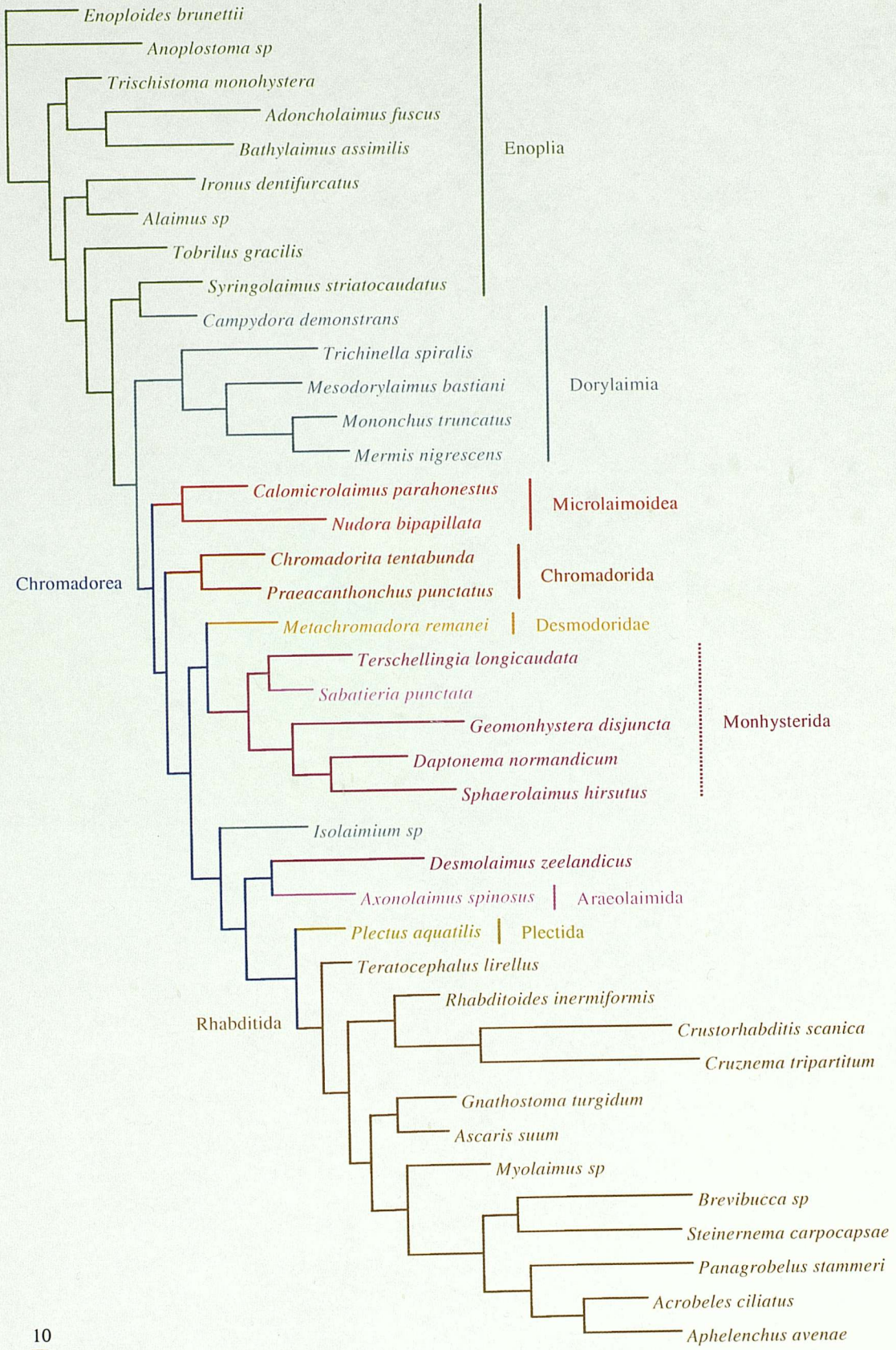


Figure 4.1: Single most parsimonious tree of the analysis of the ingroup taxa (nematodes) alone; dotted lines depict non-monophyletic taxa

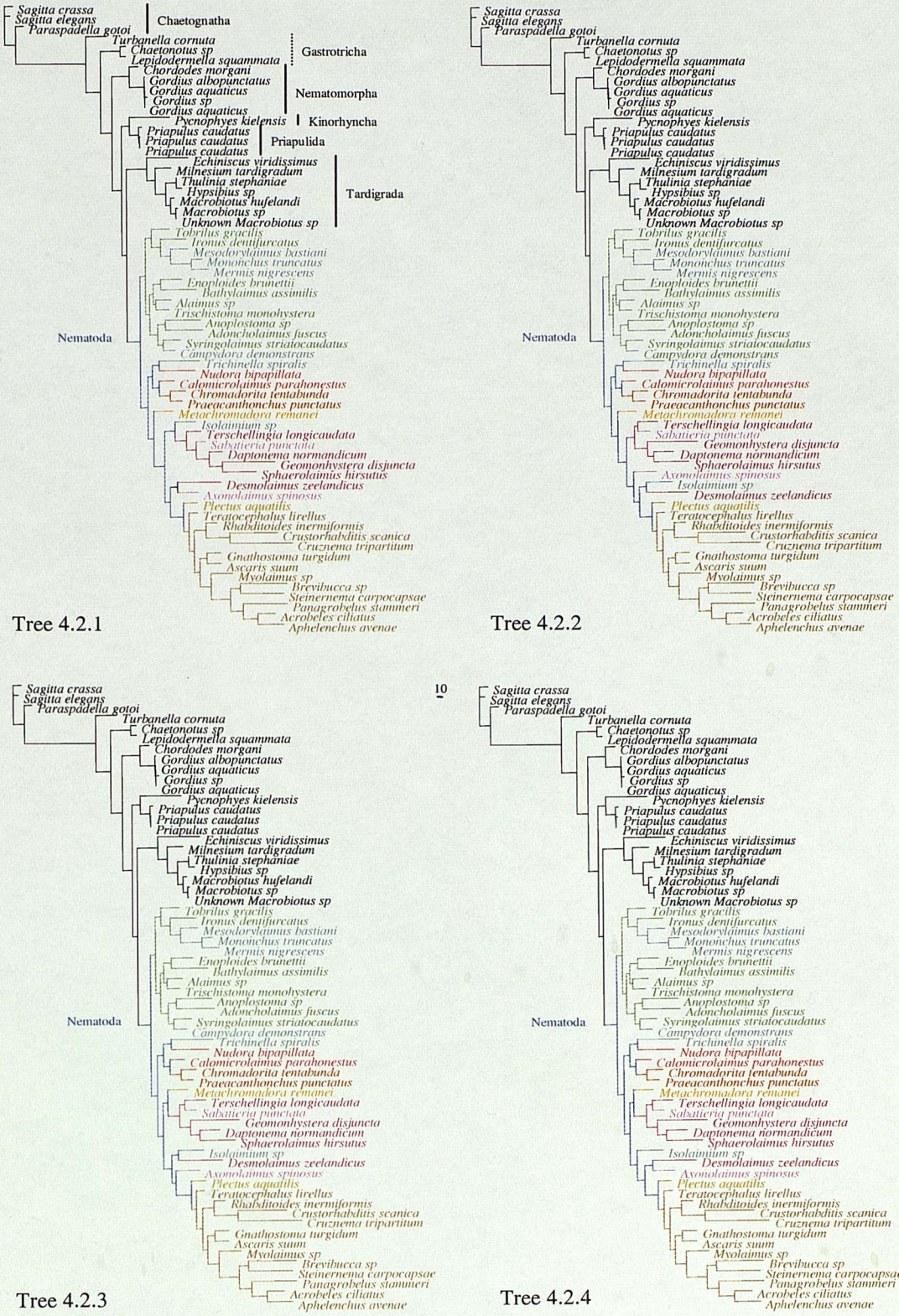


Figure 4.2: Four most parsimonious trees of the simultaneous analysis of 40 ingroup taxa (nematodes) and 22 outgroup taxa; colour coordination as in Figure 4.1, outgroups in black

5. Analysis of Molecular Sequence Data

The evolutionary history of 214 rRNA sequences from nematodes and 16 outgroup sequences were modelled using four different algorithms: parsimony (unweighted and weighted), LogDet transformation, maximum likelihood distances and Bayesian inference based on MCMCMC. The characters were divided into four different sets: using all 1769 characters, excluding 1079 positions with gaps in one or more sequences, excluding 780 parsimony-uninformative characters and excluding both, 1376 positions with gaps and parsimony-uninformative characters. Under the maximum likelihood distance model and the Bayesian inference parsimony-uninformative characters were never excluded.

Figures 5.1 to 5.5 show the strict consensus tree of the consensus trees of the four analyses using different character sets under unweighted and weighted parsimony, the LogDet distance and maximum likelihood distance criteria and under the Bayesian inference respectively. The bootstrap proportions and posterior probabilities for each analysis are superimposed on the branches leading to the relevant nodes. The differences between the individual analyses under the same algorithm are discussed in Section 5.1 and the support values are discussed for all algorithms together in Section 5.2. Variations to traditional classifications are summarised in Section 5.3 and interpreted in Chapter 8.

5.1 Differences based on the choice of model and character inclusion

5.1.1 Parsimony

Table 5.1 shows the summary of all indices for all eight parsimony analyses. The consistency index (CI), retention index (RI), rescaled consistency index (RC) and homoplasy index (HI) are identical for all fundamental trees of an individual analysis and so is the Goloboff-Fit under the weighted parsimony criterion. The values for the Goloboff-Fit vary in the value of the decimal places for the fundamental trees under all unweighted parsimony analyses. The pattern of change in the homoplasy content is similar for unweighted and weighted parsimony. The CI, RI and RC all decrease when uninformative characters are excluded. Accordingly, the HI increases with the exclusion of these characters. Although uninformative characters do not actually add any homoplasy to the data (they are either constant or apomorphic) the consistency index is dependent on the number of characters in the data set. It is this stochastic quality of the CI that results in its decrease and the increase of the HI with the inclusion of uninformative characters. The exclusion of positions with gaps on the other hand increases the consistency indices. Since

most gaps are found where the alignment is thought to be more ambiguous, it is no surprise that the homoplasy content decreases with the exclusion of the gapped positions. Clearly, a lot of the homoplasy is contained in those characters that contain gaps for certain taxa. Overall, character weighting increases all consistency indices by down-weighting those characters with high homoplasy content.

Table 5.1: Tree length (TL), consistency index (CI), CI without uninformative characters (excl uninf), retention index (RI), rescaled consistency index (RC), homoplasy index (HI), HI without uninformative characters and Goloboff-Fit (G-Fit) for all parsimony models (unweighted and weighted) including all characters (incl all), excluding positions with gaps (excl gapped), excluding uninformative characters and excluding positions with gaps and uninformative characters.

	Unweighted Parsimony				Weighted Parsimony			
	Incl all	**Excl gapped**	**Excl uninf**	**Excl gapped & uninf**	**Incl all**	**Excl gapped**	**Excl uninf**	**Excl gapped & uninf**
TL	16263	4148	15993	4017	1553	606	1279	477
CI	0.1607	0.2300	0.1464	0.2051	0.4109	0.4832	0.2858	0.3386
CI excl uninf	0.1464	0.2051	0.1464	0.2051	0.2852	0.3421	0.2858	0.3386
RI	0.6396	0.7266	0.6395	0.7267	0.7539	0.8325	0.7547	0.8302
RC	0.1028	0.1671	0.0936	0.1491	0.3097	0.4022	0.2157	0.2811
HI	0.8393	0.7700	0.8536	0.7949	0.5891	0.5168	0.7142	0.6614
HI excl uninf	0.8536	0.7949	0.8536	0.7949	0.7148	0.6579	0.7142	0.6614
G-Fit	-358	-171	-357	-171	-110	-52	-110	-50

The exclusion of positions with gaps under the unweighted parsimony results in less resolution of all major clades (Enoplia, Monhysterida, Rhabditida, Spirurina and Rhabditina). When parsimony-uninformative characters are excluded, resolution is also reduced in all major clades and most markedly, the Monhysteridae plus Comesomatidae form the sister group to the Desmodoridae, a relationship otherwise not observed. Excluding both, positions with gaps and uninformative characters, results in little loss of resolution. The best overall resolution is achieved with all characters included or excluding all potentially ambiguous and uninformative positions. The benefit of excluding both sets of characters is also mirrored in the improvement of the consistency indices. Gapped positions potentially introduce many homoplasies and uninformative characters inflate the consistency index. Therefore, a combination of excluding all of these potentially misleading characters gives the best results under the unweighted parsimony criterion.

In contrast to unweighted parsimony, under weighted parsimony the exclusion of gapped or uninformative characters does not reduce the resolution of the consensus trees. The consistency indices increase with the exclusion of positions with gaps but excluding both character sets results in a lower consistency index than including all character (but

still higher than only excluding uninformative characters). This shows that character weighting has already reduced the amount of homoplasy by down-weighting more homoplasious characters in the whole data set. However, considering the effect of positions with gaps and uninformative characters these are probably best excluded under weighted parsimony as well.

5.1.2 LogDet distances

The consensus trees for all four data sets are well resolved under the LogDet criterion. Including all characters and excluding parsimony-uninformative characters even resulted in one best tree only in contrast to other analyses that produced hundreds or even tens of thousands of best trees. Nevertheless, there is still considerable conflict in the relationships of the different orders and suborders between the four consensus trees.

The exclusion of both character sets, positions with gaps and parsimony-uninformative characters, results is an uncertain position of the Trichinelloidea which can be found most closely related to the outgroup taxon Tardigrada. The exclusion of such a number of characters may have resulted in the exclusion of informative characters for the positioning of the Trichinelloidea. On the other hand, the topology of the Enoplida, most contentious in many other analyses, is almost constant across all four character sets. In the light of the potentially misleading consequences of excluding both sets of ambiguous characters there is the suggestion of retaining uninformative characters. However, the results of all four analyses are considered in the phylogenetic interpretation in Section 5.3.

5.1.3 Maximum likelihood distances

Under the maximum likelihood distance criterion only one best tree was found when all characters were included but several hundred when gapped positions were excluded from the reconstruction. Nevertheless, in the second analysis only few polytomies are found, restricted to the position of *Rhabditoides inermiformis* (Osche, 1952), *Steinernema carpocapsae* and the internal relationships of the Rhabditidae. A comparison of the two analyses results in conflicts of relationships of families within the Enoplida, Dorylaimida and Monhysterida and between species in the Desmodoridae, Panagrolaimidae, Spiruromorpha, Ascaridomorpha, Tylenchomorpha, Rhabditidae and Strongylidae. As can be seen in Section 5.3 below these are groups where conflicts occur under most other criteria, too. Since there are no particularly dubious relationships in either analysis both topologies are retained for further analysis. One might favour the exclusion of positions with gaps because it is expected to reduce homoplasy.

5.1.4 Bayesian inference

Two character sets were analysed under the Bayesian criterion, including all characters and excluding gapped positions. Both analyses where repeated five times under the same conditions. For all ten analyses the development of all parameters, log-likelihood, tree length, substitution rates, base frequencies, α shape parameter and the proportion of invariant sites, were plotted against the number of generations to establish whether the analyses reached a constant level for each.

Under the inclusion of all characters two analyses showed constant levels for all parameters after the initial burn-in phase. In two cases a constant level for the tree length was only achieved after the preset burn-in phase of 250,000 generations. In one other analysis a sudden change occurred between 1.8 – 2.0 million generations with an increase in tree length, change in base frequencies and change in value of the α shape parameter for these 200,000 generations only. Additionally, in this analysis the outgroup sequences of *Paraspadella gotoi* (Casanova, 1990) (Chaetognatha) is found within the Rhabditida. The two analyses for which all parameters reach a constant level within the burn-in phase differ in two general details; firstly, in one of the analyses *Diplolaimella* sp. Allgén, 1929, is found with the Enoplida instead of the other Monhysteridae, and secondly, in the other of the two analyses there is no resolution between the three suborders of the Rhabditida. Therefore, the results from the analysis with the more likely position of *Diplolaimella* sp. will be used for any further comparison despite the fact that this is the analysis with the lack of resolution between the rhabditid suborders.

When positions with gaps were excluded the topologies and parameter behaviours were less constant. In three analyses one or two chaetognath sequences were found within the Nematoda, once with the Monhysteridae and twice with the Rhabditina respectively. In all five analyses the tree length never reached a constant level but kept increasing throughout the analysis. In one analysis with dubious outgroup topology the α shape parameter only reached a constant level after the pre-determined burn-in phase of 150,000 generations. In all five analyses the substitution rates between cytosine and thymine were much more variable than the rates for all other changes but the variation was restricted to a constant range after the burn-in phase. The two analyses with the best parameter behaviour (only the tree length remained unstable) differ in the level of resolution of the Enoplida, the internal topology of the Dorylaimida, the level of resolution in the Monhysterida and the lack of resolution in the Rhabditida, especially the resolution in the Cephalobidae, Tylenchomorpha and Rhabditina, and the positions for the Plectida, *Teratocephalus lirellus*

Anderson, 1969, and *Steinernema carpocapsae*. Since the lack of resolution was not restricted to only one of the two analyses both topologies were retained for further analyses.

Of the four different algorithms the phylogenetic reconstructions under Bayesian inference result in the best resolution of the consensus trees. On the one hand, this is probably due to the fact that the consensus trees of the Bayesian analyses are majority-rule consensus trees and the others are strict consensus trees. On the other hand, the steady search of the likelihood surface aided by the use of heated chains that allow the traversal of valleys under the Bayesian criterion may be significantly more efficient in finding global optima than any other heuristic search.

5.2 Bootstrap support and posterior probabilities

The significance of the bootstrap support is interpreted according to Hillis & Bull (1993). Clades with over 85% bootstrap support are very likely to be monophyletic. Clades with support between 65% and 85% are estimated to have moderate support for their monophyly and clades with less than 65% bootstrap support are discounted as monophyletic. The minimum and maximum values of the bootstrap support for monophyletic clades are recorded in Tables 5.2 to 5.4 below.

The bootstrap support was usually very similar for clades in common for the two analyses either including positions with gaps or excluding them respectively. Most clades that are found in the majority of reconstructions and that are traditionally thought of as monophyletic (see Section 5.3) are well supported, the majority having over 90% bootstrap support. However, some of the clades with generally significant bootstrap support have insignificant support when positions with gaps are excluded. The lack of support when positions with gaps are excluded is more significant under the unweighted parsimony, LogDet and maximum likelihood distance criteria than when characters have been weighted prior to parsimony analysis. There are a few occasions when the bootstrap support is higher when positions with gaps are excluded; these clades correspond to traditional monophyletic groups for the Triplonchida (excluding *Trischistoma monohystera* [de Man, 1880]) under all criteria but weighted parsimony, for the Axonolaimidae under unweighted parsimony including all characters and for the Spiruromorpha under maximum likelihood distances.

The posterior probabilities of the Bayesian analyses behave rather similarly to the bootstrap proportions. They are generally higher than or at the higher end of the range of

the bootstrap values with the exception of the support for the Monhysteridae and Spirurina where the posterior probabilities are lower then the bootstrap values when positions with gaps are excluded from the analysis. Generally, as in the case of the bootstrap proportions, the posterior probabilities are lower under the exclusion of positions with gaps than when those characters remain in the analysis.

Since positions with gaps potentially contain a lot of homoplasy (see increase of CI in Section 5.1.1 above), the high bootstrap proportions and posterior probabilities are an indication that these positions contain a lot of consistent but phylogenetically misleading information. Additionally, there is less loss of bootstrap support when positions with gaps are excluded if characters have been weighted according to their rescaled consistency index. This indicates that there may be significant character conflict in the remaining, parsimony-informative characters and that down-weighting potentially homoplastic characters reduces this conflict and therefore increases the bootstrap values and thereby the confidence in the phylogenetic reconstruction.

5.3 Consistency of current model results with previous classifications

This section contains an in-depth discussion of the phylogenetic reconstructions under the different models and characters sets. Table 5.2 lists those relationships that are consistently found under all models. Table 5.3 lists those relationships that are also consistently found in all analyses but do not form a monophyletic group according to previous interpretation of data. Table 5.4 lists those taxa with conflicting phylogenetic relationships based on previous data but which have been found consistently in the same relationships in the present study. The minimum and maximum bootstrap proportions (in percentage) and posterior probabilities (as proportions) are given in parenthesis in the tables or in the text for those groups not included in the tables. The taxa in each table are ordered according to Table A2.3 in Appendix II which gives the most recent classification of the phylum Nematoda according to de Ley & Blaxter (2002).

The monophyly of the Nematoda is well supported when positions with gaps are included (90 – 96%, 1.00) but there is insignificant support when such positions are excluded (26 – 61%, 0.82 – 0.96). Under the LogDet criterion when positions with gaps and uninformative characters are excluded the placement of the Trichinelloidea becomes uncertain (closer to the outgroup Tardigrada than the remaining Dorylaimia) but under all other conditions the Trichinelloidea are found in one clade with the Dorylaimia unless the Dorylaimia are entirely unresolved (unweighted parsimony excluding uninformative characters).

Table 5.2: Monophyletic groups that are found consistently in all analyses. Bootstrap proportions and posterior probabilities in parenthesis.

Phylum Nematoda (26 – 96%, 0.82 – 1.00)
Oncholaimoidea / Oncholaimina (100%, 1.00)
Tripyloididae / Tripyloidoidea / Tripyloidina (100%, 1.00)
Trichodoridae (96 – 100%, 1.00)
Diphtherophoroidea/Diphtherophorina (93 – 100%, 1.00)
Triplonchida excluding Tripylina (72 – 99%, 1.00)
Qudsianematidae (58 – 100%, 1.00)
Dorylaimina /Dorylaimida excluding *Campydora demonstrans* (96 – 100%, 1.00)
Mononchina / Mononchida (54 – 100%), 1.00
Trichinelloidea / Trichinellida (90 – 100%, 0.95 – 1.00)
Chromadorea (30 – 78%, 0.80 – 1.00)
Chromadoridae (95 – 100%, 1.00)
Cyatholaimidae (51 – 100%, 0.91 – 1.00)
Chromadoroidea / Chromadorina / Chromadorida (37 – 95%, 0.89 – 1.00)
Desmodoridae including *Xyzzors* sp. (99 – 100%, 1.00)
Microlaimidae (93 – 100%, 0.99 – 1.00)
Monoposthiidae (54 – 100%, 0.75 – 1.00)
Monhysteridae / Monhysteroidea (98 – 100%, 0.54 – 1.00)
Axonolaimidae (78 – 100%, 1.00)
Comesomatidae (89 – 100%, 1.00)
Plectidae / Plectoidea / Plectida including *Tripyla* sp. (66 – 100%, 0.96 – 1.00)
Filarioidea including *Setaria digitata* (29 – 100%, 1.00)
Spiruromorpha including *Setaria digitata* (58 – 89%, 0.99 – 1.00)
Spirurina (91 – 99%, 0.80 – 1.00)
Panagrolaimidae / Panagrolaimoidea (94 – 100%, 0.97 – 1.00)
Cephalobidae / Cephaloboidea / Cephalobomorpha (100%, 1.00)
Criconematoidea including *Paratylenchus dianthus* (98 – 100 %, 1.00)
Hoplolaimidae including *Globodera pallida* (57 – 100%, 1.00)
Diplogasteroidea / Diplogasteromorpha (56 – 100%, 0.77 – 1.00)
Strongyloidea (84 – 99%, 0.95 – 1.00)

Table 5.3: Relationships that are consistently found in all analyses but do not form a monophyletic group according to previous interpretation of data. Bootstrap proportions and posterior probabilities in parenthesis.

Campydora demonstrans and *Syringolaimus striatocaudatus* (68 – 96%, 1.00)
Qudsianematidae and *Pungentus* sp. (67 – 100%, 1.00)
Mermis nigrescens as sister taxon to Mononchina (100%, 1.00)
Xyzzors sp. as part of Desmodoridae (99 – 100%, 1.00)
Cyartonema elegans and *Terschellingia longicaudata* (93 – 100%, 1.00)
Rhabditophanes sp. and *Strongyloides ratti* (100%, 1.00)
Globodera pallida within or as sister taxon to Hoplolaimidae (57 – 100%, 1.00)
Paratylenchus dianthus within or as sister taxon to Criconematoidea (98 – 100 %, 1.00)
Tripyla sp. with Plectida (66 – 100%, 0.96 – 1.00)
Diploscapter sp. and *Protorhabditis* sp. (94 – 100%, 1.00)
Setaria digitata with Filarioidea (29 – 100%, 1.00)

Table 5.4: Taxa with conflicting phylogenetic relationships based on previous data but which have been found consistently in the same relationships in the present study. Bootstrap proportions and posterior probabilities in parenthesis.

Cyartonema elegans with *Terschellingia longicaudata* (93 – 100%, 1.00)
Comesomatidae in Monhysterida (67 – 85%, 0.52 – 1.00)
Plectidae as closest sister taxon to the Rhabditida
Teratocephalus lirellus as closest sister taxon to the Plectida or Rhabditida
Setaria digitata with the Filarioidea (29 – 100%, 1.00)
Paratylenchus dianthus within or as sister taxon to Criconematoidea (98 – 100%, 1.00)
Globodera pallida within or as sister taxon to Hoplolaimidae (57 – 100%, 1.00)
Rhabditophanes sp. with *Strongyloides ratti* (100%, 1.00)

Many of the traditional monophyletic groups, mainly the families and superfamilies, are recovered with the current character sets and models. However, there are a number of major groups that are not recovered under all models. The relationship between the Enoplia and Dorylaimia is uncertain. Of those consensus trees that are resolved at this level, the Dorylaimia are the most primitive of the three nematode subclasses under unweighted parsimony excluding positions with gaps and uninformative characters, under weighted parsimony excluding positions with gaps and excluding positions with gaps and uninformative characters and under Bayesian inference excluding positions with gaps; the Enoplia are the most primitive subclass under unweighted parsimony excluding positions with gaps, under weighted parsimony including all characters and excluding uninformative characters, under LogDet distances including all characters and excluding positions with gaps and uninformative characters, under maximum likelihood distances excluding positions with gaps and under Bayesian inference including all characters. The class Enoplea is monophyletic (19 – 48%), forming a sister group to the class Chromadorea, under unweighted parsimony and maximum likelihood distances including all characters and under LogDet distances including all characters and excluding uninformative characters. Under unweighted parsimony excluding uninformative characters the three subclasses are unresolved. The Chromadorea are always monophyletic unless they are unresolved at that level (unweighted parsimony excluding uninformative characters).

Within the Enoplea, the Enoplia (3 – 48%, 0.72 – 0.99) and Dorylaimia (18 – 80%, 0.79 – 1.00) are usually monophyletic if they are resolved at that level. Within the Enoplia, the order Enoplida is usually not fully resolved and many different combinations of relationships exist depending on the model and character set used. Two groups, the Oncholaimoidea and the Tripyloididae, are always monophyletic and supported 100%. The Enoploidea are usually paraphyletic or unresolved but can have moderate support when

they are monophyletic (33 – 87%, 0.57). The two species of the Ironidae are always found with the other Enoplida but are never closely related to each other. The position of *Ironus dentifurcatus* (Argo & Heyns, 1972) is very variable and *Syringolaimus striatocaudatus* de Man, 1888, is always found as the sister taxon to *Campydora demonstrans* Cobb, 1920, which was previously though to belong to the Dorylaimida.

The two species of the Tripylina (*Trischistoma monohystera* and *Tripyla* sp. Bastian, 1865) are never part of the clade of the remaining Triplonchida, which are usually fully resolved. The Trichodoridae and Diphtherophorina are always monophyletic and strongly supported and the Tobrilina form their sister taxa. *Trischistoma monohystera* is found within the Enoplida, often as the sister group to *Alaimus* sp. de Man, 1880, but never with the other species of the Triplonchida. *Tripyla* sp. likewise is never found with the other Triplonchida species but consistently with the Plectida.

Within the Dorylaimia the four orders Dorylaimida, Mononchida, Mermithida and Trichinellida are all monophyletic. *Mermis nigrescens* Dujardin, 1842, the only representative of the Mermithida, is always the sister taxon to the Mononchida with 100% support. The Trichinellida are either found as sister taxon to the Mermithida plus Mononchida or as sister taxon to the remaining Dorylaimia if the Mermithida, Mononchida and Dorylaimida form one clade. *Campydora demonstrans* and *Isolaimium* sp. are never found with the remaining Dorylaimia. *Campydora demonstrans* is always found as the sister taxon to *Syringolaimus striatocaudatus* in the Enoplida and *Isolaimium* sp. is usually found as sister taxon to the Axonolaimidae or to the Axonolaimidae plus *Desmolaimus zeelandicus* in the Chromadorea.

Within the Dorylaimida most families are paraphyletic; the Longidoridae are always monophyletic and the Qudsianematidae are monophyletic apart from one of the two alternative topologies under Bayesian inference excluding positions with gaps when *Pungentus* sp. Thorne & Swanger, 1936, is part of this family. Otherwise, *Pungentus* sp. is found as the nearest sister taxon to the Qudsianematidae. The inter-relationships between the representative species of the Mononchida are variable and therefore the Mononchoidea and Mononchidae are probably paraphyletic.

Only two of the six proposed orders of the Chromadorea, the Chromadorida and Plectida, are monophyletic. Within the Chromadorida the Chromadoridae and Cyatholaimidae are also always monophyletic apart from one case (weighted parsimony excluding gapped positions) when the Cyatholaimidae are unresolved. Either the Chromadorida or the Microlaimoidea are the most basal taxon of the Chromadorea.

The Desmodorida are polyphyletic because the Microlaimidae and Monoposthiidae never form a monophyletic group with the Desmodoridae. The Microlaimidae and Monoposthiidae are always monophyletic. These two families also tend to form the monophyletic superfamily Microlaimoidea (not under weighted parsimony excluding positions with gaps and LogDet including all characters) but it is only weakly supported. Both families are usually found as sister groups to the Chromadorida or at least closely related to this order. They are sometimes found closely related to the Desmodoridae but never as their direct sister taxon. *Xyzzors* sp. Inglis, 1966, is always included in the Desmodoridae despite being previously placed in the Cyatholaimidae (Chromadorida) rendering this family potentially paraphyletic; the internal relationships of the Desmodoridae are variable and often unresolved.

The orders Monhysterida and Araeolaimida are both paraphyletic. The Comesomatidae are always included in the Monhysterida and not in the Araeolaimida as previously suggested. The Linhomoeidae are polyphyletic because *Desmolaimus zeelandicus* is always found closely related to the Axonolaimidae (Araeolaimida) whilst *Terschellingia longicaudata* is always found in the Monhysterida. The Monhysteridae are always monophyletic. The Xyalidae (42 – 95%, 0.95) and Sphaerolaimoidea (39 – 82%, 1.00) are either monophyletic, paraphyletic or unresolved. The extended Monhysterida are consistently reconstructed and moderately to well supported (69 – 85%, 0.52 – 1.00). The Axonolaimidae are always monophyletic and together with *Desmolaimus zeelandicus* and *Isolaimium* sp. form the closest sister group to the Plectida. *Cyartonema elegans* Jayasree & Warwick, 1977, is always found as the sister taxon to *Terschellingia longicaudata* and therefore belongs most likely to the Monhysterida.

The Plectida are always monophyletic and usually well supported. They are the sister taxon to the Rhabditida under unweighted and weighted parsimony including all characters and excluding uninformative characters and under Bayesian inference including all characters. The Plectida form the sister taxon to the Rhabditina under weighted parsimony excluding positions with gaps and excluding positions with gaps and uninformative characters and under Bayesian inference excluding positions with gaps. They are the sister taxon to the Spirurina under LogDet and maximum likelihood distances for all character sets. Under unweighted parsimony excluding positions with gaps and excluding positions with gaps and uninformative characters the Plectida are found unresolved in the Rhabditida.

The Rhabditida are either monophyletic (75 – 81%, 0.65) or paraphyletic including the Plectida. Of the three suborders Spirurina, Tylenchina and Rhabditina only the Spirurina are always monophyletic, however, for the other two suborders at least the majority of species are usually found in one clade. Under unweighted parsimony including all characters and excluding uninformative characters and under weighted parsimony including all characters and excluding positions with gaps and uninformative characters the Rhabditina and Tylenchina form one clade and are sister taxa to the Spirurina. Under unweighted parsimony excluding positions with gaps and excluding positions with gaps and uninformative characters and under Bayesian inference the relationships between the subclasses are uncertain. Under weighted parsimony excluding positions with gaps the Plectida, Spirurina and Rhabditina form one clade and are the sister taxon to the Tylenchina. Under weighted parsimony excluding uninformative characters the Spirurina and Tylenchina form one clade and are the sister taxon to the Rhabditina. Under the LogDet and maximum likelihood distances the Tylenchina and Rhabditina form one clade and are the sister taxon to the Spirurina and Plectida.

Teratocephalus lirellus is the sister taxon to the Plectida under unweighted parsimony excluding positions with gaps, weighted parsimony excluding positions with gaps and uninformative characters and under LogDet and maximum likelihood distances; it is the sister taxon of the Rhabditida under unweighted and weighted parsimony including all characters and excluding uninformative characters and under Bayesian inferences including all characters; under all other conditions its position in the Rhabditida is unresolved.

The position of *Brevibucca* sp. Goddey, 1935, remains uncertain; under Bayesian inference including all characters it is the most basal taxon of the Tylenchina. Under unweighted parsimony and under maximum likelihood distance criterion including all characters and under weighted parsimony and LogDet distances including all characters and excluding uninformative characters it is found as sister taxon to the Panagrolaimomorpha (Tylenchina). In all other cases its position in the Rhabditida is unresolved.

The Spirurina and Spiruromorpha are always monophyletic and the Ascaridomorpha are either monophyletic with little support or, more often, unresolved. *Setaria digitata* (Linstow, 1906), of previously uncertain position, is always found within the Filarioidea or Onchocercidae. The Onchocercidae also include *Loa loa* (Guyot, 1778) of the Filariidae. Therefore, at least the Onchocercidae are paraphyletic. *Ascarophis arctica* Poljansky,

1952, is always found as the most basal taxon of the Spiruromorpha. The families and superfamilies of the Ascaridomorpha are either paraphyletic or unresolved altogether. Four infraorders are represented by only a single species in this data set. *Gnathostoma turgidum* Stossich, 1902, is always found as the most basal species within the Spirurina. *Philonema* sp. Kuitunen-Ekbaum, 1933, and *Dentostomella* sp. are found close to the base of the Spirurina but their precise relationships vary with different algorithms and data sets. *Brumptaemilius justini* Adamson & Anderson, 1985, is usually found basal to the Ascaridomorpha but its position is unresolved in the Spirurina in many cases.

The position of *Myolaimus* sp. Cobb, 1920, remains uncertain. It is either found as the most basal taxon of the Spirurina (all analyses including all characters, unweighted and weighted parsimony and LogDet distances excluding uninformative characters and weighted parsimony excluding gapped positions) or its position in the Rhabditida is unresolved.

The Tylenchina are usually monophyletic but weakly supported (8 – 56%, 0.58 – 1.00). They include *Brevibucca* sp. as sister taxon to the Panagrolaimomorpha under all algorithms including all characters apart from Bayesian inferences and under weighted parsimony and LogDet distances excluding uninformative characters. Under unweighted parsimony excluding uninformative characters and under LogDet distances excluding positions with gaps and uninformative characters the position of *Steinernema carpocapsae* is uncertain in the Rhabditida. The relationships between the three infraorders, Panagrolaimomorpha, Cephalobomorpha and Tylenchomorpha are unresolved under Bayesian inference including all characters and unweighted parsimony excluding uninformative characters. Otherwise the Cephalobomorpha and Tylenchomorpha are sister taxa and the sister taxon to the Panagrolaimomorpha.

The internal relationships within the Panagrolaimomorpha vary substantially. The Panagrolaimidae are always monophyletic. The Strongyloidoidea are at least paraphyletic because *Strongyloides ratti* Sandground, 1925, and *Rhabditophanes* sp. Fuchs, 1930, (Rhabditina) are always sister taxa. *Steinernema carpocapsae* is usually found as their sister taxon under parsimony models. Under weighted parsimony excluding positions with gaps and under Bayesian inference including all characters it is the most basal taxon of the Tylenchina. Under LogDet distances including all characters and excluding uninformative characters and under maximum likelihood distances *Steinernema carpocapsae* is the most basal taxon of the Panagrolaimomorpha. Its position is unresolved within the Panagrolaimomorpha under LogDet distances excluding positions with gaps and under one

of the two options of Bayesian inference excluding positions with gaps and unresolved within the Rhabditida under LogDet distances excluding positions with gaps and uninformative characters and under the other of the two options of Bayesian inference excluding positions with gaps. *Aphelenchoides fragariae* (Ritzema Bos, 1890) (Tylenchomorpha) is often, but not always, part of the clade of Panagrolaimomorpha (see below). The Cephalobomorpha are always monophyletic but the internal relationships vary between models.

The Tylenchomorpha are only monophyletic under Bayesian inference (0.63 – 0.97). The Tylenchomorpha excluding the Aphelenchoidea (see below) are monophyletic and moderately to well supported (45 – 90%, 0.97 – 1.00) under all other models apart from unweighted parsimony excluding uninformative characters in which case they are unresolved. *Aphelenchus avenae* Bastian, 1865, is the sister group to the remaining Tylenchomorpha under all unweighted parsimony models apart from excluding uninformative characters, under weighted parsimony including all characters and excluding uninformative characters, LogDet transformation excluding gapped positions and excluding gapped and uninformative characters and under maximum likelihood distance excluding gapped positions. It is the sister group to the Cephalobomorpha and the remaining Tylenchomorpha under weighted parsimony excluding gapped positions and excluding gapped and uninformative characters and the sister taxon to the Cephalobomorpha under LogDet distances including all characters and excluding uninformative characters and under maximum likelihood distances including all characters. *Aphelenchoides fragariae* is the sister taxon to the Panagrolaimidae under all unweighted parsimony models and all weighted parsimony models apart from excluding gapped positions, or the sister group to the Cephalobomorpha and remaining Tylenchomorpha including *Aphelenchus avenae* under weighted parsimony excluding gapped positions. It is found within the Panagrolaimomorpha under all LogDet distance models apart from excluding gapped positions and under both maximum likelihood distance models. It is found with the Rhabditidae partim under LogDet distances excluding gapped positions.

The Criconematoidea are mostly monophyletic (52 – 71%, 0.83 – 1.00) with *Paratylenchus dianthus* Jenkins & Taylor, 1956, as their direct sister taxon; under unweighted parsimony and weighted parsimony including all characters and excluding uninformative characters *Paratylenchus dianthus* is part of this superfamily. The Anguinidae are either paraphyletic including *Deladenus* sp. Thorne, 1941, and the Tylenchid from the seaweed or unresolved. The Hoplolaimidae are monophyletic (33 –

92%, 0.75 – 1.00) with *Globodera pallida* Stone, 1972, as their sister taxon under unweighted parsimony and under weighted parsimony including all characters and excluding uninformative characters. In all other cases *Globodera pallida* is part of this family. *Meloidogyne incognita* (Kofoid & White, 1919) is found with *Nacobbus aberrans* (Thorne, 1935), *Pratylenchus goodeyi* Sher & Allen, 1953, *Radopholus similis* (Sher & Allen, 1953) or *Zygotylenchus guevarai* (Tobar Jiménez, 1963) or its position is unresolved within the Tylenchoidea. The Belonolaimidae and Pratylenchidae are always paraphyletic or unresolved.

The Rhabditina are either monophyletic (12 – 45%, 1.00) or unresolved. *Bunonema franzi* Andrássy, 1971, is usually the sister taxon to the Diplogasteroidea or very closely related to them; under unweighted parsimony excluding positions with gaps, uninformative characters or both its position in the Rhabditina is unresolved. The Diplogasteroidea are always monophyletic but the Diplogasteridae are never monophyletic because *Diplogaster lheritieri* Maupas, 1919 is more closely related to *Pristionchus pacificus* Sommer *et al.*, 1996 (Neodiplogasteridae) than *Aduncospiculum halicti*. The Rhabditidae, Rhabditoidea and Rhabditomorpha are never monophyletic because *Cuticularia* sp. Linde, 1938, and *Poikilolaimus regenfussi* Sudhaus, 1980, are found more basal to the Rhabditina than the Diplogasteroidea and *Rhabditoides inermiformis* is often unresolved within the Rhabditida. There is a deep split between the remaining Rhabditomorpha; *Crustorhabditis scanica* (Allgén, 1949), *Distolabrellus veechi* Anderson, 1983, *Mesorhabditis anisomorpha* Sudhaus, 1978, *Parasitorhabditis* sp. Fuchs, 1937, *Pelodera teres* (Schneider, 1866), *Rhabditoides regina* Schulte & Poinar, 1991, and *Teratorhabditis synpapillata* Sudhaus, 1985, form one consistent group and *Caenorhabditis elegans, Choriorhabditis dudichi* Andrássy, 1970, *Cruznema tripartitum* Linstow, 1906, *Diploscapter* sp. Cobb, 1913, *Dolichorhabditis* sp. Andrássy, 1983, *Oscheius insectivora* (Körner, 1954), *Pellioditis marina* (Bastian, 1865), *Protorhabditis* sp. Osche, 1952, *Rhabditella axei* (Cobbold, 1884), *Rhabditis blumi* Sudhaus, 1974, *Rhabditis myriophila* Poinar, 1986, and the Strongyloidea form the other group. The Strongyloidea are always monophyletic. The Trichostrongylidae are monophyletic (32 – 66%, 1.00) under unweighted parsimony excluding positions with gaps and excluding positions with gaps and uninformative characters, under weighted parsimony and LogDet distances including all characters and excluding uninformative characters and under Bayesian inference, under the other models they are paraphyletic including *Strongylus equinus* (Müller, 1780) and *Syngamus trachea* (Montague, 1811).

In most respects, the molecular data set of the 18S rRNA gene agrees with recent phylogenetic reconstructions (de Ley & Blaxter, 2002). It firmly establishes the monophyly of the Nematoda and the existence of three main clades therein, the Enoplia, Dorylaimia and Chromadoria. However, there is no resolution for the relative relationships between these three subclasses other than that it is most likely that either the Enoplia or Dorylaimia are most primitive and that the Chromadorea are derived from one of those two taxa. There is also proof for the establishment of the Rhabditida as an order of the Chromadorea as opposed to its status as an independent clade of a higher rank. On a more specific level, the molecular data provide good estimates of the phylogenetic position of some taxa for which the closest relationships were previously either unknown or debatable.

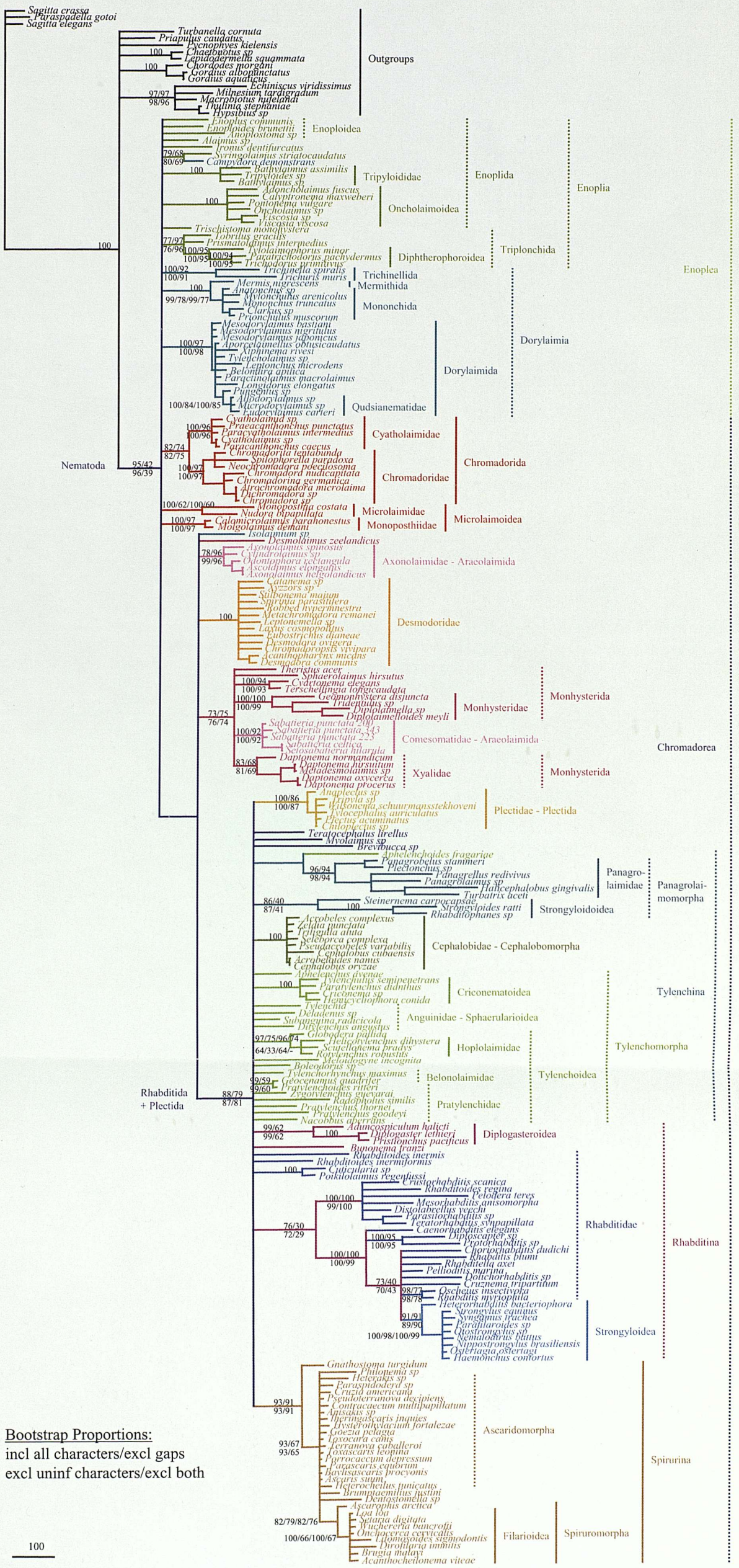


Figure 5.1: Strict consensus tree of the molecular character analyses under unweighted parsimony; Bootstrap Proportions (BP) of <65% support only given for major clades, no BP given for clades in monophyletic families; dotted lines depict paraphyletic or unresolved taxa

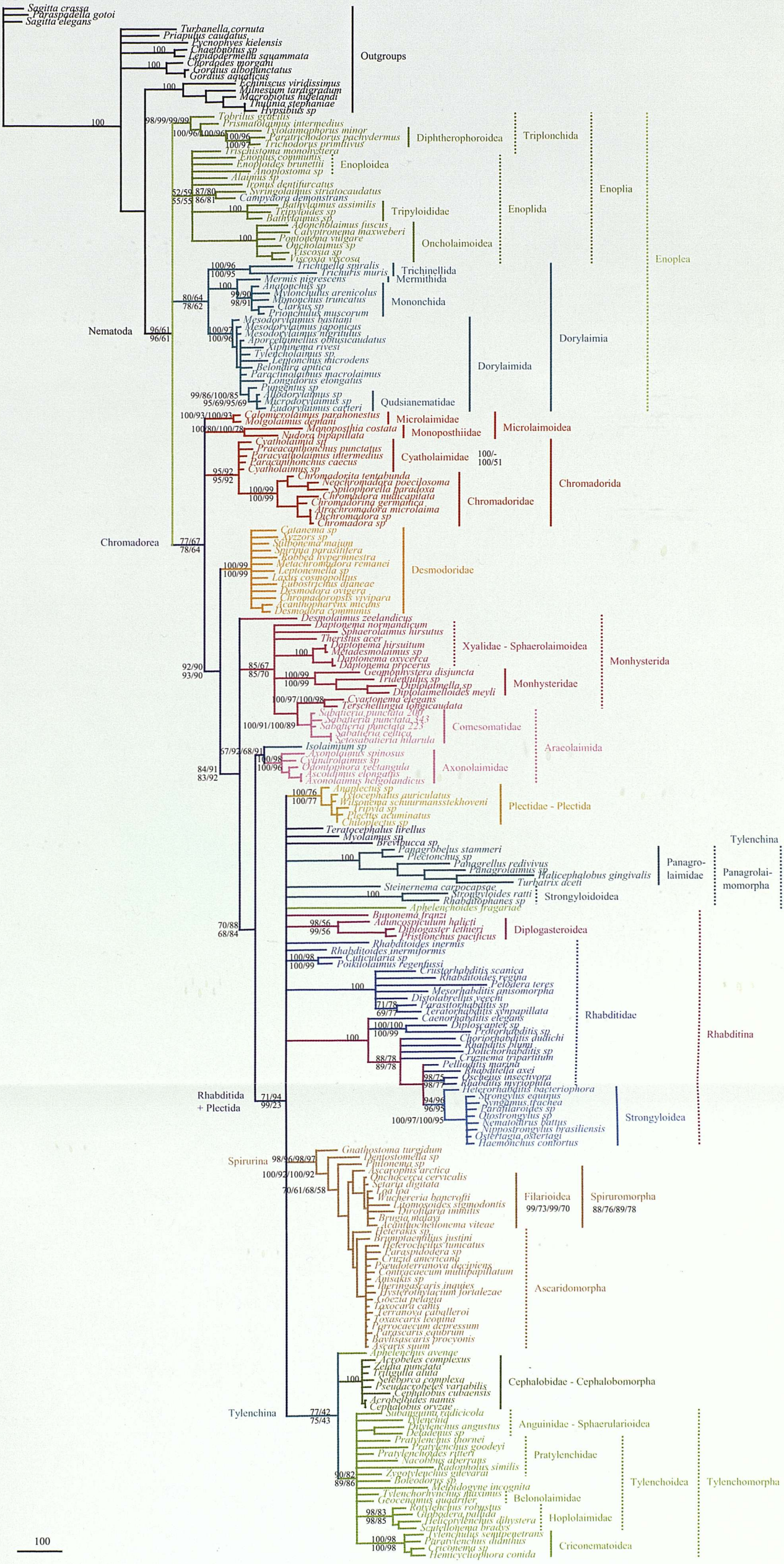


Figure 5.2: Strict consensus tree of the molecular character analyses under weighted parsimony; BP and legend as in Figure 5.1

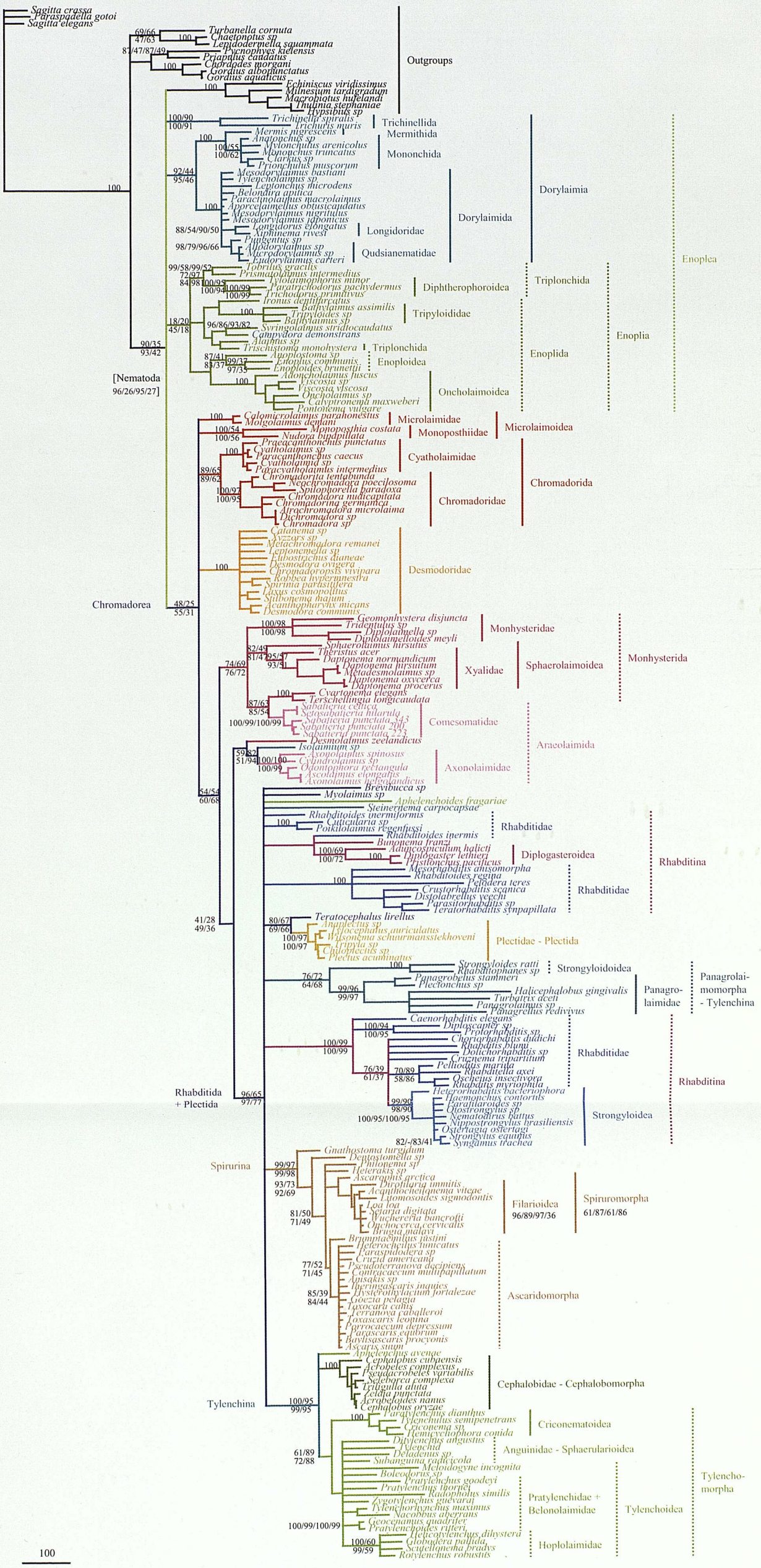


Figure 5.3: Strict consensus tree of the molecular character analyses under LogDet transformation; BP and legend as in Figure 5.1

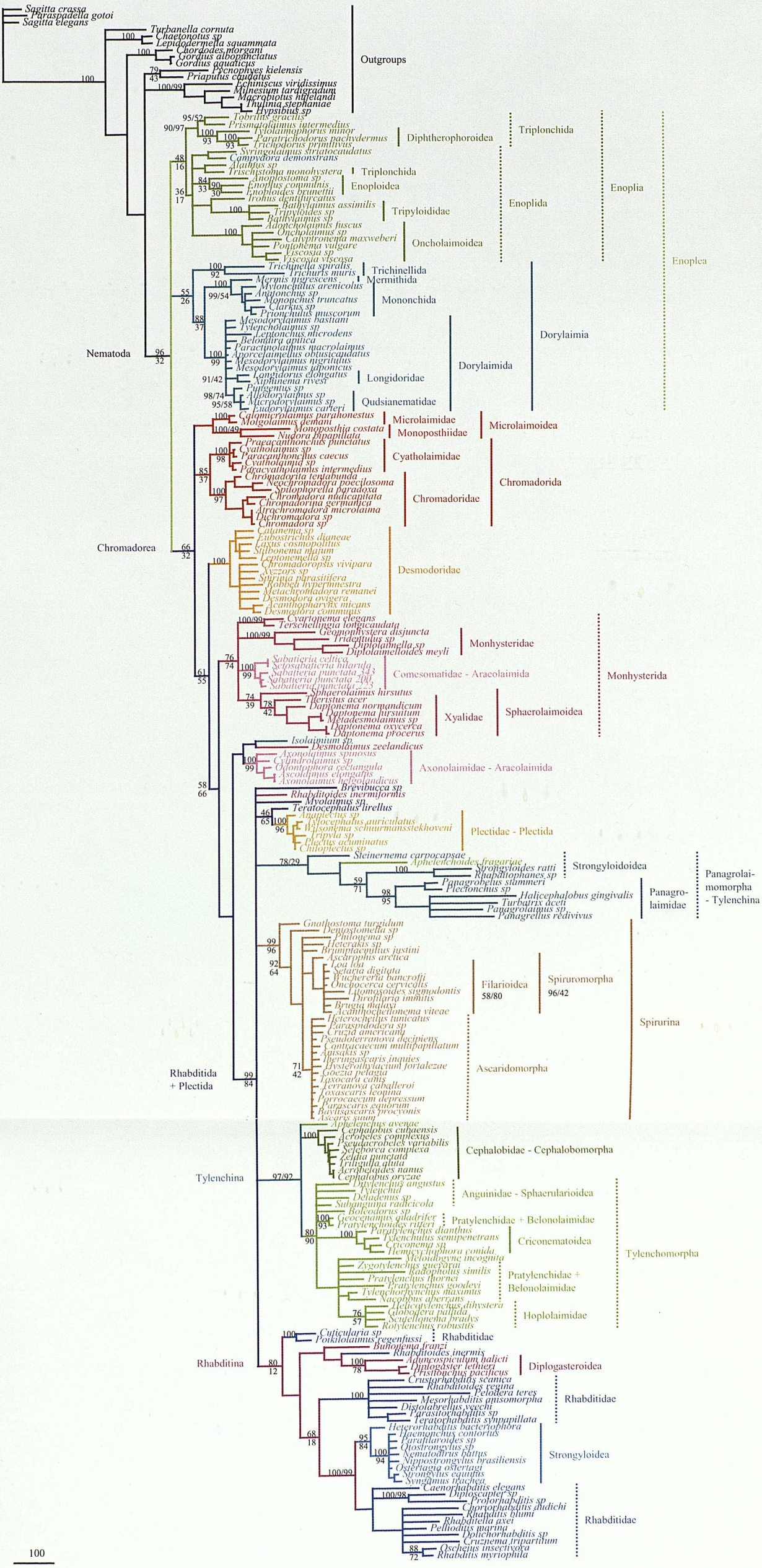


Figure 5.4: Strict consensus tree of the molecular character analyses under the maximum likelihood distance criterion; BP and legend as in Figure 5.1, but only for 'all' and 'excluding gaps'

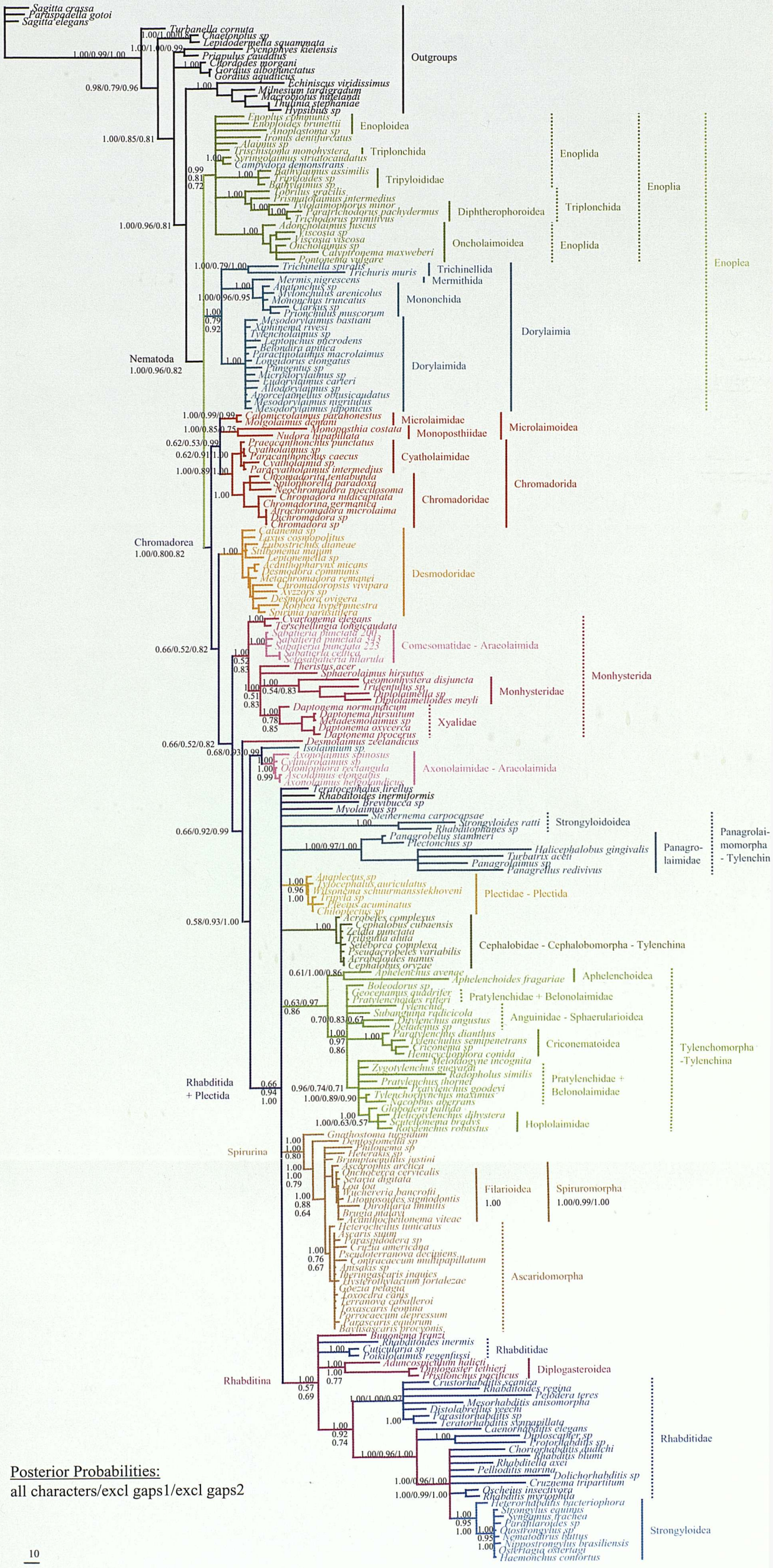


Figure 5.5: Strict consensus tree of the molecular character analyses under Bayesian inference; posterior probabilities (PP) of <0.65 only given for major clades, no PP given for clades in monophyletic families; dotted lines depict polyphyletic or unresolved taxa

6. Analysis of Morphological Characters

After an initial analysis of 180 characters, a suite of 159 characters was analysed for 167 nematode species. The analysis was performed only under the parsimony criterion because software for an analysis under the likelihood or Bayesian criterion is still under development. In the first analysis all 159 characters were included and in a second analysis 21 parsimonious-uninformative characters were excluded. Subsequently, both character sets were re-weighted under the rescaled consistency index (RC). The following interpretations are made using the strict consensus trees of all fundamental trees per analysis.

6.1 Differences based on the choice of model, character inclusion and bootstrap support

Figures 6.1 to 6.4 show the strict consensus trees for the four analyses of the morphological characters with the exception of Figure 6.2 which is the Adam consensus of the fundamental trees recovered analysing only parsimony-informative characters under equal character weights because the strict consensus tree gave little resolution. The numbers along the branches are the respective bootstrap values.

There is a lack of resolution in the analysis of informative characters alone under unweighted parsimony. This is possibly a result of the inclusion of uninformative characters increasing branch length for taxa with apomorphic characters. The exclusion of these characters results in a collapse of zero-length branches and a lack of resolution. However, weighting of informative characters increases resolution again, probably as a result of reduction in homoplasy (see Table 6.1).

Comparing the phylogenetic relationships between the four consensus trees shows that there is a lot of similarity between the results from the two analyses under character weighting and the two analyses excluding uninformative characters but less similarity between the topologies of the two trees under unweighted parsimony and the two including all characters. The major difference lies in the position of the Dorylaimia. They are only found as sister groups to the Enoplia under unweighted parsimony including all characters. In the other three cases they are found within the Chromadorea as sister group to the Rhabditida. A detailed discussion of the taxonomic differences between the four analyses is given in Section 6.2.

Table 6.1 summarises the total consistency and homoplasy content of the data as measured on the strict consensus trees of all fundamental trees for the four analyses of the morphological data. Under both weighting schemes the trees are slightly shorter and the three consistency indices lower after the 21 parsimony-uninformative characters are excluded from the data set. Fewer characters almost always produce shorter trees because the maximum number of character changes is lower and uninformative characters inflate the consistency index. The tree lengths and homoplasy content are both drastically reduced after characters are weighted on their rescaled consistency index because the weighting process reduces the weight assigned to characters that introduce homoplasy and the maximum possible number of character changes.

Table 6.1: Tree length (TL), consistency index (CI), CI without uninformative characters, retention index (RI), rescaled consistency index (RC), homoplasy index (HI), HI without uninformative characters and Goloboff-Fit (G-Fit) for all four analyses of the morphological data set, including all characters (all), excluding parsimony-uninformative characters (excl uninf), under unweighted (upars) and weighted (wpars) parsimony.

	TL	CI	CI excl uninf	RI	RC	HI	HI excl uninf	G-Fit
all, upars	1050	0.2343	0.2179	0.7151	0.1675	0.7657	0.7821	-69
all, wpars	158.51	0.5160	0.4380	0.8418	0.4343	0.4840	0.5620	-29
excl uninf, upars	1029	0.2177	0.2177	0.7147	0.1556	0.7823	0.7823	-69
excl uninf, wpars	142.69	0.4748	0.4748	0.8660	0.4112	0.5252	0.5252	-31

The significance of the bootstrap support is interpreted according to Hillis & Bull (1993). Clades with over 85% bootstrap support are very likely to be monophyletic. Clades with support between 65% and 85% are estimated to have moderate support for their monophyly and clades with less than 65% bootstrap support are discounted as monophyletic.

The bootstrap proportions are generally relatively similar for the two analyses under unweighted and weighted parsimony respectively. Although exclusion of uninformative characters increases homoplasy and reduces the resolution of the consensus tree under unweighted analyses the commonly reconstructed clades have similar support. Where there is a difference between the bootstrap supports, the value for the exclusion of uninformative characters is usually higher under unweighted parsimony; under weighted parsimony there is no pattern. Under most circumstances the difference is less than 5% and when the difference is higher the actual bootstrap proportions are below 65% and therefore the support is insignificant altogether. For the same character set, the bootstrap support is generally higher under weighted parsimony. If the support is higher under unweighted

parsimony, the difference is usually below 5% or the bootstrap support is below 65% and therefore insignificant.

There are only four exceptions to the above findings where the support for the alternative condition is significantly higher with over 65%: Under weighted parsimony the support for the (*Nacobbus aberrans* (*Globodera pallida* plus *Meloidogyne incognita*)) clade is significantly higher under the inclusion of all characters. Under the inclusion of all characters the support for the monophyly of *Sphaerolaimus* Bastian, 1865, is significantly higher under unweighted parsimony. Under the exclusion of uninformative characters the support for the Cyatholaimidae and for the *Chromadora nudicapitata* Bastian, 1865, plus *Chromadorina germanica* (Bütschli, 1874) clade is higher under unweighted parsimony.

6.2 Consistency of current model results with previous classifications

Contrary to the molecular analysis, the two classes and three subclasses of the Nematoda are never recovered as monophyletic groups using the 159 characters of the present analysis. However, many other relationships are still found in the majority of analyses including many families. Tables 6.2 and 6.3 list those groups that have been recovered consistently in all analyses; Table 6.2 contains traditional monophyletic taxa and Table 6.3 contains groups that do not belong to traditional monophyletic taxa or only form part of monophyletic groups. The taxa in each table are ordered according to Table A2.3 in Appendix II which gives the most recent classification of the phylum Nematoda according to de Ley & Blaxter (2002).

In the following discussion of the internal relationships of the Nematoda bootstrap values are only given for those groups that are not consistently recovered or where the support varies highly between the two weighting schemes because the values for consistently recovered taxa are already given in Tables 6.2 and 6.3.

The genus *Enoplus* Dujardin, 1845, is a monophyletic group. *Enoploides brunettii* Gerlach, 1953, is always the sister taxon to *Enoplus* and *Anoplostoma vivipara* (Bastian, 1865) is their sister taxon. Therefore, the Enoplina/Enoploidea are also a monophyletic group using the current taxa and characters. The Oncholaimidae are probably the sister taxon to the Enoploidea. *Adoncholaimus fuscus* and *Viscosia viscosa* (Bastian, 1865) are always sister taxa and *Pontonema vulgare* (Bastian, 1865) is closely related but they never form a monophyletic group. However, *Calyptronema maxweberi*, the other Oncholaimoidea taxon, is found more closely related to the Dorylaimia and therefore the Oncholaimina/Oncholaimoidea are paraphyletic. The two taxa of the Ironidae, *Ironus*

dentifurcatus and *Syringolaimus striatocaudatus*, also do not form a sister group relationship. *Ironus* is almost always found with other Dorylaimia, most often as sister taxon to *Isolaimium papillatum* Cobb, 1920, *Mermis nigrescens* and the Trichinelloidea. Under weighted parsimony *Syringolaimus striatocaudatus* is found as the sister taxon to *Metachromadora remanei* (12 – 18% support) situated closely to the Dorylaimia; under unweighted parsimony it is closely related to the Rhabditida. The Tripyloididae are always monophyletic and found closely related to the Enoploidea and Oncholaimidae but not directly as their sister taxon. Therefore, the order Enoplida is non-monophyletic in all analyses.

Table 6.2: Classic monophyletic groups consistently recovered in at least three of the four analyses of the morphological character sets. Range in values of bootstrap proportions in parenthesis.

Enoplidae (single genus *Enoplus*) (39 – 59%)
Enoplina / Enoploidea (70 – 92%)
Tripyloidina / Tripyloidoidea / Tripyloididae (95 – 99%)
Trichodoridae (84 – 98%)
Dorylaimida (excluding *Campydora demonstrans*) (39 – 62%)
Dorylaimina (39 – 62%)
Longidoridae (61 – 65%)
Xiphinematidae (single genus *Xiphinema*) (89 – 98%)
Mononchidae (74 – 92%)
Trichinellida/Trichinelloidea (95 – 97%)
Chromadoridae (45 – 68%)
Cyatholaimidae (paraphyletic including all characters, weighted) (39 – 67%)
Laxus (97 – 99%)
Microlaimidae (unresolved excluding uninformative characters, unweighted) (10 - 19%)
Monoposthiidae (91 – 98%)
Sphaerolaimoidea (44 – 69%)
Sphaerolaimidae (63 – 78%)
Comesomatidae (58 – 61%)
Sabatieria (92 – 94%)
Axonolaimus (68 – 84%)
Plectida (92 – 98%)
Rhabditida (3 – 20%)
Spiruromorpha (38- 61%)
Filarioidea (including *Setaria digitata*) (48 – 84%)
Cephalobomorpha / Cephaloboidea / Cephalobidae (78 – 92%)
Tylenchomorpha (83 – 98%)
Criconematoidea (67 – 79%)
Sphaerularioidea/Anguinidae (76 – 91%)

The Trichodoridae are always monophyletic and tend to form a sister taxon to the majority of the Dorylaimia. *Tylolaimophorus minor* (Thorne, 1939) is the sister taxon to the Trichodoridae in the two analyses under weighted parsimony (48 – 58% support).

Under weighted parsimony *Tylolaimophorus minor* is found more closely related to the Dorylaimia making the Diphtherophoroidea polyphyletic. The two species of the Tobrilina, *Tobrilus gracilis* (Bastian, 1865) and *Prismatolaimus intermedius* (Bütschli, 1873), are never sister taxa. *Tobrilus gracilis* is usually found closely related to the Oncholaimidae while *Prismatolaimus intermedius* is found as the sister taxon to the Cyatholaimidae (Chromadorea!) in the analyses of all characters (47% support under weighted parsimony). Excluding uninformative characters leaves the relationships of *Prismatolaimus intermedius* unresolved. The position of *Trischistoma monohystera* (Tripylina) is entirely inconsistent. It is found in different positions in the Enoplea and in the analysis of all characters under weighted parsimony it is found with the non-rhabditid Chromadorea. Therefore, the order Triplonchida, like the Enoplida, is non-monophyletic, rendering the subclass Enoplia also non-monophyletic.

Table 6.3: Relationships that do not form traditional monophyletic groups but have been recovered consistently in the four analysis of the morphological data set.

(*E. meridionalis* (*E. communis* and *E. anisospiculum*)) (43 – 71%)
Enoploides brunettii and Enoplidae (81 – 85%)
Adoncholaimus fuscus and *Viscosia viscosa* (25 – 64%)
Chromadora nudicapitata and *Chromadorina germanica* (52 – 68%)
Chromadoropsis vivipara and *Desmodora communis* (37 - 50%)
Eubostrichus topiarus and *Laxus* (41 – 60%)
Eubostrichus parasiticus, E. topiarus and *Laxus* (51 – 63%)
Daptonema setosum and *Theristus acer* (38 – 44%)
Plectus acuminatus and *P. aquatilis* (85 – 89%)
Tylocephalus auriculatus and *Wilsonema schuurmansstekhoveni* (85 – 90%)
Gnathostoma turgidum and *Goezia pelagia* (50 – 99%)
Acanthocheilonema viteae and *Brugia malayi* (53 – 74%)
Baylisascaris procyonis and *Toxascaris leonine* (57 – 61%)
Tylenchomorpha without Aphelenchoidea (58 – 63%)
Criconematoidea including *Paratylenchus dianthus* (52 – 75%)
Globodera pallida and *Meloidogyne incognita* (55 – 74%)
Nacobbus aberrans and (*Globodera pallida* and *Meloidogyne incognita*) (46 – 68%)
Pratylenchoides ritteri and *Radopholus similes* (38 - 48%)
Haemonchus contortus and *Ostertagia ostertagi* (59 – 86%)
Nematodirus battus and *Nippostrongylus brasiliensis* (77 – 83%)

In the Dorylaimia relationships are a little more stable but the subclass is also non-monophyletic. *Campydora demonstrans* (Campydorina) is found in different positions but always near the base of the Rhabditida. The remaining Dorylaimina/Dorylaimida form a monophyletic clade but with insignificant support. The Longidoridae and the genus *Xiphinema* are always monophyletic. *Mesodorylaimus bastiani* (Bütschli, 1873) and *Paractinolaimus macrolaimus* (de Man, 1880) (47% support under unweighted, 71 – 72%

support under weighted parsimony) and the Longidoridae are sister taxa respectively. *Aporcelaimellus obtusicaudatus* (Bastian, 1865), *Eudorylaimus carteri* (Bastian, 1865) and *Belondira apitica* Thorene, 1939, are found in different positions but always within the same clade as the aforementioned species. The Dorylaimoidea are monophyletic including all unweighted characters and paraphyletic in all other analyses as they contain *Belondira apitica* (Belondiroidea).

The Mononchidae are always monophyletic. They are the sister taxon to *Mylonchulus arenicolus* Clark, 1961, (34% support) and together the sister taxon to all other Dorylaimia when the characters are treated as unweighted. Weighting of all characters leaves the Mononchidae as sister taxon to the Dorylaimida (54% support) and *Mylonchulus* more basal to the Dorylaimia plus Rhabditida while character weighting under the exclusion of uninformative characters renders both families as sister taxa to the Dorylaimida but the Mononchidae slightly more closely related to the latter (61% and 28% support respectively). *Isolaimium papillatum* (Isolaimida) and *Mermis nigrescens* are always found within the Dorylaimia but in different sister relationships in the four analyses and with little support for the various relationships. The Trichinellida are always monophyletic and also always found within the Dorylaimia but in different sister relationships in the four analyses.

The subclass Enoplea is only recovered as a monophyletic clade but with little support when all characters are included and treated as unweighted with the exclusion of *Syringolaimus striatocaudatus* and *Campydora demonstrans*. In all other analyses the Dorylaimia are found as a sister taxon to the Rhabditida or at least part of the Dorylaimia are closest related to the Rhabditida. The findings also show that the other class, the Chromadorea, are non-monophyletic apart from the analysis of all characters under unweighted parsimony.

The Chromadoridae are always monophyletic and the Cyatholaimidae are monophyletic with the exception of weighting all characters. However, the Chromadoroidea are always polyphyletic because its two families are found in different relationships amongst the majority of freeliving Chromadorea. *Cyartonema elegans*, previously of uncertain relationship, is found either basal to the Rhabditida together with *Metachromadora remanei*, *Syringolaimus striatocaudatus* and *Campydora demonstrans* (all characters unweighted) or as sister taxon to the Monhysterida (both analyses with weighted characters). Excluding uninformative characters leaves this species in an uncertain position. The Desmodoridae are polyphyletic. *Metachromadora remanei* is

always found more closely related to the Rhabditida, either close to *Syringolaimus striatocaudatus* (including all characters) or as its sister taxon (weighted characters, 12 – 18% support). The remaining representatives of the Desmodoridae form a monophyletic clade apart from the analysis of weighted informative characters only when they are split into three groups amongst the majority of freeliving Chromadorea. This also renders the Desmodorida polyphyletic. The Microlaimidae are always monophyletic except when uninformative characters are excluded under weighted parsimony (10 – 19% support); in this case their internal relationships are unresolved. The Microlaimidae are the sister taxon to the Monhysterida plus Araeolaimida under unweighted parsimony and they are the sister taxon to the Araeolaimida under weighted parsimony. However, there is little support for these relationships. The Monoposthiidae are always monophyletic. They are the sister taxon to the Monhysterida (excluding the Monhysteridae) under weighted parsimony. Under unweighted parsimony including all characters they are the nearest taxon to the Plectida plus Rhabditida. When uninformative characters are excluded under unweighted parsimony the position of the Monoposthiidae is unresolved within a clade containing the majority of the freeliving Chromadorea.

The Sphaerolaimidae and Sphaerolaimoidea are always monophyletic. The genus *Daptonema* Cobb, 1920 is paraphyletic as *Daptonema setosum* (Bütschli, 1874) is most closely related to *Theristus acer* Bastian, 1865. Under unweighted parsimony *Daptonema hirsuitum* (Vitiello, 1967) is unresolved within the Sphaerolaimoidea leaving the Xyalidae also unresolved under these conditions. Under weighted parsimony the Xyalidae are monophyletic (45 – 46% support). The Linhomoeidae are the sister taxon to the Sphaerolaimoidea but only monophyletic when all characters are treated as unweighted (35% support); in all other cases they are paraphyletic. The Monhysteridae are never closely related to the remaining Monhysterida and the latter is therefore polyphyletic. The Monhysteridae are always found more closely related to the Dorylaimia but in varying positions. *Diplolaimella stagnosa* Lorenzen, 1966, and *Geomonhystera disjuncta* are sister taxa (61 – 65% support) except under weighted parsimony excluding uninformative characters; in this case *Diplolaimella stagnosa* is more closely related to *Diplolaimelloides meyli* Timm, 1961. *Diplolaimelloides meyli* only forms a sister relationship with the other two taxa when the characters are weighted (66 – 70% support).

The Comesomatidae and the genus *Sabatieria* Rouville, 1903, are always monophyletic. They are the sister taxon to the monophyletic Axonolaimidae when the characters are weighted rendering the Araeolaimida also monophyletic (14 – 20% support).

In these cases the Araeolaimida are the sister taxon to the Monoposthiidae which together form the sister taxon to the Monhysterida and Microlaimidae, however, none of these relationships are significantly well supported. Otherwise the Axonolaimidae are unresolved forming an unresolved clade with the Comesomatidae plus Monhysterida (not including the Microlaimidae and Monoposthiidae).

The Plectida / Plectoidea / Plectidae are always monophyletic. They are either the closest sister taxon to the Rhabditida (all characters unweighted) or the sister taxon to the Chromadoridae (weighting of characters). Under unweighted parsimony excluding uninformative characters the position of the Plectida is unresolved. None of the relationships of the Plectida are well supported.

The order Rhabditida is always monophyletic and either the sister taxon to the Plectida (all characters unweighted) or to the Dorylaimia. Within the Rhabditida, however, the relationships vary as much as in the non-Rhabditida taxa. Of the higher taxa, only the Spiruromorpha, Cephalobomorpha and Tylenchomorpha are always monophyletic. *Teratocephalus lirellus* is found near the base of the Rhabditida (including all characters), usually as sister taxon to *Panagrellus redivivus* (Linnaeus, 1767) (15 – 19% support, not if all characters are weighted) and closely related to the remaining Panagrolaimidae. When uninformative characters are excluded (under both weighting schemes) *Teratocephalus lirellus* and the Panagrolaimidae form the sister taxa to the Cephalobidae.

The Spiruromorpha are always monophyletic. They always form one clade with the Ascaridoidea and Strongyloidea but the relationships between these groups vary between all analyses. None of the clades are significantly supported either. Within the Spiruromorpha, *Ascarophis arctica* is always the most basal taxon. The remaining Spiruromorpha are only insignificantly supported under unweighted parsimony (48 – 49%) but moderately-well supported under weighted parsimony (82 -84%). The Filarioidea and Onchocercidae are paraphyletic as they include the taxon of uncertain position, *Setaria digitata* (Setariidae). *Loa loa* (Filariidae) is also part of the Onchocercidae.

The Ascaridomorpha are polyphyletic as *Cruzia americana* Rudolphi, 1819, is usually found at the base of the Spirurina and not in the same clade as the Ascaridoidea except under weighted parsimony excluding uninformative characters. In the latter case, *Cruzia americana* is found as the sister taxon to the Ascaridoidea. In both analyses with unweighted characters, *Cruzia americana* is found as the sister taxon to *Steinernema carpocapsae*. The Ascaridoidea themselves are paraphyletic as *Gnathostoma turgidum* (Gnathostomatomorpha) is always found as the sister taxon to *Goezia pelagia* Deardorff &

Overstreet, 1980. None of the families of the Ascaridoidea represented in this study (Anisakidae, Ascarididae, Rhaphidascaridae) are monophyletic. The variable position of *Cruzia americana* and the possible inclusion of the Strongyloidea make the suborder Spirurina paraphyletic.

The Tylenchina are always paraphyletic because the Panagrolaimomorpha are always polyphyletic. *Panagrobelus stammeri* Rühm, 1956, is always the sister taxon to the Cephalobidae. *Panagrellus redivivus* is the sister taxon to *Teratocephalus lirellus* (15 – 19% support) except under weighted parsimony including all characters; they are always found close to the base of the Rhabditida. Both these taxa together with *Turbatrix aceti* (Müller, 1783) form one clade with *Panagrobelus stammeri* and the Cephalobidae when uninformative characters are excluded (under both weighting schemes). *Steinernema carpocapsae* and *Strongyloides stercoralis* are only sister taxa under weighted parsimony excluding uninformative characters (45% support) rendering the Strongyloidoidea polyphyletic. When all characters are included (under both weighting schemes) the two species are found basal to the Spirurina. *Halicephalobus gingivalis* (Stefanksi, 1954) is found in various positions amongst the Rhabditina. The Cephalobidae are always monophyletic. They are found either closely related to the Rhabditidae (under unweighted parsimony including all characters) or more basally to the Rhabditida amongst some of the Panagrolaimidae.

The Tylenchomorpha are always monophyletic but they are found in varying positions within the Rhabditida. The Aphenchoidea are always the most basal taxon of the Tylenchomorpha. They are paraphyletic if the characters are unweighted and monophyletic when the characters are weighted (52 – 58% support). The Criconematoidea, Sphaerularioidea and Tylenchoidea are always monophyletic with the exception of the Tylenchoidea under unweighted parsimony; in the latter case they are paraphyletic including the other two superfamilies. The Criconematoidea are always the sister taxon to *Paratylenchus dianthus*; together they are the sister taxa to the Tylenchoidea when characters have been weighted (28 – 29% support). Under unweighted parsimony including all characters the Criconematoidea and *Paratylenchus dianthus* are the sister taxon to the Sphaerularioidea but with insignificant support. The Sphaerularioidea are the sister taxon to the other two superfamilies under weighted parsimony (38 – 63% support) and they are the sister taxon to the Criconematoidea plus *Paratylenchus dianthus* if all characters are included and unweighted. The Tylenchoidea are monophyletic and the sister taxon to the Criconematoidea plus *Paratylenchus dianthus* when characters have been

weighted (28 – 29% support). When characters are treated as unweighted they are mixed with the other superfamilies. The taxa of the Hoplolaimidae and Pratylenchidae are always mixed together rendering these families paraphyletic. The Belonolaimidae are monophyletic when characters are treated as unweighted (64 – 65%) and paraphyletic as the basal taxa of the Tylenchoidea when characters are weighted. *Globodera pallida* and *Meloidogyne incognita* are always sister taxa and together the sister taxa to *Nacobbus aberrans*. These three taxa are always found within the clade of the Hoplolaimidae and Pratylenchidae.

The Rhabditina are always polyphyletic. *Bunonema franzi* (Bunonematomorpha) is always found with the Rhabditidae except under unweighted parsimony excluding uninformative characters when its position in the Rhabditida is unresolved. The Diplogasteroidea (Diplogasteromorpha) are always paraphyletic. They are the sister taxa to the Tylenchomorpha in all cases apart from when all characters are included and unweighted; in this case *Pristionchus pacificus* (Neodiplogasteridae) is the basal taxon to a clade of Rhabditidae and Cephalobidae while *Aduncospiculum halicti* (Diplogasteridae) is basal to a clade containing the aforementioned clade plus the Tylenchomorpha.

The Rhabditoidea are always at least paraphyletic including *Heterorhabditis bacteriophora* (Strongyloidea). When all characters are included and unweighted they also include the Cephalobidae and if uninformative characters are excluded and the remaining characters weighted they also include *Bunonema franzi*. The intra-relationships of the Rhabditidae are highly variable as well as their position within the Rhabditida. The clades also find little bootstrap support. The Strongyloidea are never monophyletic because *Heterorhabditis bacteriophora* is always found with the Rhabditidae. The remaining Strongyloidea are always found most closely related to the Spirurina and therefore the Rhabditomorpha are polyphyletic. The Strongyloidea (excluding *Heterorhabditis bacteriophora*) form a single clade as sister taxon to the Spiruromorpha or Spirurina when characters are weighted but they have little support (< 31%). Within the Trichostrongylidae, *Haemonchus contortus* (Rudolphi, 1803) and *Ostertagia ostertagi* (Stiles, 1892) as well as *Nematodirus battus* Croften & Thomas, 1951, and *Nippostrongylus brasiliensis* (Travassos, 1914) respectively form sister taxa relationships but the family is never monophyletic. If all characters are included and unweighted *Syngamus trachea*, *Nematodirus battus* and *Nippostrongylus brasiliensis* are the sister clade to the Ascaridoidea and *Strongylus equinus*, *Haemonchus contortus* and *Ostertagia ostertagi* the sister clade to the Spirurina. If uninformative characters are excluded and

characters treated as unweighted, these six species of the Strongyloidea are found in unresolved positions in one clade with the Spirurina.

In summary, many of the classic relationships of the Nematoda are not reconstructed with the current data set despite the use of the same characters in both the classic, intuitive analysis and mathematical modelling. In fact, some relationships resemble more closely the biology of the species. For example, in all analyses but the one including all characters without weighting a cut can be made between mostly aquatic species and mostly terrestrial and parasitic species. Similarly, the close relationship of the Strongyloidea with the Spirurina reflects the common habitat as internal vertebrate parasites. In some preliminary analyses when individual characters were tested by respectively including or excluding them, the Trichinelloidea were also found in one group with the Spirurina.

These results are particularly interesting because an objective, mathematical system is expected to reflect the true evolutionary history more truthfully. However, the failure to recover so many established groups highlights the problem of homology estimations during character coding. Although characters with dubious homology were tested and excluded, a more thorough analysis of characters will have to be undertaken in a future analysis.

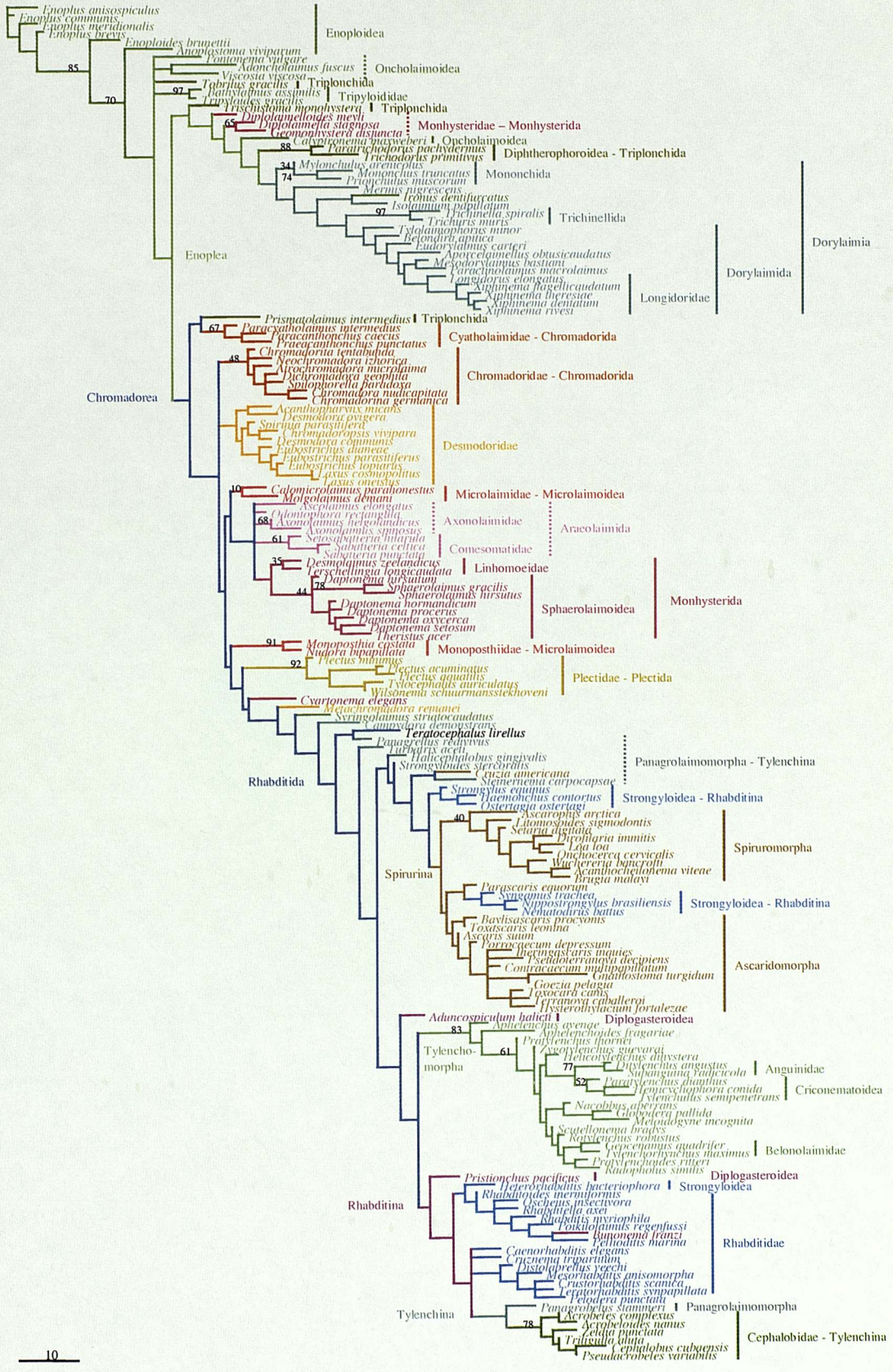


Figure 6.1: Strict consensus tree of the morphological character analysis under unweighted parsimony including all characters; BP of < 65% support only given for major clades; dotted lines depict polyphyletic or unresolved taxa

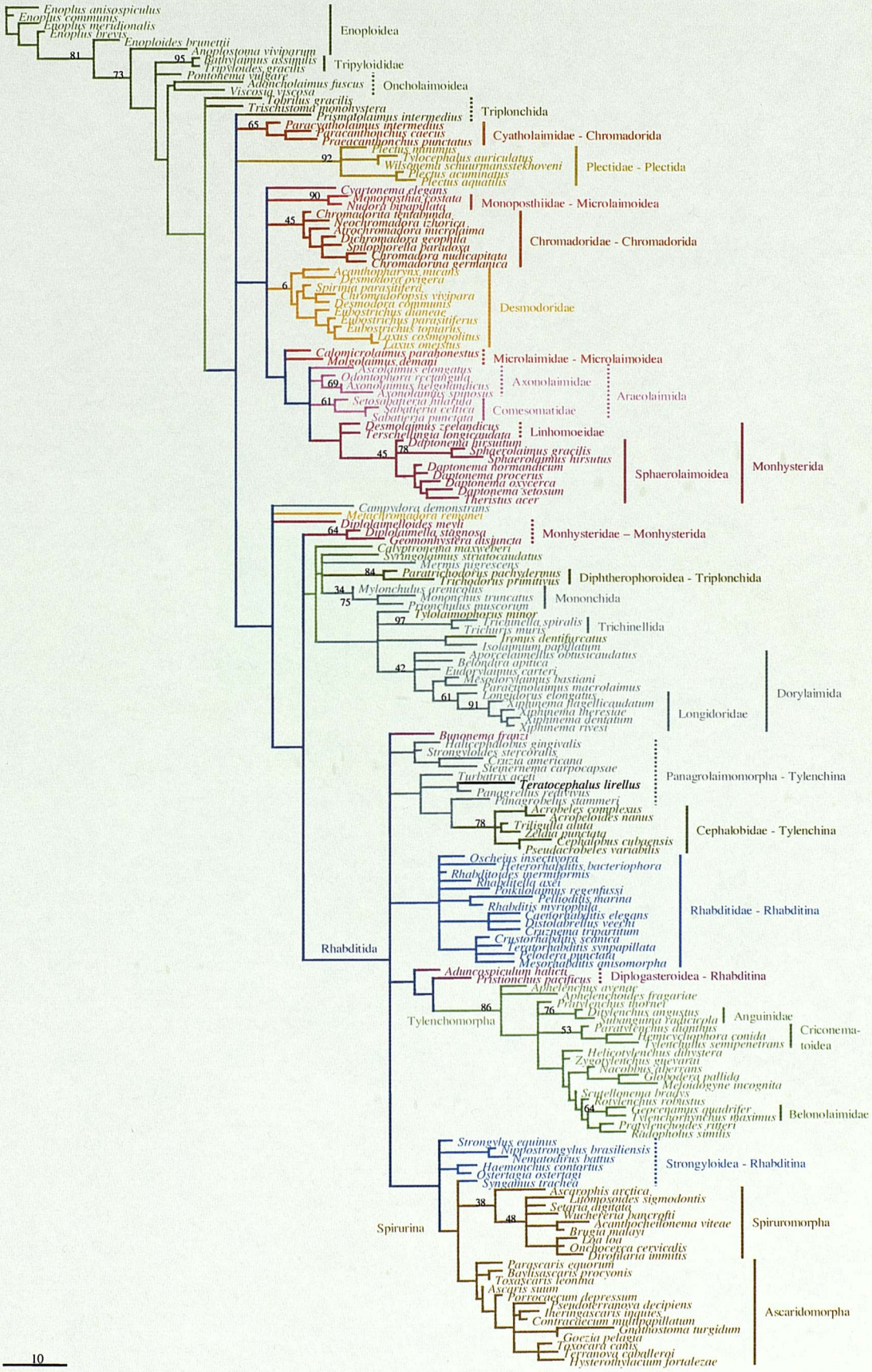


Figure 6.2: Adam consensus tree of the morphological character analysis under unweighted parsimony excluding all uninformative characters; BP and lines as in Fig. 6.1

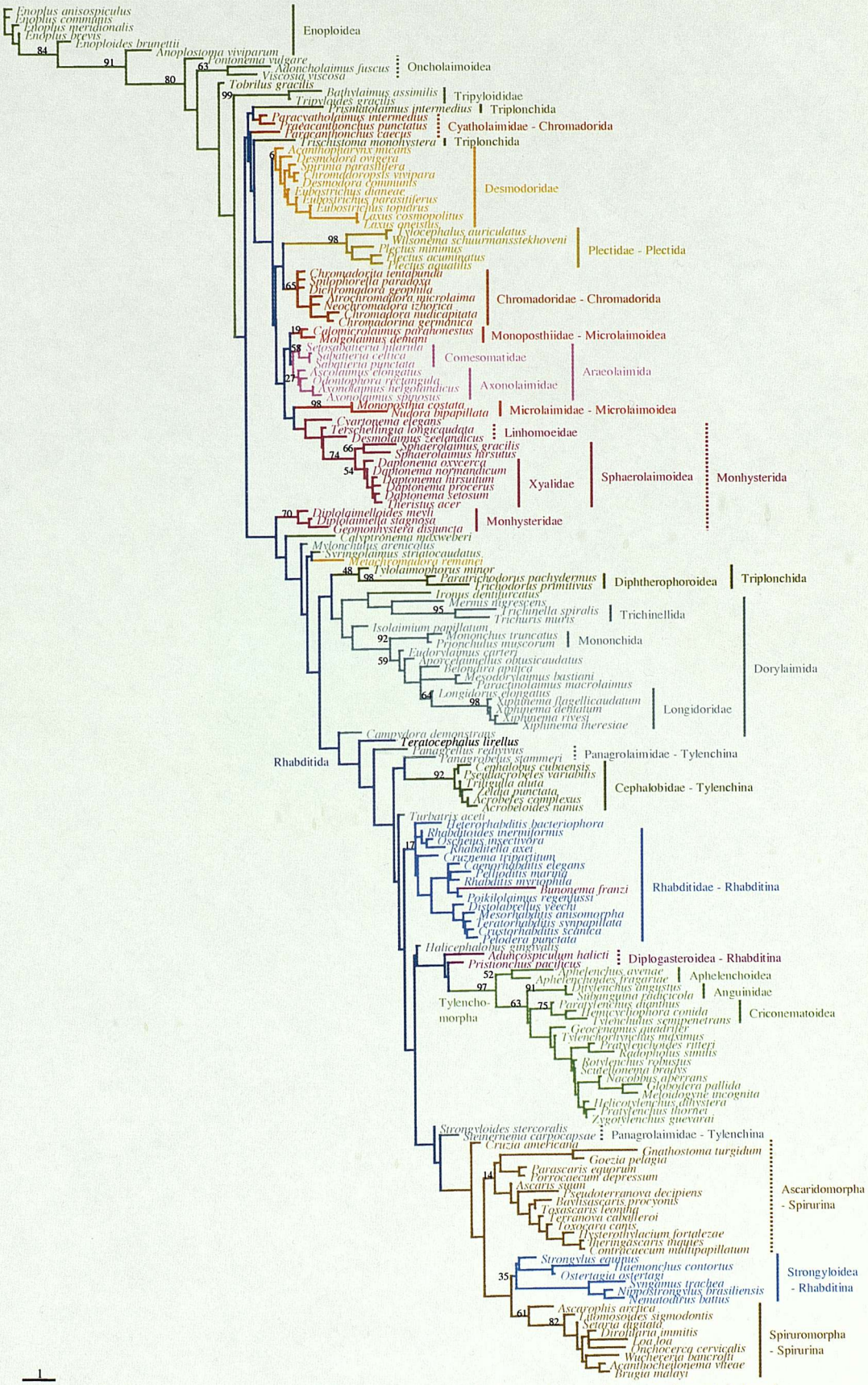


Figure 6.3: Strict consensus tree of the morphological character analysis under weighted parsimony including all characters; BP and lines as in Fig. 6.1

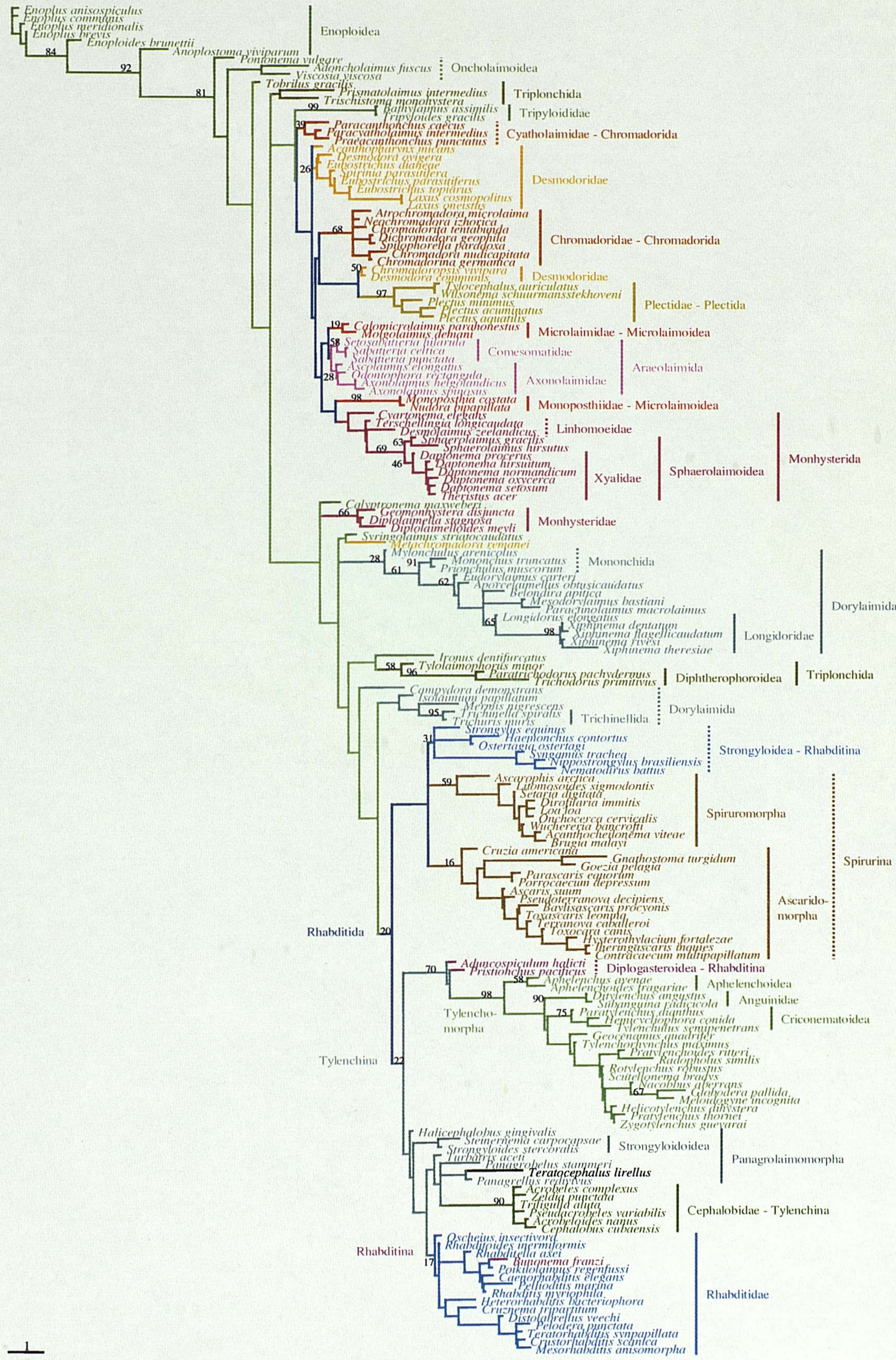


Figure 6.4: Strict consensus tree of the morphological character analysis under weighted parsimony excluding all uninformative characters; BP and lines as in Fig. 6.1

7. Analysis of the Combined Character Set

All characters from the molecular and morphological analyses where combined in one single data set and analysed under unweighted and weighted parsimony (on the rescaled consistency index, RC). Providing both sets of data are compatible (see below) the simultaneous analysis of all characters presents the possibility of a more refined analysis. The combined analysis can increase the phylogenetic signal if the characters from the different independent data sets provide evidence for different parts of the phylogeny. A detailed discussion of the advantages of a combined analysis is given in the Introduction (see Section 1.2.7.2).

Four character sets were created: using all 1928 characters, excluding all positions with gaps in the molecular data (752 characters remaining), excluding all parsimony-uninformative characters (1068 characters remaining) and excluding all positions with gaps in the molecular data and all parsimony-uninformative characters (440 characters remaining). The same 16 outgroup taxa were included as in the analysis of the molecular data, however, no morphological characters were available and gaps were included instead. A combined analysis was justified because no significant incongruence was found between the two character sets using the 'Incongruence Length Difference Test' (Farris *et al.*, 1995). The interpretations are made on the strict consensus trees of all fundamental trees recovered in each analysis.

7.1 Differences based on the choice of model, character inclusion and bootstrap support

Figures 7.1 to 7.8 show the strict consensus trees for the eight different analyses. All consensus trees are very well resolved apart from Figure 7.2 constructed under unweighted parsimony excluding positions with gaps where the relationships amongst the families of the Enoplea are unresolved. A comparison between the tree topologies shows a high degree of congruence. The Nematoda are consistently recovered as a monophyletic taxon. The class Enoplea and its two subclasses, Enoplia and Dorylaimia, are never recovered as a monophyletic clade but the other class, Chromadorea, is always monophyletic. Major differences are only found amongst the relationships of the different orders of the Enoplea, the families of the Enoplida, the position of the Plectida relative to the Rhabditida and the position of the Strongyloidoidea amongst the Rhabditida. The internal relationships of the Monhysterida are identical in all analyses. Overall, the phylogenetic reconstructions are more similar to those using molecular characters alone but some aspects of the

morphological topologies are also recovered. A detailed analysis of the topologies is found in Section 7.2.

Table 7.1 shows the details of the consistency and homoplasy indices for the strict consensus trees of the eight analyses of the combined data set. The behaviour of the indices is comparable to the findings for the analysis of both data sets separately. The trees are shorter when fewer characters are included and when characters are weighted on their rescaled consistency index because both actions reduce the homoplasy of the data and therefore also reduced the maximum possible tree length. The consistency indices increase – and the homoplasy index decreases - after character weighting because homoplastic characters are given less weight. The exclusion of parsimony-uninformative characters decreases the consistency index slightly as the inclusion of such characters inflates this index because it is dependent on the number of characters in the data set. The exclusion of positions with gaps in the molecular data set increased the consistency indices and reduced the homoplasy index because at such positions levels of homoplasy tend to be high as a result of difficulties in alignment. The highest consistency index is found when only positions with gaps are excluded but because uninformative characters inflate the index the best possible analysis might be the one excluding both sets of potentially ambiguous characters.

Table 7.1: Tree length (TL), consistency index (CI), CI without uninformative characters, retention index (RI), rescaled consistency index (RC), homoplasy index (HI), HI without uninformative characters and Goloboff-Fit (G-Fit) for all four analyses of the combined data set, including all characters (all), excluding positions with gaps in molecular data (excl gaps), excluding parsimony-uninformative characters (excl uninf) and excluding positions with gaps and uninformative characters, under unweighted (upars) and weighted (wpars) parsimony.

	TL	CI	CI excl uninf	RI	RC	HI	HI excl uninf	G-Fit
all, upars	14331	0.1875	0.1702	0.6398	0.1200	0.8125	0.8298	-414
all, wpars	1624.80	0.4381	0.3049	0.7542	0.3289	0.5639	0.6906	-132
Excl gaps, upars	4130	0.2412	0.2149	0.7130	0.1720	0.7588	0.7851	-199
Excl gaps, rpars	647.47	0.5064	0.3726	0.8278	0.4192	0.4936	0.6274	-67
Excl uninf, upars	14035	0.1702	0.1702	0.6397	0.1089	0.8298	0.8298	-141
Excl uninf, wpars	1326.53	0.3098	0.3098	0.7544	0.2337	0.6902	0.6902	-132
excl gaps uninf, upars	3992	0.2149	0.2149	0.7130	0.1533	0.7851	0.7851	-199
excl gaps uninf, wpars	506.76	0.3699	0.3699	0.8286	0.3065	0.6301	0.6301	-67

The significance of the bootstrap support is interpreted according to Hillis & Bull (1993). Clades with over 85% bootstrap support are very likely to be monophyletic. Clades with support between 65% and 85% are estimated to have moderate support for their

monophyly and clades with less than 65% bootstrap support are discounted as monophyletic.

The bootstrap support is generally higher in the analyses under weighted parsimony compared to unweighted parsimony under the same character inclusions. This might be a consequence of the reduction of homoplasy in the data resulting in a more consistent analysis. As in the separate analyses of the two data sets the bootstrap proportions are very similar between the two analyses including positions with gaps and excluding such positions. Under unweighted parsimony the bootstrap support is more than 5% higher amongst the alternative conditions at a level above 65% for the following groups: the Nematoda and (*Syringolaimus striatocaudatus* plus *Campydora demonstrans*) when uninformative characters are excluded compared to when all characters are included; Rhabditidae plus Strongyloidea when all characters are included compared to excluding uninformative characters; and Hoplolaimidae plus *Globodera pallida* when positions with gaps are excluded compared to when positions with gaps and uninformative characters are excluded. Under weighted parsimony the support for the genus *Daptonema* is significantly higher when uninformative characters are excluded compared to when all characters are included.

Table 7.2: Taxa and groups that have higher bootstrap support when positions with gaps are excluded. Significant = support above 65%, insignificant = support below 65%, blank = support higher when positions with gaps included, inapplicable = clade does not exist in this analysis.

Taxon	**Unweighted Parsimony**	**Weighted Parsimony**
Enoploidea	Insignificant	Significant
Diphtherophoroidea + *Prismatolaimus intermedius*	Inapplicable	Significant
Mononchidae	Significant	Significant
Monhysterida + Comesomatidae	Significant	Significant
Monhysterida	Insignificant	Significant
Monhysteridae + Sphaerolaimoidea + *Desmolaimus zeelandicus*	Insignificant	Significant
Sphaerolaimoidea + *Desmolaimus zeelandicus*	Significant	Significant
Sphaerolaimidae + *Desmolaimus zeelandicus*	Significant	Significant
Xyalidae	Significant	
Filarioidea		Significant
Ascaridomorpha	Insignificant	
Halicephalobus gingivalis + *Turbatrix aceti*	Significant	
Tylenchomorpha	Inapplicable	Significant
Aphelenchoidea	Inapplicable	Significant
Criconematoidea	Significant	Significant
Hoplolaimidae + *Globodera pallida*		Significant
Strongyloidea	Significant	Significant
Trichostrongylidae	Significant	
Haemonchus contortus + *Ostertagia ostertagi*	Significant	
Nematodirus battus + *Nippostrongylus brasiliensis*	Significant	

The exclusion of positions with gaps generally reduces the bootstrap support. Table 7.2 lists those groups that have more support at a level above 65% under the exclusion of positions with gaps. As explained in Chapter 5 for the separate analysis of the molecular data the exclusion of positions with gaps reduces the amount of homoplasy in the data and one would also expect an increase in support (cf. increase after character weighting). However, the higher support under the inclusion of those characters with potential homoplasies indicates that these characters contain a lot of consistent but potentially misleading phylogenetic information. Therefore, the high bootstrap support under the inclusion of positions with gaps should be interpreted with care.

7.2 Consistency of current model results with previous classifications

The phylogenetic reconstructions using both data sets results in topologies that are more similar to those recovered using only molecular data. However, some relationships from the morphological data are recovered, especially some of those that were suggested as dubious in the molecular analysis. A detailed analysis is given below.

Table 7.3 lists classic monophyletic taxa that are consistently found under all models and Table 7.4 lists those relationships that are also consistently found in all analyses but do not form a monophyletic group according to previous interpretation of data. The minimum and maximum bootstrap proportions (in percentage) are given in parenthesis in the tables. The taxa in each table are ordered according to Table A2.3 in Appendix II which gives the most recent classification of the phylum Nematoda according to de Ley & Blaxter (2002).

The phylum Nematoda is always monophyletic. However, it is only weakly supported under unweighted parsimony (36 – 42%) with the exception when parsimony-uninformative characters are excluded (96%) but it is moderately to well-supported after the characters are weighted on their RC (64 – 99%). In the following discussion of the internal relationships of the Nematoda bootstrap values are only given for those groups that are not consistently recovered or where the support varies highly between the two weighting schemes because the values for consistently recovered taxa are already given in Tables 7.3 and 7.4.

Table 7.3: Monophyletic groups that are found consistently in all analyses. Bootstrap proportions are in parenthesis.

Phylum Nematoda (35 – 99%)
Enoplidae (single genus *Enoplus*) (93 – 100%)
Oncholaimina / Oncholaimoidea (99 – 100%)
Tripyloidina / Tripyloidoidea / Tripyloididae (100%)
Diphtherophorina / Diphtherophoroidea (98 – 100%)
Trichodoridae (99 – 100%)
Dorylaimida / Dorylaimoidea (excl. *Campydora demonstrans*) (100%)
Longidoridae (76 – 100%)
Xiphinematidae (single genus *Xiphinema*) (92 – 99%)
Mononchida / Mononchoidea (81 – 100%)
Mononchidae (46 – 95%)
Trichinelloidea (96 – 100%)
Chromadorea / Chromadoria (36 – 88%)
Chromadorida / Chromadorina / Chromadoroidea (69 – 96%)
Chromadoridae (99 – 100%)
Cyatholaimidae (80 – 100%)
Desmodoroidea / Desmodoridae (99 – 100%)
Laxus (98 – 100%)
Microlaimidae (97 – 100%)
Monoposthiidae (100%)
Monhysteroidea / Monhysteridae (100%)
Xyalidae (79 – 95%)
Daptonema (83 – 95%)
Sphaerolaimidae (single genus *Sphaerolaimus*) (100%)
Axonolaimidae (90 – 100%)
Comesomatidae (98 – 100%)
Plectida / Plectoidea / Plectidae (97 – 100%)
Spirurina (96 – 100%)
Spiruromorpha (96 – 100%)
Filarioidea (85 – 100%)
Ascaridomorpha (49 – 78%)
Panagrolaimoidea / Panagrolaimidae (82 – 100%)
Strongyloidoidea (53 – 96%)
Cephalobomorpha / Cephaloboidea / Cephalobidae (95 – 100%)
Criconematoidea (69 – 90%)
Sphaerularioidea / Anguinidae (67 – 98%)
Diplogasteroidea (73 – 100%)
Strongyloidea (69 – 83%)
Trichostrongylidae (54 – 77%)

The genus *Enoplus* is always monophyletic and is the sister taxon to *Enoploides brunettii*. Both relationships are well supported. The Enoploidea are only monophyletic when positions with gaps (and uninformative characters) are excluded in which case they are moderately-well supported (55 – 79%). When all characters are included and when uninformative characters are excluded *Anoplostoma vivipara* is the sister taxon to the

Oncholaimoidea rendering the Enoploidea paraphyletic. The Oncholaimoidea are always monophyletic and very well supported. The Oncholaimidae are the monophyletic sister taxon to *Calyptronema maxweberi* only under weighted parsimony when positions with gaps (and uninformative characters) are excluded in which case they are well supported (93%). In all other six cases *Calyptronema maxweberi* is found within the Oncholaimidae rendering this family paraphyletic. The Enoploidea and Oncholaimoidea are found in one single clade when all characters are included and when uninformative characters are excluded under unweighted parsimony. Under all other conditions they are found in different clades in the Enoplea or their relationships are unresolved. The two representatives of the Ironidae never form a monophyletic clade. *Syringolaimus striatocaudatus* is always the sister taxon to *Campydora demonstrans* (formerly of the Dorylaimida). They are either closely related to the Dorylaimia and Triplonchida or their position in the Enoplea is unresolved. *Ironus dentifurcatus* is usually found closely related to the Dorylaimia and under weighted parsimony excluding positions with gaps (and uninformative characters) it is the sister taxon to the Enoploidea. The Tripyloididae are always monophyletic and supported 100%. Under weighted parsimony and under unweighted parsimony excluding uninformative characters they are found as the closest sister taxon to the Chromadorea, under all other conditions under unweighted parsimony they are either closely related to the Enoploidea, Oncholaimoidea and *Trischistoma monohystera*. Therefore, the Enoplida and Enoplia are never recovered as a monophyletic group. However, with the exception of *Isolaimium* sp. and the Trichinelloidea the remaining Enoplea form a single clade, however, with little support (16%) when all characters are included under unweighted parsimony.

The Trichodoridae and Diphtherophoroidea are always monophyletic and well supported. When all characters are included the Tobrilina are monophyletic (89% support under unweighted, 47% support under weighted parsimony) and the sister taxon to the Diphtherophoroidea (42% support under unweighted, 96% support under weighted parsimony). Under unweighted parsimony excluding positions with gaps and uninformative characters *Tobrilus gracilis* is the sister taxon to the Diphtherophoroidea and *Prismatolaimus intermedius* their sister taxon (50% support). Under weighted parsimony with the exception when all characters are included, *Prismatolaimus intermedius* is the sister taxon to the Diphtherophoroidea (67 – 86% support) and *Tobrilus gracilis* their sister taxon (95 – 97% support). Under unweighted parsimony excluding uninformative characters the Tobrilina (88% support) form a different clade to the

Diphtherophoroidea (100% support). *Trischistoma monohystera* is never found in the same clade rendering the Triplonchida paraphyletic.

Table 7.4: Relationships that are consistently found in all analyses but do not form a monophyletic group according to previous interpretation of data. Bootstrap proportions are in parenthesis.

Enoploides brunettii and Enoplidae (83 – 100%)
Syringolaimus striatocaudatus and *Campydora demonstrans* (38 – 99%)
Mononchida and *Mermis nigrescens* (97 – 100%)
Monhysterida including *Cyartonema elegans* (23 – 90%)
Monhysterida (including *Cyartonema elegans*) and Comesomatidae (32 – 86%)
Terschellingia longicaudata and *Cyartonema elegans* (77 – 100%)
Sphaerolaimidae and *Desmolaimus zeelandicus* (5 – 78%)
Sphaerolaimoidea including *Desmolaimus zeelandicus* (47 – 70%)
Sphaerolaimoidea (including *Desmolaimus zeelandicus*) and Monhysteridae (52 – 90%)
Plectus acuminatus and *P. aquatilis* (92 – 99%)
Rhabditida including Plectida (95 – 100%)
Spiruromorpha and Ascaridomorpha (80 – 100%)
Setaria digitata and Filarioidea (85 – 100%)
Filariidae including *Loa loa* (38 – 68%)
Tylenchomorpha excluding Aphelenchoidea (40 – 99%)
Criconematoidea and *Paratylenchus dianthus* (99 – 100%)
Hoplolaimidae including *Globodera pallida* (40 – 93%)
Bunonema franzi and Diplogasteroidea (39 – 84%)
Strongyloidea excluding *Heterorhabditis bacteriophora* (98 – 100%)
Haemonchus contortus and *Ostertagia ostertagi* (39 – 85%)
Nematodirus battus and *Nippostrongylus brasiliensis* (36 – 87%)

As mentioned above, *Campydora demonstrans* is never found with the remaining Dorylaimida but as sister taxon to *Syringolaimus striatocaudatus* in the Enoplia. The remaining Dorylaimida always form one clade and are 100% supported. The internal relationships vary but the Longidoridae and Xiphinematidae are always monophyletic and moderately to well supported. The Dorylaimoidea are only recovered as monophyletic group and as sister taxon to *Belondira apitica* (Belonolaimoidea) under unweighted parsimony excluding positions with gaps (and uninformative characters). However, the Dorylaimoidea are still only weakly supported (40% support) in this case. The Mononchidae are moderately to well supported and they are the sister taxon to *Mylonchulus arenicolus* (Mylonchulidae). Therefore, the Mononchida are also always monophyletic and they are the sister taxon to *Mermis nigrescens*. This clade is almost always the sister taxon to the Dorylaimida (36 – 77% support) except under weighted parsimony excluding uninformative characters (and positions with gaps). The Trichinelloidea are also always monophyletic and well supported. Excluding uninformative characters (under both weighting schemes) and under unweighted parsimony including all

characters they are the most basal taxon of the Nematoda. Under unweighted parsimony excluding positions with gaps and uninformative characters and under weighted parsimony including all characters they are the sister taxon to the Dorylaimida but with weak support (< 51% support). Under weighted parsimony excluding positions with gaps they are the sister taxon to the Mermithida and Mononchida (34% support) and excluding positions with gaps and uninformative characters they are the sister taxon to the remaining Dorylaimia (67% support). *Isolaimium* sp. is never found with the remaining Dorylaimia but it is always closely related to the Axonolaimidae (Chromadorea). If the majority of the Dorylaimia (Dorylaimida, Mononchida, Mermithida and Trichinellida) are reconstructed as a monophyletic clade this clade is moderately-well supported (30% under unweighted parsimony, 64 – 78% under weighted parsimony).

The Chromadorea are always monophyletic. They are weakly supported under unweighted parsimony (36 – 56%) and moderately-well supported after character weighting (79 – 88%). The Chromadorida as well as their two families, the Chromadoridae and Cyatholaimidae, are also always monophyletic and well supported. The internal relationships in the two families vary between different analyses. The Desmodoridae are also always monophyletic and 100% supported but their internal relationships vary among analyses. The genus *Laxus* Cobb, 1894, is always monophyletic and very well supported. The genus *Eubostrichus* Greeff, 1869, is usually monophyletic and moderately-well supported (68 – 84%) except under weighted parsimony excluding gaps (and uninformative characters) when it is paraphyletic including the genus *Laxus* (88 – 90%). The Desmodorida are polyphyletic because the other two potential families never form a monophyletic clade with the Desmodoridae. The Microlaimidae and Monoposthiidae are always monophyletic. Under weighted parsimony they are sister taxa forming the Microlaimoidea. They are weakly supported (42 – 43%) when positions with gaps (and uninformative characters) are excluded, and moderately-well supported when all characters are included and only uninformative characters are excluded (80 – 81 %). Under unweighted parsimony the Microlaimoidea are monophyletic when all characters are included and when uninformative characters are excluded but they are only weakly supported (36 – 37%). When positions with gaps are excluded their relationships to each other and to the other Chromadorea are unresolved. When positions with gaps and uninformative characters are excluded the Monoposthiidae are the sister taxon to the Chromadorida and the Microlaimidae are the most basal taxon of the Chromadorea.

The Monhysterida are almost monophyletic; they always include *Cyartonema elegans* as the sister taxon to *Terschellingia longicaudata*. The relationships within the Monhysteridae are constant across all eight analyses: *Cyartonema elegans* and *Terschellingia longicaudata* are well supported and form the sister taxon to the remaining Monhysterida, which are only weakly to moderately-well supported. The Monhysteridae are 100% supported and form the sister group to the Sphaerolaimoidea including *Desmolaimus zeelandicus*. This clade is also only weakly to moderately-well supported. *Desmolaimus zeelandicus* is found as the sister taxon to the Sphaerolaimidae (genus *Sphaerolaimus*) and they are weakly to moderately-well supported; they are the sister taxon to the Xyalidae. The Xyalidae and the genus *Daptonema* are monophyletic and moderately to well-supported.

The Araeolaimida are never monophyletic. The Comesomatidae are always well supported and form the sister taxon to the Monhysterida. The Axonolaimidae are also well supported and they are the closest taxon to the Plectida. *Isolaimium* sp. (Dorylaimia) is the direct sister taxon to the Axonolaimidae when all characters are included and when uninformative characters are excluded (23 – 26% support under unweighted parsimony, 61 – 65% under weighted parsimony); otherwise it is the closest taxon to the Plectida and Rhabditida (70 – 72% support under weighted parsimony).

The Plectida are always monophyletic and very well supported but the genus *Plectus* (Bastian, 1865) is only monophyletic when positions with gaps are included. When all characters are included and when uninformative characters are excluded (under both weighting schemes), the Plectida are the sister taxon to the Rhabditida and *Teratocephalus lirellus* is the most basal taxon of the Rhabditida. When positions with gaps (and uninformative characters) are excluded *Teratocephalus lirellus* and the Plectidae are sister taxa (45 – 66% support). Under unweighted parsimony their relationships to the other orders of the Rhabditida are unresolved. Under weighted parsimony they are the sister taxon to the Spirurina when only positions with gaps are excluded (38% support) and the sister taxon to the Rhabditina and Tylenchina when positions with gaps and uninformative characters are excluded.

The clade of Rhabditida and Plectida is always well supported (95 – 100%). When the Rhabditida are monophyletic under the inclusion of all characters and under the exclusion of uninformative characters they are also well supported (91 – 92%). The relationships of the suborders of the Rhabditida are always constant. The Tylenchina and Rhabditina form one clade and the Spirurina are their sister taxon. Only the Spirurina are

always monophyletic. The Strongyloidoidea are usually found with the Rhabditina rather than the Tylenchina and *Rhabditoides inermiformis, Poikilolaimus regenfussi, Bunonema franzi* and the Diplogasteroidea are usually most basal to the Rhabditina and Tylenchina.

The Spirurina, Spiruromorpha and Filarioidea are always monophyletic and well supported and the Ascaridomorpha are also monophyletic but only moderately-well supported. The Spiruromorpha and Ascaridomorpha are always sister taxa and their relationship is well supported. *Gnathostoma turgidum* is always the sister taxon to those two infraorders. Within the Spiruromorpha, *Loa loa* (Filariidae) is always part of the Onchocercidae rendering this family paraphyletic. The Filarioidea are always monophyletic and well supported and *Setaria digitata* is always their sister taxon. The internal relationships of the Filarioidea and Ascaridomorpha vary amongst analyses. None of the families of the Ascaridomorpha are recovered in any of the analyses.

The Tylenchina are only monophyletic under weighted parsimony excluding positions with gaps and uninformative characters (47% support). In all other cases the Strongyloidoidea are the sister taxon to the Rhabditomorpha except under weighted parsimony excluding only positions with gaps when they are the most basal taxon of the Rhabditida. Hence, the Panagrolaimomorpha are also only monophyletic under weighted parsimony excluding positions with gaps and uninformative characters. On the other hand, the two superfamilies, Panagrolaimoidea and Strongyloidea are always monophyletic and moderately to well supported. The Panagrolaimidae are the sister taxon to either one of the other infraorders or the relationships of the infraorders and superfamilies are unresolved in the Tylenchina.

The Cephalobomorpha are always monophyletic and well supported. The Tylenchomorpha and Aphelenchoidea are monophyletic and the Aphelenchoidea are the sister taxon to the remaining Tylenchomorpha when positions with gaps (and uninformative characters) are excluded (65 – 92% support). However, the Aphelenchoidea find little support under unweighted parsimony and they are only moderately-well supported under weighted parsimony (66%). When all characters are included and when uninformative characters are excluded *Aphelenchoides fragariae* is the sister taxon to the Panagrolaimidae (44 – 64% support) and *Aphelenchus avenae* is the most basal taxon of the remaining Tylenchomorpha (62 – 79% support).

The Criconematoidea are always monophyletic and moderately to well supported. *Paratylenchus dianthus* is always the sister taxon to the Criconematoidea and this relationships is very well supported. They are the sister taxon to the Tylenchoidea when all

characters are included and when uninformative characters are excluded (30 – 32% support under unweighted parsimony, 78 – 79% support under weighted parsimony). Under all other conditions their relationship to the Tylenchoidea is unresolved. The Sphaerularioidea are always the sister taxon to the Criconematoidea and Tylenchoidea (40 – 67% support under unweighted parsimony, 90 – 99% support under weighted parsimony) except under unweighted parsimony when positions with gaps are excluded; in this case their position in the Tylenchomorpha is unresolved.

The Tylenchoidea are monophyletic and the sister taxon to the Criconematoidea when all characters are included and when uninformative characters are excluded. They are weakly supported under unweighted parsimony (25 – 32%) but moderately-well supported under weighted parsimony (70%). The internal relationships of the Tylenchoidea vary. Under most conditions *Globodera pallida* is the sister taxon to the Hoplolaimidae which are monophyletic but weakly to moderately-well supported (31 – 68% under unweighted parsimony, 82 – 88% under weighted parsimony). Under weighted parsimony excluding positions with gaps (and uninformative characters) *Globodera pallida* is part of the Hoplolaimidae and this clade is weakly to well-supported. The Belonolaimidae and Pratylenchidae are never monophyletic.

The Rhabditina are never monophyletic. *Bunonema franzi* is always the sister taxon to the monophyletic Diplogasteroidea. Under all conditions except unweighted parsimony including all characters and excluding uninformative characters *Poikilolaimus regenfussi* is their sister taxon. Together with *Rhabditoides inermiformis* these taxa are always found most basal to the Rhabditina plus Tylenchina clade except under weighted parsimony excluding positions with gaps when they form the direct sister clade to the remaining Rhabditomorpha (except *Rhabditoides inermiformis* which is still most basal). Therefore, the Rhabditidae and Rhabditomorpha are paraphyletic. The Strongyloidea form a monophyletic clade of moderate support amongst the majority of the Rhabditidae. However, the clade of the Rhabditidae and Strongyloidea is only weakly to moderately-well supported (45 – 88%). It is split into two distinct and well supported clades (88 – 99% and 100% support respectively). *Heterorhabditis bacteriophora* is always the most basal taxon to the Strongyloidea. *Strongylus equinus* and *Syngamus trachea* are the sister taxa to the monophyletic Trichostrongylidae forming a very well supported clade. However, *Strongylus equinus* and *Syngamus trachea* are sister taxa to each other only under unweighted parsimony including all characters and excluding positions with gaps (67 – 69% support).

In conclusion, it can be said that the reconstructed phylogenies using the combined data set are more similar to those found using only the molecular data. However, there are some relationships that resemble more closely the phylogenies resulting from the morphological characters. Contrary to the analysis of the molecular characters alone but in tune with the analysis of the morphological characters, the Enoplea and Enoplia are never recovered as monophyletic taxa. The Dorylaimia are only tentatively monophyletic because they never contain *Campydora demonstrans* and *Isolaimium* sp. and often the Trichinelloidea are also missing from the clade. The Monhysterida are as good as monophyletic (including *Cyartonema elegans*). They include *Desmolaimus zeelandicus*, which were found with the Axonolaimidae using only molecular data. A closer similarity to the molecular analysis is found in the relationship of *Isolaimium* sp. with the Axonolaimidae rather than with the Dorylaimia where it was most frequently found in the morphological analysis. Within the Rhabditida the relationships of the suborders are much clearer in the combined analysis with the exception of the Panagrolaimomorpha that are almost always polyphyletic. In the molecular analysis the latter taxon is always found in one clade but may be paraphyletic including *Aphelenchoides fragariae* whilst they are never monophyletic using only morphological characters. Furthermore, the combined analysis clarifies the relationship between the Spiruromorpha and Ascaridomorpha, rendering both infraorders monophyletic and as sister taxa to each other.

Since elements from both separate analyses are found in the combined analysis, it is very unlikely that the higher number of molecular characters overpowers the effect of the morphological characters. As discussed above, some relationships only become clear after combining all available data.

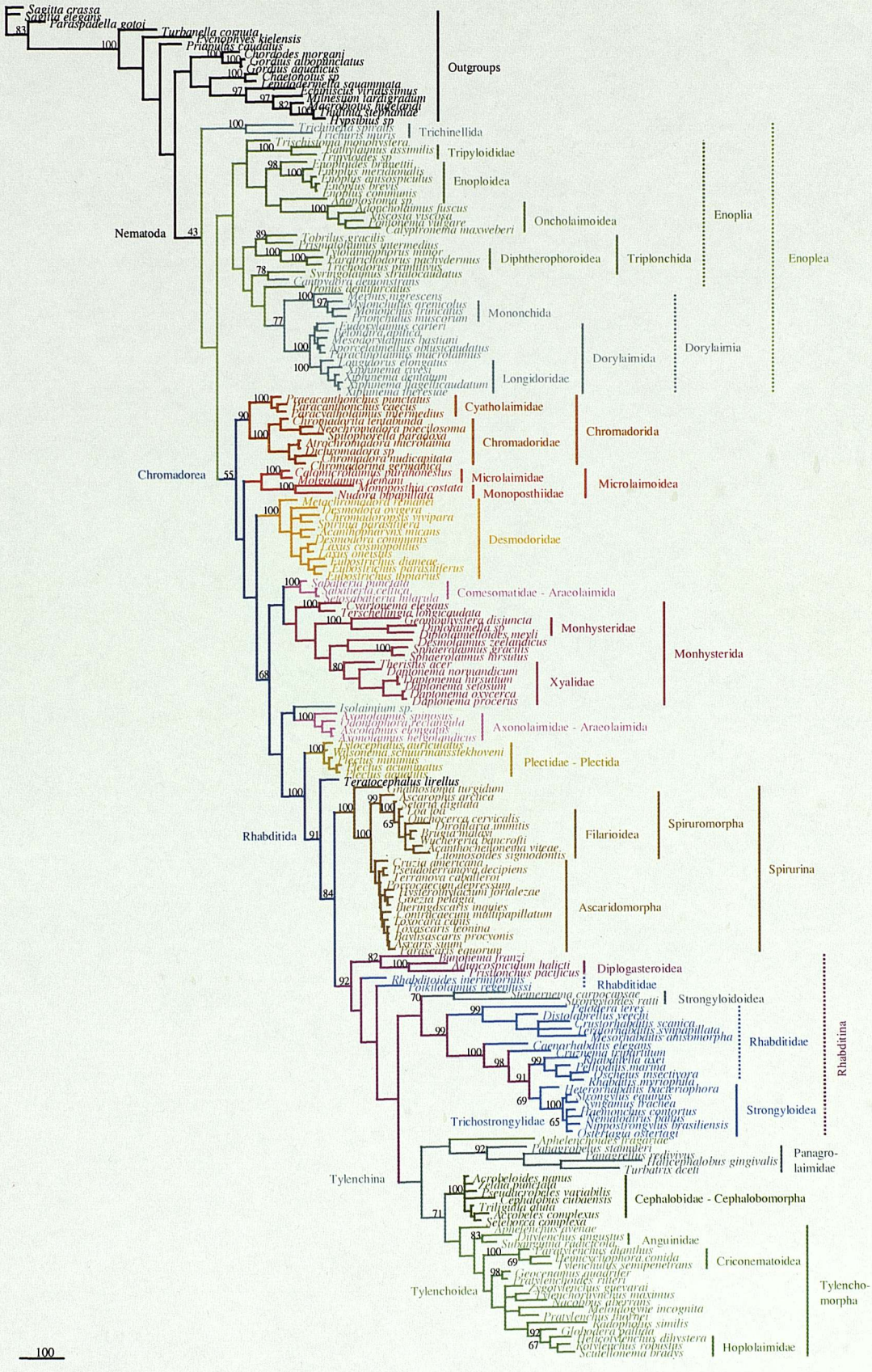


Figure 7.1: Strict consensus tree of the combined character analysis under unweighted parsimony including all characters; BP of < 65% support only given for major clades, no BP given for clades in monophyletic families; dotted lines depict polyphyletic or unresolved taxa

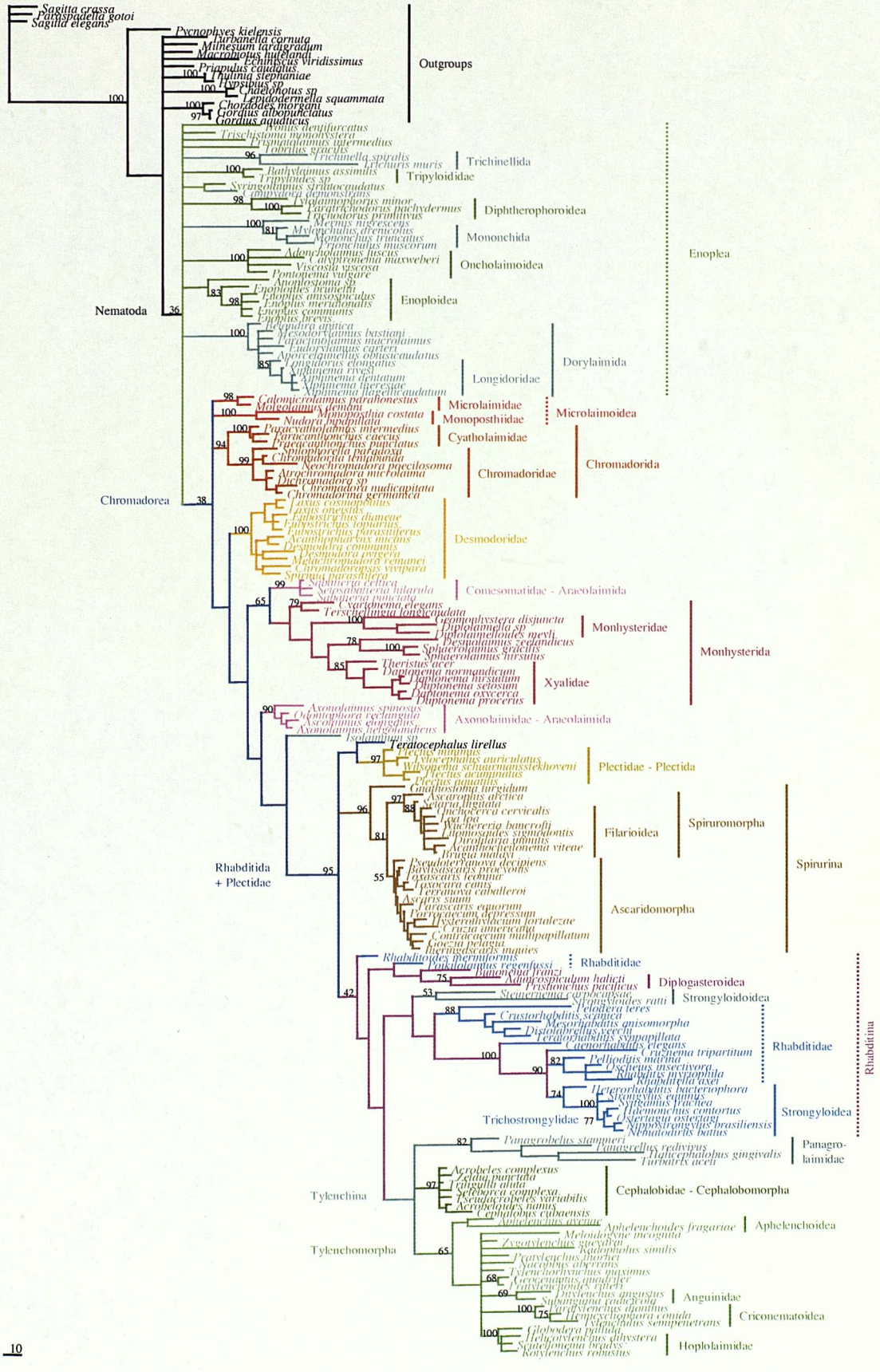


Figure 7.2: Strict consensus tree of the combined character analysis under unweighted parsimony excluding all molecular characters with positions with gaps; BP and lines as in Figure 7.1

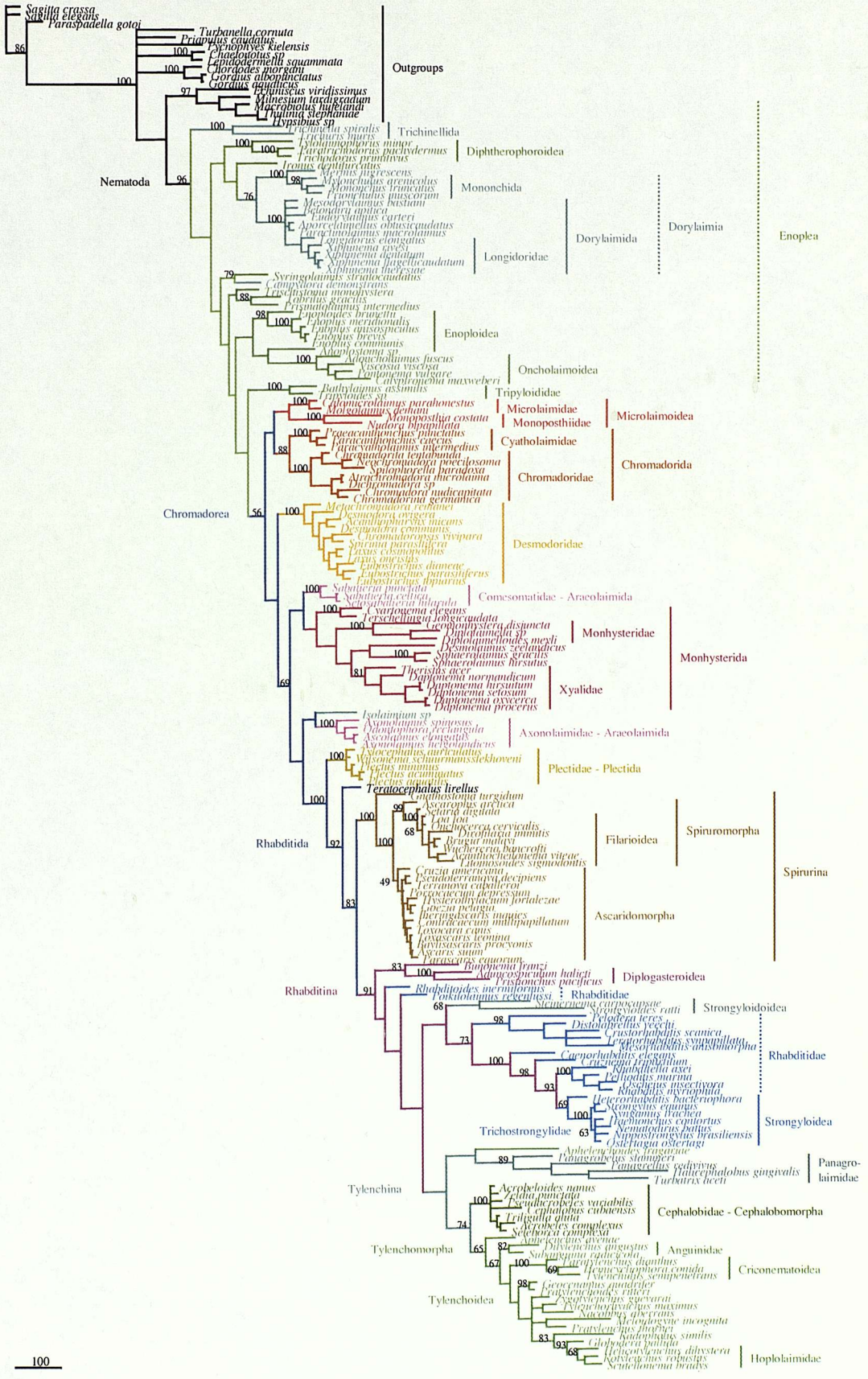


Figure 7.3: Strict consensus tree of the combined character analysis under unweighted parsimony excluding all uninformative characters; BP and lines as in Figure 7.1

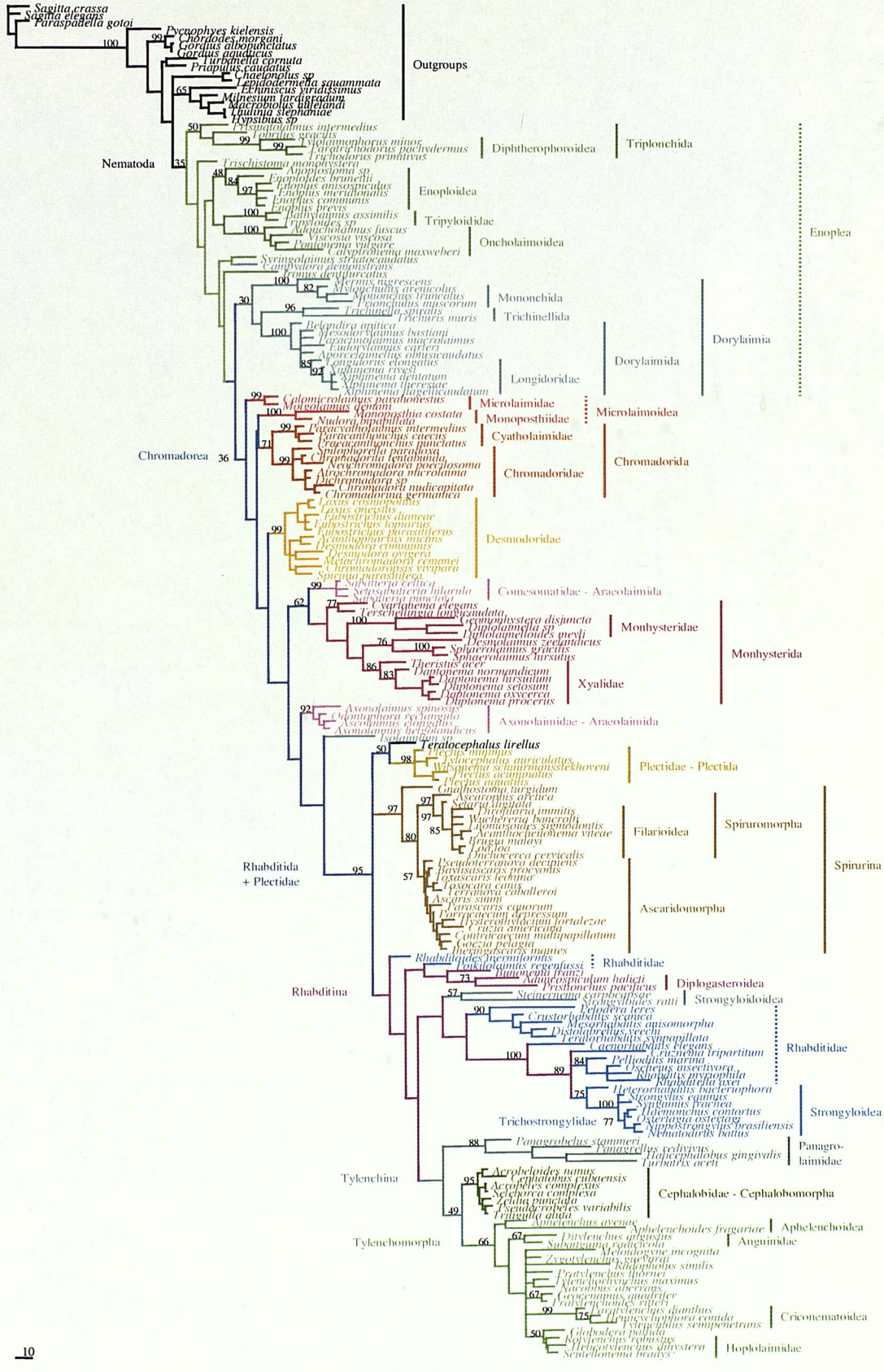


Figure 7.4: Strict consensus tree of the combined character analysis under unweighted parsimony excluding all molecular characters with positions with gaps and all uninformative characters; BP and lines as in Figure 7.1

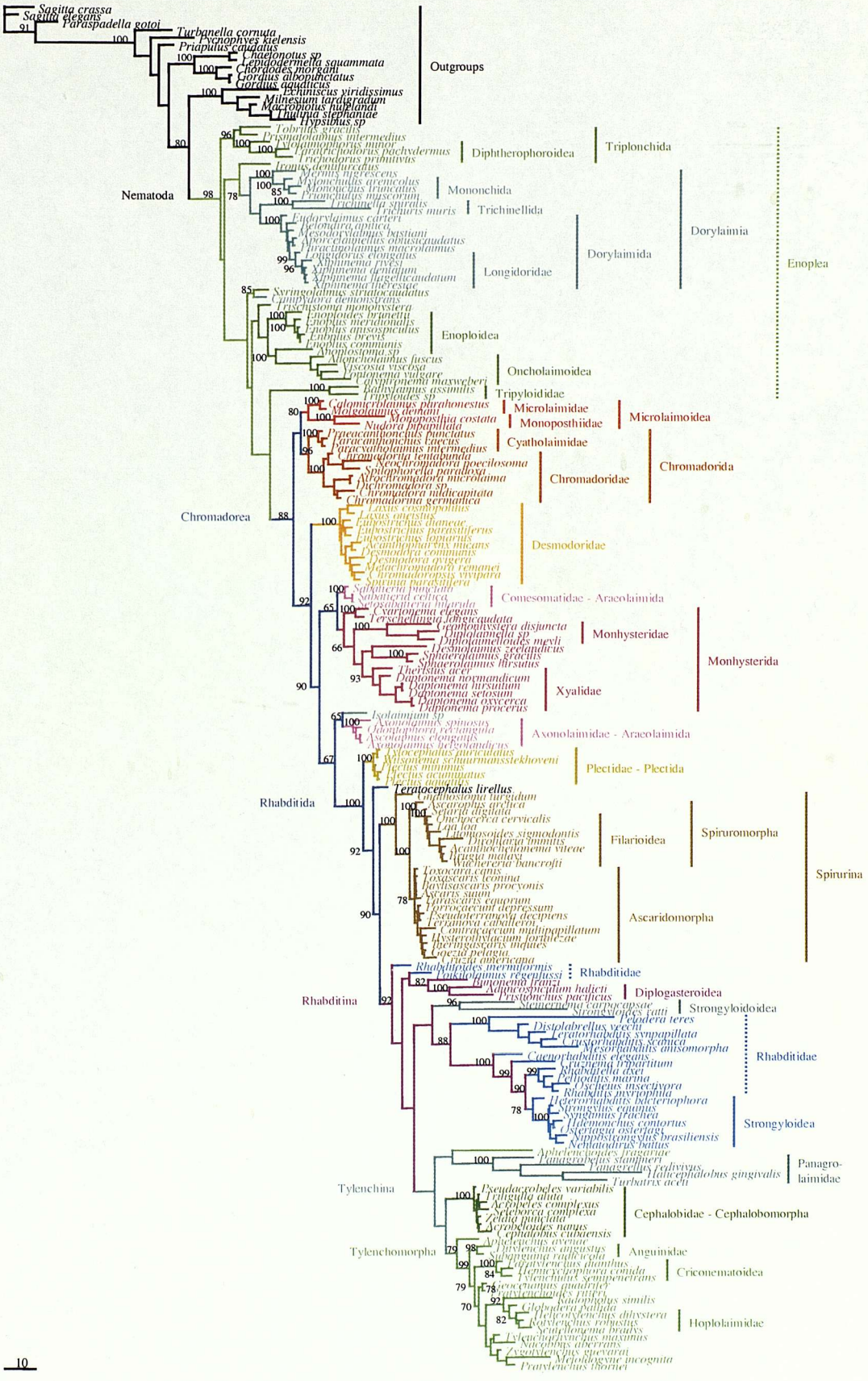


Figure 7.5: Strict consensus tree of the combined character analysis under weighted parsimony including all characters; BP and lines as in Figure 7.1

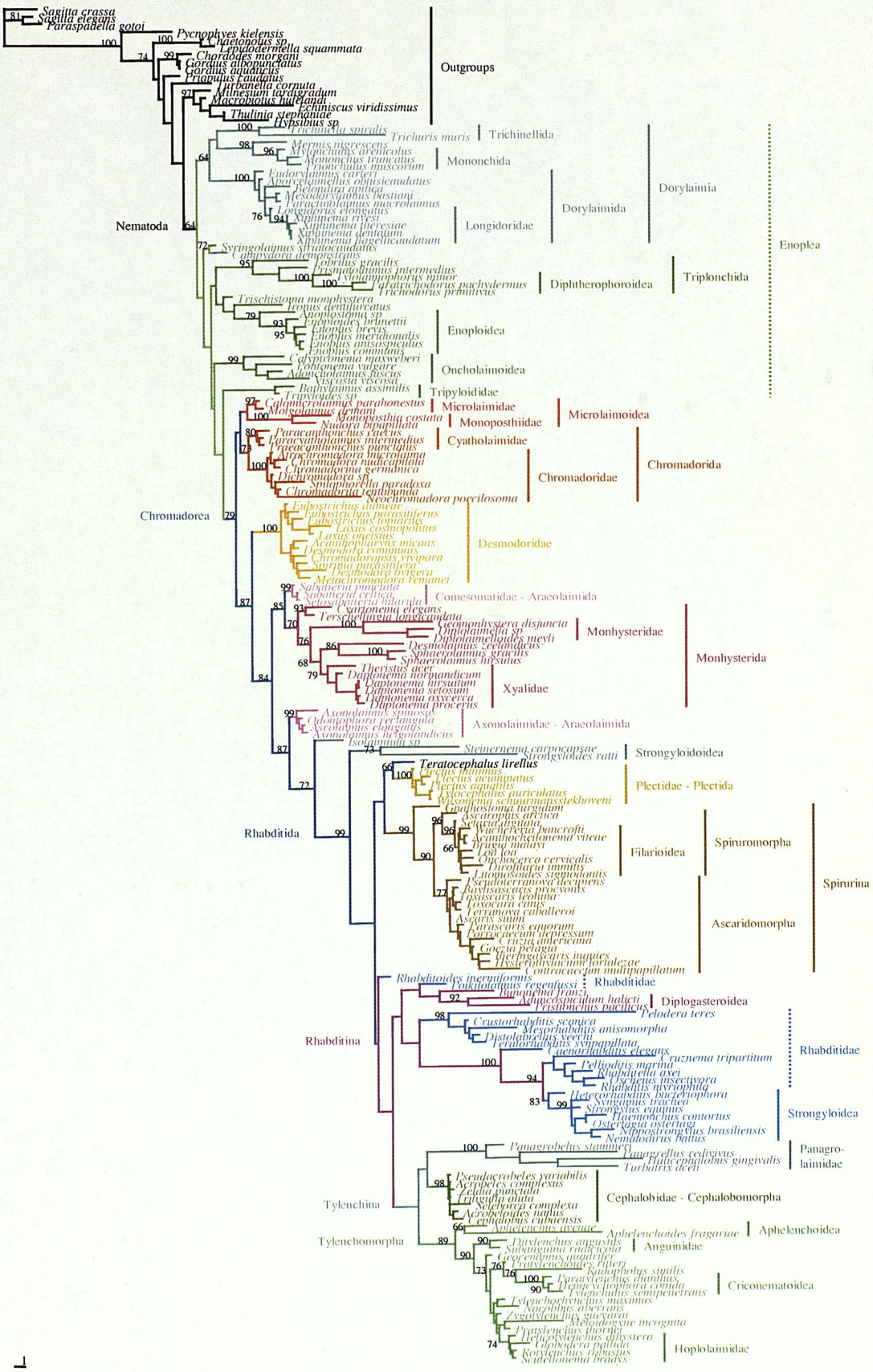


Figure 7.6: Strict consensus tree of the combined character analysis under weighted parsimony excluding all molecular characters with positions with gaps; BP and lines as in Figure 7.1

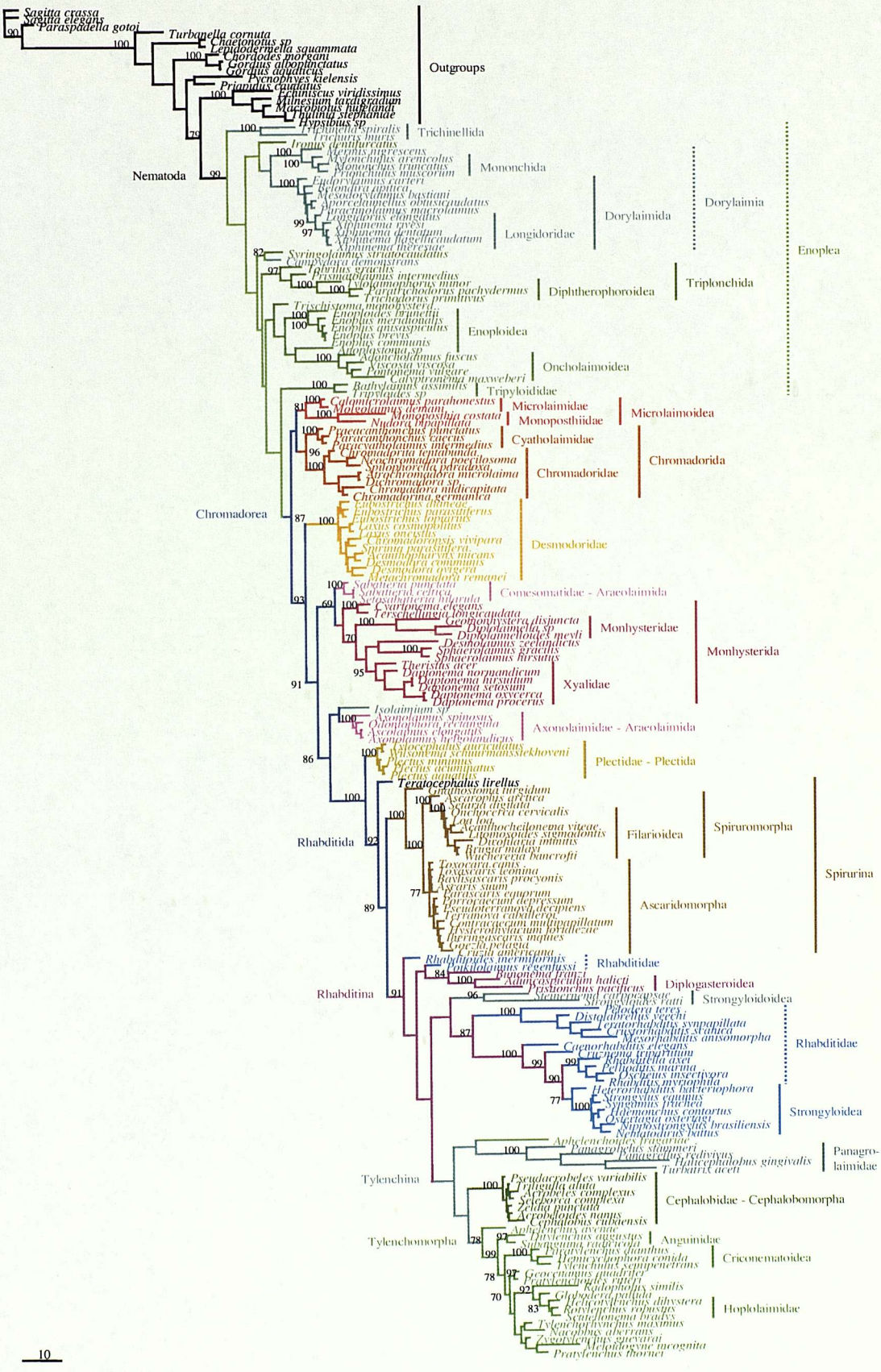


Figure 7.7: Strict consensus tree of the combined character analysis under weighted parsimony excluding all uninformative characters; BP and lines as in Figure 7.1

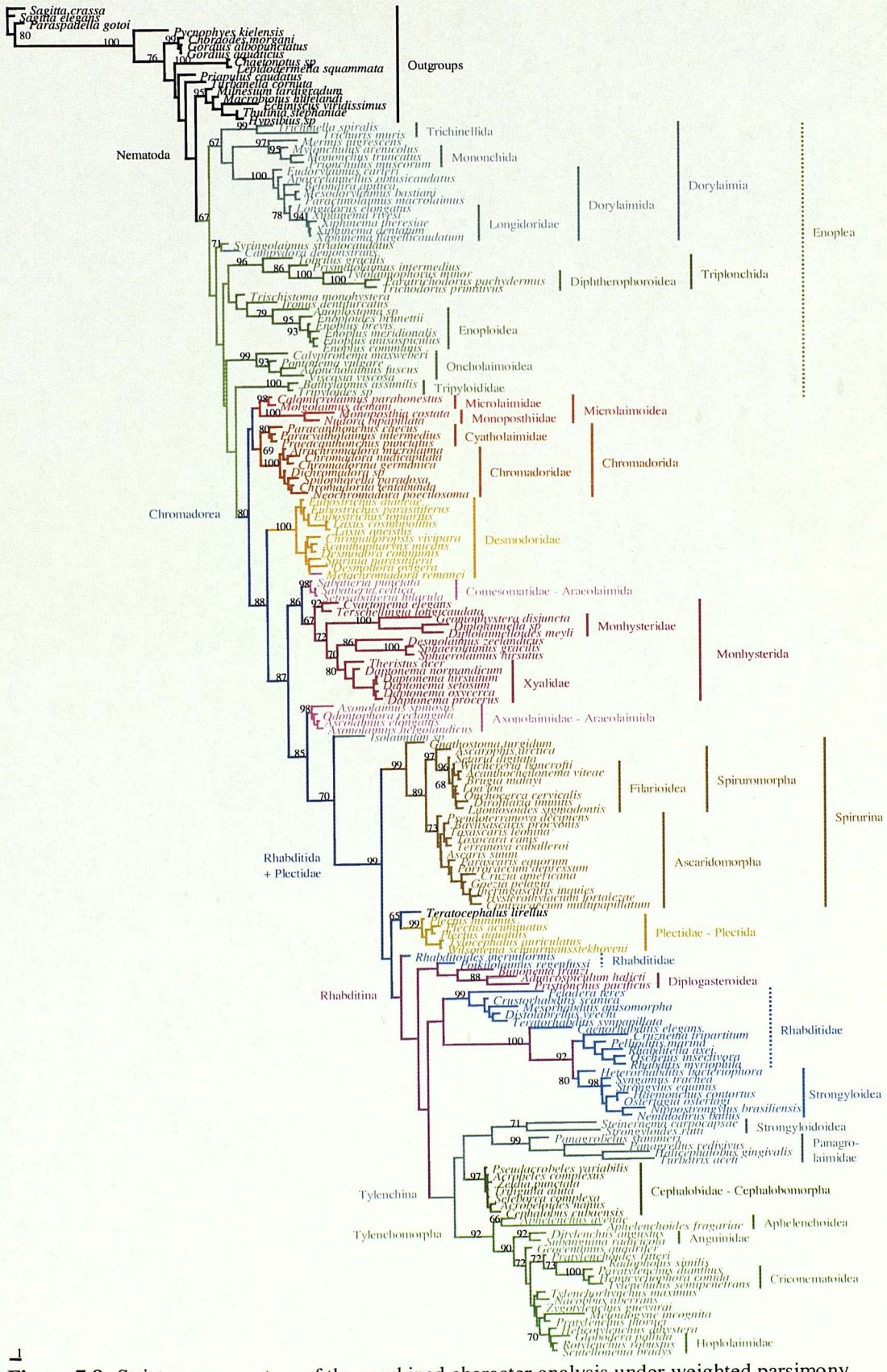


Figure 7.8: Strict consensus tree of the combined character analysis under weighted parsimony excluding all molecular characters with positions with gaps & all uninformative characters; BP and lines as in Figure 7.1.

8. Discussion of the Revised Phylogeny of the Nematoda

8.1 Fundamental differences and similarities between the analyses of the new data sets

In most respects, the phylogenetic reconstructions using the molecular characters on their own and in combination with the 159 morphological characters agree with recent phylogenetic reconstructions. The monophyly of the Nematoda is firmly established and the phylum is often highly supported by both bootstrap proportions and posterior probabilities. Of the major clades, the Chromadorea and the Rhabditida plus Plectida are also always recovered. The Enoplia and Dorylaimia, however, are only monophyletic when molecular characters are analysed on their own. In the combined analysis, the internal relationships within the Enoplea vary according to the state of character inclusion and character weighting. The phylogenetic reconstruction of the morphological characters only provides a very different picture. None of the three major clades are monophyletic but the phylogeny resembles much more the ecological relationships of the species than the proposed phylogenetic relationships. Dorchin *et al.* (2004) also recovered an eco-morphological cladogram in an analysis of morphological characters for gall-midges.

In this study, the analysis of the same morphological characters that have been used for over a century to assess the evolutionary relationships of the Nematoda resulted in a system with little similarity to traditional classifications or phylogenetic reconstructions using molecular characters. This came as a bit of a surprise but also opens a very interesting discussion. The homology assessment of morphological characters is inherently difficult and inundated with problems. As long as the true underlying genetic origin is unknown the homology of any morphological character is open to debate. As already mentioned in the Introduction (Section 1.2.1) and in the discussion of the character coding in Materials and Methods (Section 2.3.2.3), many morphological characters exhibit strong convergence because they are subject to natural selection and similar environmental factors cause the selection of similar phenotypes. Additionally, a number of coding mistakes were discovered during the character interpretation (see Section 9.1). Therefore, relying on morphological characters alone can be grossly misleading when trying to infer phylogenetic relationships (Whiting, 1998; Alvarez *et al.*, 2000; Radford *et al.*, 2001; deBry, 2003).

Why should, therefore, the classical and molecular phylogenies be preferred over the alternative system that results from the empirical analysis of the morphological characters?

Many molecular biologists would argue that the molecular phylogeny is the most robust one because it is based on the real genetic information and because the mathematical models of the evolutionary history of molecular characters have become very sophisticated. However, this does not explain the striking similarities found between the molecular reconstruction and the classical phylogenies. For the last decade taxonomists using morphological characters have been under heavy criticism from the molecular community because there are no clear models and algorithms that can explain their experience and intuition in making inferences about the relationships between species. However, molecular analysis has proven them right for much of the time. To answer this conundrum we need to answer a second question: Why do the empirical and intuitive analyses of morphological characters lead to such drastically different phylogenies?

Classic phylogeneticists did not apply empirical mathematical systems to reconstruct phylogenies but relied on their experience and instinct as to which character states were homologous and which homoplastic. In the rigorous empirical analysis of this study, each state that was seemingly of the same evolutionary origin was treated as a single, homologous character state amongst all species that had this state. In a classic, intuitive analysis however, this distinction did not have to be made *a priori*. A state could have been described in the same way in several groups and served as symplesiomorphy for each group but the character could then be ignored for the reconstruction of the phylum if it was thought to have had several independent evolutionary origins. In the empirical analysis, each independent origin would have to be coded as a separate character and in those species that did not acquire the character at all it would have to be coded either as absent or as missing data. In either case, the empirical analysis then runs into the problem of how to treat missing data and any solution will introduce misleading information (see Section 1.2.1.5). The only way out of this dilemma in an empirical analysis is to go through a series of coding, analysis and re-coding of homoplastic characters. Many characters will be lost in the final data set for the analysis of the phylum because they will introduce misleading relationships. In a second scenario, however, the same character state in two distantly related clades may in fact be of the same evolutionary origin because this state is plesiomorphic for a much larger clade. In this larger clade, the character has undergone many changes including two independent reversals to the ancestral state. In this case, no iterative coding approach can prevent a mathematical algorithm from inferring that the same states originate from the same change in evolutionary time. Algorithms are not yet able to distinguish between a recent and ancient common evolutionary history unless the ancestral states are defined *a priori*. In this analysis, it was not possible to define ancestral

states *a priori* because they were unknown. An iterative analysis has not been performed because the aim was to show what happens when morphological characters are treated in an empirical analysis. The last chapter (Chapter 9) however contains a detailed discussion of each morphological character, including its possible homoplasies and suggestions for alternative coding.

Now, we can answer our initial question in two parts. Firstly, the results from an intuitive analysis of morphological characters should be preferred over an empirical analysis of the same data because today's mathematical models cannot yet handle all the information on the evolutionary history of morphological characters that we already have from experience. The current study is a prime example where apparent empirical and objective mathematical algorithms are misleading and intuitive human assessments provide a better understanding of the evolutionary history of a set of characters (see also Barriel *et al.*, 1993; Cantrell & Hanlin, 1997; Whiting, 1998; Meier & Wiegmann, 2002; Garnica *et al.*, 2003). Secondly, molecular characters (a) carry a lot of fundamental information, (b) their homology assessment is much easier than that for the morphological characters and (c) the mode of molecular evolution is much better understood than that of morphological characters. Therefore, the phylogenetic reconstruction of molecular characters is more robust than that of morphological characters when empirical methods are employed.

The above observations also lead to the following conclusions about the similarities between the empirical reconstructions of the phylogenetic systematics of the Nematoda using molecular characters and traditional classifications based on morphology. The straightforward answer is that intuition and experience are very good methods of inferring phylogenies based on morphological characters. The human mind is able to grasp complex issues of evolutionary change reasonably accurately and many of the intuitive decisions cannot be replicated even in the most sophisticated mathematical models. Some molecular phylogeneticists may argue against this observation but there is no denying the fact that this and other studies have demonstrated significant similarities between classic and molecular phylogenies (e.g. Shoshani *et al.*, 1998; Bradford, 2002; Perez-Lozada *et al.*, 2002). And there is another advantage to using traditional inferences in the analysis of morphological characters: Many molecular phylogenies – this one included – are based solely on a single gene. A single gene is a very small representation of the total genetic information contained in the genome of any species. On the other hand, any analysis using morphological characters takes into account a wide variety of features and therefore a wide range of underlying genetic material. However, this is not saying that genetic evidence has

not got its advantages. However, as can be seen in the remaining part of this chapter, molecular evidence can be of great use in the resolution of conflicting hypothesis from the evidence of morphological characters.

In this study, the combined data resulted in a phylogeny that resembled that from the molecular data alone but morphological characters nevertheless left their fingerprints. Here, while reducing the clarity of the relationships in the Enoplea, they produced some strong relationships for taxa that seemed potentially misplaced in the molecular phylogeny. Other studies also found that adding morphological characters to a molecular phylogeny increases resolution and support in areas where the molecular phylogeny was weak (e.g. Tehler, 1995; McDowell & Bremer, 1998; Doyle & Endress, 2000; Thompson *et al.*, 2001; Winterton *et al.*, 2001, Won *et al.*, 2001).

In accordance with the above discussion the reconstructions using the molecular characters alone and those using the combined data set are used in the following sections for the revision of the phylogenetic systematics of the Nematoda.

8.2 Comparison of the phylogenies from the different analyses and a partial revision of the phylogenetic systematics of the Nematoda

Despite the numerous agreements between the reconstructions using molecular characters alone, combining these with morphological characters and with traditional classifications, there are a number of significant differences. The differences are based on character conflict within and between data sets as well as violations of model assumptions. The following section discusses the differences in the phylogenetic reconstructions together with the potential changes to the phylogenetic systematics of the Nematoda. With respect of the above discussion the reconstructions using only the morphological characters are not included in this discussion. However, the reconstructions using the combined data set will be considered because the morphological characters provide interesting additional information when analysed in combination with the molecular characters. Tables 8.1 and 8.2 list all those taxa that are in common between the molecular and combined analyses. The first table contains those taxa that are also traditional monophyletic groups and the second contains other consistent relationships. Figure 8.1 displays a partial revision of the phylogenetic systematics of the Nematoda based on the taxa included in this study.

From the discussion in the individual analyses chapters (Chapters 5 – 7) it became apparent that the best analysis is probably the one excluding both sets of potentially ambiguous characters. Furthermore, under the parsimony criterion, character weighting is

preferred to unweighted characters. Therefore, in the following discussion about the potential relationships of species and taxa of variable phylogenetic position the results under the analyses of potentially higher quality (excluding ambiguous characters, weighted vs. unweighted parsimony) are regarded as more likely.

Table 8.1: Traditional monophyletic groups also recovered in the analyses of the molecular characters and the combined data set.

Phylum Nematoda
Enoplidae
Oncholaimoidea
Tripyloididae
Trichodoridae
Diphtherophoroidea
Dorylaimida (excluding *Campydora*)
Longidoridae
Xiphinematidae
Mononchida
Trichinelloidea
Chromadoridae
Cyatholaimidae
Desmodoridae
Laxus
Microlaimidae

Monoposthiidae
Monhysteridae
Sphaerolaimidae
Axonolaimidae
Comesomatidae
Plectidae
Spirurina
Spiruromorpha
Panagrolaimidae
Cephalobomorpha
Tylenchomorpha
Criconematoidea (incl. *Paratylenchus*)
Anguinidae
Diplogasteroidea
Strongyloidea

Table 8.2: Groups that have been recovered consistently in the analyses of the molecular characters and the combined data set but that do not form traditional monophyletic groups

Syringolaimus striatocaudatus and *Campydora demonstrans*
Triplonchida excluding Tripylina
Mononchida and *Mermis nigrescens*
Cyartonema elegans and *Terschellingia longicaudata*
Comesomatidae as sister taxon to or in same clade as Monhysterida
Teratocephalus lirellus as closest sister taxon to the Plectida or Rhabditida
Spiruromorpha and Ascaridomorpha as sister taxa
Setaria digitata with Filarioidea
Tylenchomorpha excluding Aphelenchoidea
Globodera pallida with or as sister taxon to Hoplolaimidae
Paratylenchus dianthus with or as sister taxon to Criconematoidea

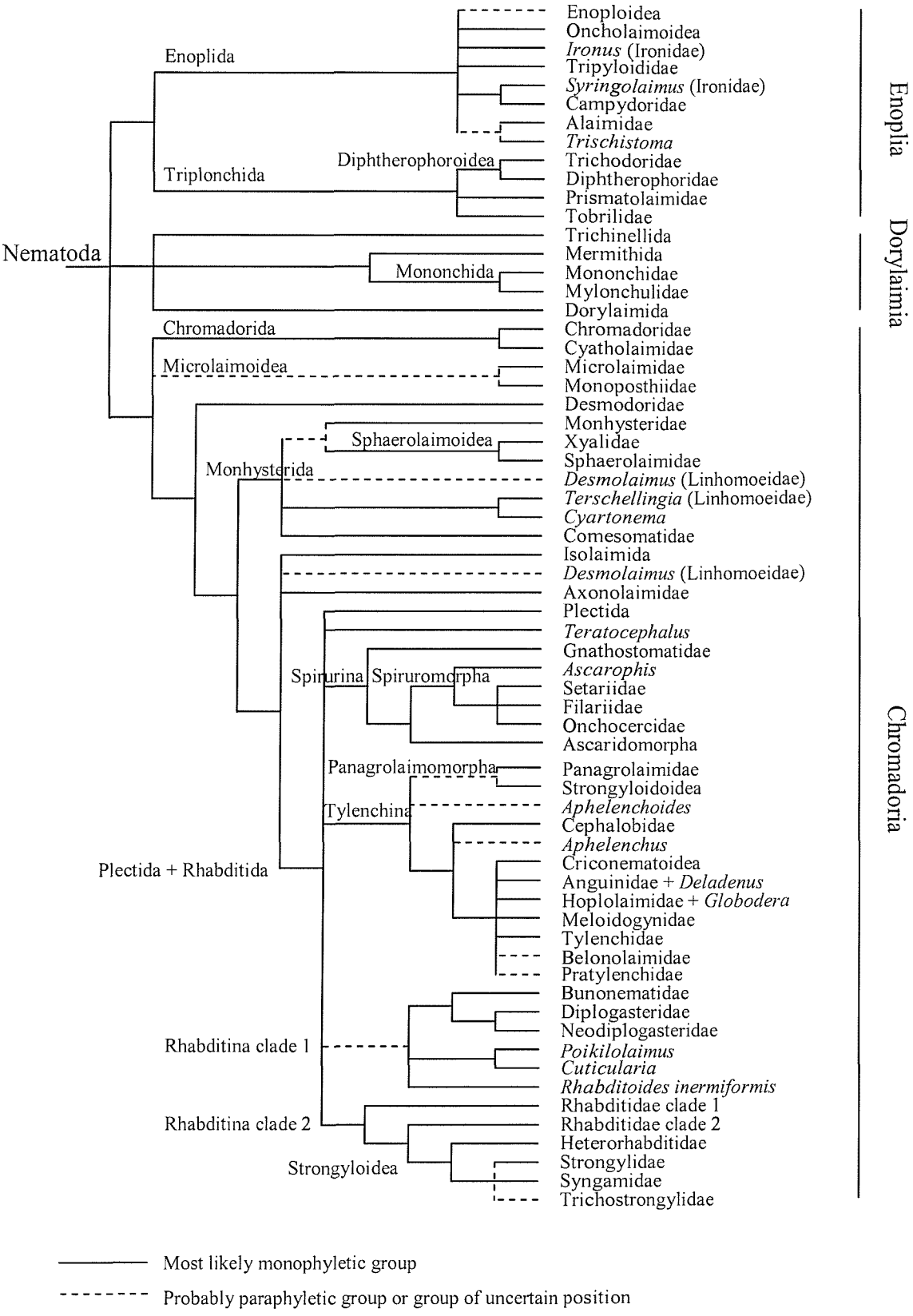


Figure 8.1: Proposed phylogenetic systematics of the Nematoda based on the taxa included in this study.

The Nematoda are most certainly monophyletic and they probably contain three major clades because almost all reconstructions using the molecular characters alone support the monophyly of de Ley & Blaxter's (2002) three subclasses Enoplia, Dorylaimia and Chromadoria. The class Chromadorea with the only subclass Chromadoria is monophyletic in all analyses except when uninformative characters are excluded under unweighted parsimony in the analysis of the molecular characters alone (no resolution). However, the other class, the Enoplea, is only monophyletic under unweighted parsimony and maximum likelihood distances including all characters and under LogDet distances including all characters and excluding uninformative characters when the molecular characters are analysed on their own. The inclusion of the morphological characters resulted in a variety of reconstructions of the genera of the Enoplea. As discussed above, the morphological characters of this study introduce a lot of homoplasy. Therefore, it can be suggested that the Enoplia and Dorylaimia are monophyletic but that the monophyly of the Enoplea remains uncertain.

The current set of molecular data cannot confirm which of the two subclasses, the Enoplia or Dorylaimia, are more likely to be the most ancestral taxon of the Nematoda. Only few investigators in the past have made proposals about the most primitive nematode. Chitwood & Chitwood (1933) recognised the similarities in the pharyngeal structures of the Plectida and Rhabditida and proposed that their common ancestor may have been the primitive nematode; therefore, they rooted the nematode tree within the Chromadorea. Maggenti (1963) disagreed with this notion and suggested that the similarities are based on convergent evolution. He offers a tentative solution for the primitive nematode to be found as the ancestor of the Enoplea and Chromadorea, that the Secernentea derived from an ancestor of the Monhysterida and that the Plectoidea are the most derived taxon of the Monhysterida. However, he did not make any final decision of the phylogeny of the Nematoda but instead referred to the need for a better understanding of comparative morphology.

Most authors agree with Filipjev (1934) that the ancestor is probably a marine nematode and found somewhere between the Enoplea and Chromadorea; the debate concerns which of the two groups is more primitive. Filipjev (1934), Coomans (1977) and Lorenzen (1981/1994) proposed an enoplid ancestry whilst Hyman (1951), Inglis (1983) and Andrássy (1976) suggested that the primitive nematode was a chromadorid ancestor. However, most investigators were still using the old terms of Adenophorea and

Secernentea and a plectoid-rhabditoid ancestry remained in the debate. It was not until Blaxter *et al.* (1998) used DNA sequences that it became finally accepted that the primitive nematode was not found between the Adenophorea and Secernentea but somewhere in the enoplean or chromadorean ancestry. It has become clear from the study of Blaxter *et al.* (1998) and the current investigation that the nematode ancestry lies in the Enoplea. Unfortunately, current molecular data are not sufficiently resolved to decide whether the Enoplia or Dorylaimia are the older of the two groups in the Enoplea. However, if the theory of a marine ancestry is valid, the primitive nematode should be sought amongst an enoplean ancestor. Maggenti (1963) already suggested that some enoplid nematodes bear many of the potentially plesiomorphic characters listed by Chitwood & Chitwood (1933) and that they are mostly marine. He also proposed that the Dorylaimida are probably derived from the Enoplia.

Within the Enoplia, there are probably two orders, the Enoplida and the Triplonchida as introduced by de Ley & Blaxter (2002). The majority of the subtaxa of the Triplonchida are monophyletic and the internal relationships only vary in the relative relationships between the Diphtherophoroidea, *Tobrilus gracilis* and *Prismatolaimus intermedius*. Traditionally, the Diphtherophoroidea were classified with the Dorylaimida (e.g. Thorne, 1939; Clark, 1961; Siddiqi, 1961, 1973; de Coninck, 1965; Coomans & Loof, 1970) and only Siddiqi (1983) proposed the order Triplonchida for the Diphtherophoroidea and Trichodoroidea. The Tobrilidae and Prismatolaimidae were traditionally associated with the Tripylidae in the Enoplida (de Coninck, 1965, Siddiqi, 1983). Lorenzen (1981/1994) moved the Prismatolaimidae into the Chromadoria (Leptolaimina) on account of the dorsally spiral amphids. Molecular data showed a closer affinity to the Diphtherophoroidea and led de Ley & Blaxter (2002) to move them into the Triplonchida. This position is confirmed with this extended data set.

Since a close affinity was proposed between the Tobrilidae and Tripylidae, the latter taxon was moved into the Triplonchida by de Ley & Blaxter (2002). However, in the molecular analysis *Trischistoma monohystera* never groups with the Triplonchida but always with the Enoplida and *Tripyla* sp. is found in the Plectidae. Therefore, the Tripylidae should remain in the Enoplida but more molecular data are needed to establish its precise relationships with the other taxa of the Enoplida. The position of *Tripyla* sp. in the Plectidae is dubious. The sequence for *Tripyla* sp. was given to the author by Mark Blaxter and originates from Paul de Ley's lab. Its identity has not been checked by the

author and is therefore questionable. There were no morphological data available for this taxon because the specimen was only identified to genus level.

The internal relationships of the Enoplida remain uncertain. *Syringolaimus striatocaudatus* and *Campydora demonstrans* consistently form a sister relationships and *Enoploides brunettii* is probably the closest sister taxon to the genus *Enoplus*. *Alaimus* sp. was not included in the combined analysis but in the molecular analysis it was usually found closely related to *Trischistoma monohystera*. The Oncholaimoidea and the Tripyloididae are always monophyletic. The position of *Ironus dentifurcatus* varies significantly. In the molecular analysis it is always found in the Enoplida and usually closely related to the Tripyloididae; in the combined analysis however, this species is found in very different positions and even basal to the Dorylaimia rather than with the Enoplida.

The Anoplostomatidae used to be classified with the Oncholaimidae (de Coninck, 1965) until Lorenzen (1981/1994) moved them into the Enoploidea. The molecular data strongly support the latter affinity. The Tripyloididae have been classified in the Enoplida on account of the similarity of the head structure with the Tobrilidae (Filipjev, 1918, 1934; Gerlach & Riemann, 1974) and in the Chromadorea (de Coninck & Schuurmans Stekhoven, 1933; Chitwood & Chitwood, 1950; de Coninck, 1965, Andrássy, 1976) on account of the spiral amphids. Since Lorenzen (1978) discovered the presence of metanemes in the Tripyloididae they were retained in the Enoplida. The molecular data support this position despite the shape of the amphids.

The Ironidae were traditionally classified as Tripyloidea (Chitwood & Chitwood, 1950; de Coninck, 1965) or Tripylina (Andrássy, 1976). Siddiqi (1983) placed them into their own suborder and Lorenzen (1981/1994) moved them into the Enoplacea together with the Enoploidea. Lorenzen established the monophyly of the Ironidae based on the buccal cavity armature but at the same time already pointed out considerable differences in the members of this family. Since the evolution of the buccal armature is largely directed by its functionality it is quite possible that the three or four movable teeth evolved numerous times in the evolution of the Enoplia. The evidence from the molecular data suggests that the Ironidae are paraphyletic and more sequences from other genera are needed to revise their position. A basal position of the Ironidae in the Dorylaimia as it was found in some of the analyses with the combined data set is in accordance with suggestions by Filipjev (1918) and Maggenti (1963). Maggenti proposed that the Dorylaimida evolved from an ancestor of *Ironus*.

The phylogenetic position of the Alaimidae has always been controversial. Originally, they were placed in the Dorylaimida (Filipjev, 1934; Thorne, 1939). Chitwood & Chitwood (1950) moved them into the Tripyloidea (Enoplina). Clark (1961) raised them to subordinal rank where they remained (de Coninck, 1965) until Hooper (1973) raised them to ordinal level. Andrássy (1976) retained them in the Tripylina (Oxystomonoidea). Lorenzen (1981/1994) moved them back into the Dorylaimida (Bathyodontina) on account of the posterior position of the oesophageal gland outlets and a lack of metanemes and Maggenti (1982) and Inglis (1983) retained them there. Siddiqi (1983) supported Hooper's (1973) decision to give them the rank of order. The molecular data clearly support a position in the Enoplia and even suggest the old affinity with the Tripylidae (*Trischistoma monohystera*). Clearly, more sequences from Alaimidae are needed to clarify their phylogenetic relationships.

The phylogenetic position of the genus *Campydora* Cobb, 1920, has been uncertain. Thorne (1939) classified the Campydorinae in the Leptonchidae (Dorylaimoidea). Clark (1961) raised them to family level. Jairajpuri (1964) emended their diagnosis and de Coninck (1965) moved the Campydoridae into the newly established Nygolaimoidea (Dorylaimina). Ferris (1971) moved them back into the Leptonchoidea (Dorylaimina) where Andrássy (1976) retained them. Jairajpuri *et al.* (1976) raised the Campydorida to the rank of superfamily and Ahmad & Jairajpuri (1979) moved them back into the Nygolaimina (Dorylaimida). Jairajpuri (1983) established the rank of Campydorina and in the same year Siddiqi (1983) moved the Campydorina into the Enoplida. Jairajpuri & Ahmad (1992) retained the Campydorina in the Dorylaimida. De Ley & Blaxter (2002) followed their decision. However, Mullin *et al.* (2001), using partial 18S sequences, had already established that this genus does not belong to the Dorylaimia but the Enoplia with a close affinity to the Alaimida and Tripylida. The current analysis with the extended data set confirms that *Campydora* belongs in the Enoplida.

The relationships in the Dorylaimia are a little more consistent than those in the Enoplia in this analysis. There are at least three clades, consisting of the orders Trichinellida, Dorylaimida and (Mermithida [only represented species *Mermis nigrescens*] plus Mononchida). The relative relationship between the three clades is uncertain as all three combinations occur in the different analyses. Within the Mononchida, the Mononchidae are paraphyletic in the molecular analysis that contained more species than the combined analysis. The interrelationships of the Dorylaimida are not very consistent

with the exception of the Longidoridae and Qudsianematidae that are probably monophyletic.

Few authors have attempted a classification of all potential taxa of the Enoplea (*sensu* de Ley & Blaxter, 2002). Chitwood & Chitwood (1950) classified the Enoplidae and Oncholaimidae in the Enoploidea, the Alaimidae and Mononchidae together with the Ironidae and Tripylidae in the Tripyloidea, the Dorylaimidae, Diphtherophoridae and Belonolaimidae in the Dorylaimoidea, the Mermithidae and Tetradonematidae in the Mermithoidea and the Trichuridae, Trichinellidae and Cystoopsidae in the Trichuroidea, all in the Enoplina (Enoplida, Aphasmidia).

Clark (1961) divided the order Enoplida into five suborders, Enoplina, Dorylaimina (including Dorylaimoidea and Mononchoidea), Alaimina, Trichosyringina (including Mermithoidea, Trichinellidae and Trichuridae) and the Dioctophymatina, raising many superfamilies to the rank of suborder. Later, Clark (1962) emended the diagnosis of the Alaimina to include the Diphtherophoroidea (formerly of the Dorylaimina). De Coninck's (1965) classification was very similar to that of Clark. He included the Dorylaimoidea, Mononchoidea, Nygolaimoidea (including Campydoridae) and Diphtherophoroidea in the Dorylaimina with the Alaimina as their sister taxon. However, he classified the Mermithoidea and Trichuroidea under the Enoplida.

Jairajpuri (1969) raised the Mononchida to the rank of order. Coomans & Loof (1970) included the Dorylaimina, Mononchina, Diphtherophorina, Bathydontina and Trichosyringina under the Dorylaimida, thereby lowering the Mononchida back to suborder level, raising the Diphtherophoroidea to suborder level and re-including the Mermithoidea and Trichuroidea (*sensu* de Coninck, 1965) into the Dorylaimida. Andrássy (1976) grouped the Mononchina, Dorylaimina, Diphtherophorina and Mermithina in the Dorylaimida and the Trichocephalida (including Trichinellidae and Trichuridae) as sister taxa to the former in the Penetrantia.

Maggenti (1982) raised all groups one rank to include the Enoplida, Dorylaimida, Isolaimida, Mermithida, Mononchida, Muspiceida and Trichochephalida under the Enoplia. Siddiqi (1983) raised the Oncholaimina (de Coninck, 1965), Alaimina, Bathyodontina and Tripylina to ordinal rank, introduced the Triplonchida and moved the Trichinellida and Mermithida into the Stichosomia (Stichosomida, Maggenti, 1982).

All the above changes are mainly reorganisations of the ranking levels and some authors moved the insect (Mermithidae) and vertebrate parasitic families (Trichuridae and Trichinellidae) out of the Dorylaimia and into a separate group within the Enoplea.

However, Lorenzen (1983) moved the latter groups, the Trichosyringida (Trichuroidea, Mermithoidea) and Dioctophymatoidea into the Secernentea on account of the presence of phasmid-like structures (Kaiser, 1977) and the non-homology of the juvenile stylets (Richter, 1971) in the Mermithidae. Based mainly on morphological data but with some information from DNA sequencing analysis, de Ley & Blaxter (2002) suggested the erection of two subclasses, the Enoplia and Dorylaimia, of the Class Enoplea. The extended sequencing analysis of this study supports the division of the Enoplea into two subclasses as well as the orders of the Dorylaimia, the Dorylaimida, Mononchida, Mermithida and Trichinellida. However, still no decision can be made on the interrelationships of the orders within the Dorylaimia.

Isolaimium sp. is consistently found closely related to the Axonolaimidae (Chromadorea). Traditionally, the Isolaimiidae were placed in the Dorylaimida (de Coninck, 1965; Timm, 1969; Coomans & Loof, 1970; Andrássy, 1976) and Maggenti (1982) established a separate order, the Isolaimida (Enoplia). Inglis (1983) moved the Isolaimiidae into the Enoplida. However, Filipjev (1934) placed the genus *Isolaimium* into the Plectidae (Chromadorida) and Goodey (1963a) and Gerlach & Riemann (1973) followed his suggestion (Araeolaimida). Lorenzen (1981/1994), however, retained the Isolaimiidae in the Dorylaimida (Bathyodontina) on account of the position of the oesophageal gland outlets. The identities of two sequences of *Isolaimium* sp. have been re-checked by Peter Mullin who provided the sequences to the author. There is no suggestion that the specimens have been wrongly identified or that the sequences have been mixed up. Two different specimens from two different locations were sequenced. The two sequences are not identical but very similar and are always found in the same position irrespectively of the character set and model used. It is unlikely that the position of *Isolaimium* sp. results from homoplasy because its position is constant even if positions with gaps and uninformative characters are excluded. Either, the Isolaimiidae are indeed closely related to the Axonolaimidae as suggested by Filipjev (1934), Goodey (1963a) and Gerlach & Riemann (1973) or the mathematical algorithms are unable to produce the true results. Whatever the cause, more sequences from this family are required to establish its true phylogenetic position.

The sequences from *Trichinella spiralis* and *Trichuris muris* were found to be more closely related to outgroups than the other nematodes under the LogDet criterion when positions with gaps and parsimony-uninformative characters were excluded from the data set. This is probably not because their sequences are more similar to those of the outgroups

but because the sequences of the Trichinelloidea and outgroups are very dissimilar from those of the remaining nematodes. The DNA sequences of these two species are very different from those of the remaining nematodes and were difficult to align in places. The sequences also exhibit long branches that may be a result of increased evolution in the rRNA gene compared to the remaining nematode sequences. The exclusion of all potentially ambiguous positions left only 393 characters to be analysed. It is likely that the characters that are informative for the position of the Trichinelloidea amongst the Dorylaimia are found amongst characters that are potentially ambiguous for the relationships of the remaining Nematoda and are thus excluded in many of the analyses. In a future analysis the Trichinelloidea should be analysed only with the potentially closely related sequences from other Dorylaimia or Enoplea and ambiguous positions should be excluded more carefully. The addition of other sequences from the Trichinellida as well as the Dioctophymatida and Muspiceida can also potentially reduce the problem. If the new sequences were intermediate they can firstly make alignment easier and secondly they may cut the long branches between the Trichinelloidea and the remaining Nematoda. Recently, Rusin *et al.* (2003) performed a phylogenetic reconstruction of the Dorylaimia using the same gene as this study (SSU rRNA) and found that *Soboliphyme baturini* (Dioctophymatida) forms the sister taxon to the Trichinellida.

However, although long-branch attraction may play a role in the potentially incorrect placement of the Trichinelloidea, this problem does not seem to be particularly great. The sequences of the Trichinelloidea and *Pelodera punctata* (Cobb, 1914) are the only ones where considerably long branches have been observed in the whole data set. Since *Trichinella spiralis* and *Trichuris muris* are expected to be closely related and since *Pelodera punctata* is always found in the same clade of Rhabditidae it is unlikely that the topology is an artefact of long-branch attraction but that it is the true topology.

Within the Chromadorea *sensu* de Ley & Blaxter (2002), the Chromadorida are the only order that is most definitely monophyletic. Its two families, the Chromadoridae and Cyatholaimidae, are also monophyletic. The Microlaimoidea may or may not be monophyletic because the relationships between the monophyletic Microlaimidae and Monoposthiidae vary in the different analyses. The Chromadorida, Microlaimidae or Monoposthiidae are the most basal taxon of the Chromadorea. The Desmodoridae are a monophyletic family but their internal relationships remain unresolved. The Desmodoridae never form a monophyletic group with the Microlaimoidea and therefore the Desmodoroidea (Desmodorida) (*sensu* de Ley & Blaxter, 2002) is not a valid taxon.

The sister relationship of the Chromadoridae and Cyatholaimidae has been well established, only the level of the common rank varied between different classifications. Chitwood & Chitwood (1950) combined the two taxa at the superfamily level (Chromadoroidea), de Coninck (1965) raised this level to the order Chromadorida, Andrássy (1976) and Malakhov (1994) concurred with de Coninck whilst Lorenzen (1981/1994) lowered the common rank back to the superfamily level. De Ley & Blaxter (2002) retained the common level at the rank of superfamily but introduced the order Chromadorida with the single suborder Chromadorina in order to maintain the hierarchical system and allowing the Rhabditida to be an inferior rank in the Chromadorea. The current molecular data do not make any suggestion as to which level in the phylogeny is the equivalent of a certain rank and before any conclusion about ranks can be drawn it would be useful to have more related taxa represented in the data. The molecular data consistently retained the same two groups in the Chromadoridae but they do not correspond to any subfamilies of other classifications.

Chitwood & Chitwood (1950) regarded the Microlaimidae as the sister taxon to the Chromadoridae and Cyatholaimidae (Chromadoroidea, Chromadorina) and the Desmodoridae as more distantly related in a separate superfamily (Desmodoroidea, Chromadorina). De Coninck (1965) split the Desmodoridae *sensu* Chitwood & Chitwood (1950) into at least two superfamilies (Desmodoroidea, Spirinoidea) and moved the Microlaimidae and Monoposthiidae also into the Desmodorida. Andrássy (1976) combined the Desmodoridae *sensu* Chitwood & Chitwood (1950) and raised their rank to superfamily. He followed de Coninck (1965) in the affinity of the Microlaimidae (Richtersioidea, Desmodorina) with the Desmodoroidea (Desmodorina). Lorenzen (1981/1994) reorganised these taxa and gave the Microlaimidae plus Monoposthiidae (Microlaimoidea) the same rank as the Desmodoroidea and Chromadoroidea. De Ley & Blaxter (2002) retained the Microlaimoidea but reversed Lorenzen's system back to combining the Desmodoridae (Desmodoroidea) and Microlaimoidea in the Desmodorina (Desmodorida) as sister taxon to the Chromadorida. The new molecular data do not support the Desmodorida *sensu* de Ley & Blaxter (2002) but retain more of Lorenzen's system of the separation between the Desmodoridae and Microlaimoidea. In fact, it is more likely that the Microlaimoidea are more closely related to the Chromadorida than to the Desmodorida.

With regards to the subfamilies of the Desmodoridae, the molecular data consistently recover the sister relationship of *Acanthopharynx micans* (Eberth, 1873) and *Desmodora*

communis (Bütschli, 1874) which equates to the Desmodoridae *sensu* de Coninck (1965) and to the Desmodorinae *sensu* Lorenzen (1981/1994). The combined data also supports the monophyly of the genus *Laxus*. However, *Desmodora ovigera* Ott, 1976, never groups with *Desmodora communis* and the genus *Eubostrichus* is not always monophyletic but often includes the genus *Laxus*. Equally, the apparently synonymous genera *Metachromadora* Filipjev, 1918, and *Chromadoropsis* Filipjev, 1918, as well as *Eubostrichus* and *Catanema* Cobb, 1920 (all *sensu* Lorenzen, 1981/1994) never form sister relationships. On the other hand, *Xyzzors* sp. is consistently found in the Desmodoridae (see also Kampfer *et al.*, 1998) while de Coninck (1965) and Lorenzen (1981/1994) grouped it with the Cyatholaimidae. Since the specimen was only identified to genus level no morphological characters were obtained in this study but the morphological analysis of Kampfer *et al.* (1998) also grouped it in the Desmodoridae. Nothing could be found about its apparent change in phylogenetic affinity.

The order Monhysterida must be extended in their content. In all analyses the Comesomatidae (formerly of the Araeolaimida) are part of this order and *Cyartonema elegans* is always the sister taxon to *Terschellingia longicaudata* (Linhomoeidae). The Xyalidae are probably monophyletic and form the sister taxon to the Sphaerolaimidae in which case the Sphaerolaimoidea are also monophyletic. The Monhysteridae are also monophyletic and probably the sister taxon to the Sphaerolaimoidea.

De Coninck (1965) placed the genus *Cyartonema* Cobb, 1920, in the Siphonolaimidae (Monhysterida) and Jayasree & Warwick (1977) retained it there when they described *Cyartonema elegans*. However, Lorenzen (1981/1994) moved this genus to the classic Chromadorida. The molecular data consistently found *Cyartonema elegans* as sister taxon to *Terschellingia longicaudata* of the Linhomoeidae (Siphonolaimoidea, Monhysterida). Therefore, this genus is most likely part of the Siphonolaimoidea (Monhysterida) in accordance with de Coninck (1965) and Jayasree & Warwick (1977).

The Monhysteridae, Xyalidae and Sphaerolaimidae were always regarded as closely related. The molecular data suggest that there may be a slightly closer affinity between the latter two families which would be in accordance with Andrássy (1976), Lorenzen (1981/1994) and de Ley & Blaxter (2002). According to Chitwood & Chitwood (1950) the Linhomoeinae and Sphaerolaiminae are closer related to each other than to the Monhysteridae and Siphonolaimidae. De Coninck (1965) suggested a close relationship between the Sphaerolaimidae and Monhysteridae (including *Theristus* Bastian, 1865, of the later Xyalidae) as Monhysteroidea and sister taxa to the Linhomoeoidea and

Siphonolaimoidea. Andrássy (1976) proposed the closest relationship of these taxa between the Monhysteridae and Siphonolaimidae in the Monhysteroidea which is the sister taxon to the Diplolaimelloidea, Sphaerolaimoidea and Xyaloidea. He raised the Linhomoeoidea to suborder level (Linhomoeina, Monhysterida).

Lorenzen (1981/1994) suggested that the Monhysteridae, Xyalidae and Sphaerolaimidae form the Monhysteroidea as sister taxon to the Siphonolaimoidea including the Siphonolaimidae and Linhomoeidae. De Ley & Blaxter (2002) closely followed Lorenzen but they raised the two groups to suborder level (Monhysterina and Linhomoeina) of the Monhysterida. The molecular data, however, do not support any one of those proposals in all respects. In particular, the Linhomoeidae appear to be paraphyletic because *Desmolaimus zeelandicus* groups either with the Monhysterida or with the Axonolaimidae (Araeolaimida) but never as sister taxon to *Terschellingia longicaudata*. On the other hand, if *Cyartonema elegans* is a species of the Siphonolaimidae or Siphonolaimoidea the molecular data support Lorenzen (1981/1994) most closely.

The Araeolaimida *sensu* de Ley & Blaxter (2002) are not supported by the molecular data because the Comesomatidae always group with the Monhysterida and the Axonolaimidae are the closest sister taxon to the Plectida plus Rhabditida. The latter group is always accompanied by *Isolaimium* sp. (formerly of the Dorylaimia). In the analysis of the molecular characters alone *Desmolaimus zeelandicus* (Linhomoeidae) is also the sister taxon to the Axonolaimidae. In the combined analysis this species is found in the Monhysterida as sister taxon of *Sphaerolaimus*.

Filipjev (1934) placed the Axonolaimidae and Comesomatidae in the Monhysterata. Wieser (1954) moved the Comesomatidae into the Chromadorida on account of their punctuated cuticle and spiral amphids. Chitwood & Chitwood (1950) moved them back into the Monhysterina as sister taxon to the Axonolaimidae (Axonolaimoidea). De Coninck (1965), Andrássy (1976) and Platt (1985) followed Wieser's earlier move and retained them in the Chromadorida (Chromadoroidea and Cyatholaimina respectively). Lorenzen (1981/1994) returned them into the Axonolaimoidae (Monhysterida). Inglis (1983) agreed with Lorenzen (1981) that the Comesomatidae are not part of the Chromadorida (*sensu* Lorenzen, 1981) but cased doubt over their close relationship with the Axonolaimidae; nevertheless, he retained their sister relationship for the time being. The molecular data provide evidence of the position of the Comesomatidae as close relatives of the Monhysterida rather than the Chromadorida but also support the doubt raised by Inglis (1983) as to its close relationship with the Axonolaimidae.

Traditionally, the Plectida are always classified as an "Adenophorean" taxon, often as part of the Monhysterida (Chitwood & Chitwood, 1950) or Araeolaimida (de Coninck, 1965; Andrássy, 1976; Inglis, 1983). The molecular data support this classic concept that the Plectidae form the intermediate taxon between the "Adenophorea" and "Secernentea". Depending on the different data sets, models and character inclusions the Plectida and Rhabditida are either sister taxa or the Plectidae form a clade within the Rhabditida. In the latter case, they can be found with either one of the three suborders of the Rhabditida. Based on the quality of the different analyses it is more probable that the Plectida are in fact part of the Rhabditida and the closeness to the Araeolaimida in some analyses is a result of symplesiomorphies of the Chromadorea. Despite the uncertainties of the true placement of the Plectidae, molecular data support the origin of the Rhabditida from the Araeolaimida.

Teratocephalus lirellus is probably a species of the Rhabditida but its relationship to the other suborders remains uncertain. In the different analyses it is the sister taxon to the Plectidae if these are part of the Rhabditida, otherwise it is found as the most basal taxon of the Rhabditida. Based on the quality of the analyses, *Teratocephalus lirellus* is probably a close relation to the Plectida. The Teratocephalidae have previously been placed in the Plectoidea/Araeolaimida (de Coninck, 1965), Leptolaimina/Chromadoria (Lorenzen, 1981/1994), Leptolaimina/Araeolaimida (Inglis, 1983) or its own suborder Teratocephalina in the Rhabditida (Andrássy, 1976). More data are needed to clarify the phylogenetic position of the Teratocephalidae in the Nematoda.

In the current molecular analysis, the Spirurina are always monophyletic and the Spiruromorpha and Ascaridomorpha are usually sister taxa. *Gnathostoma turgidum*, *Dentostomella* sp. Schulz & Krepkogorskaja 1932 and *Philonema* sp. are basal taxa of the Spirurina whilst the relationship of *Brumptaemilius justini* to the other taxa of this suborder remains uncertain. *Ascarophis arctica* is the sister taxon to the Filarioidea which also include *Setaria digitata*. The last species is currently classified in its own family Setariidae in the Filarioidea (Anderson & Bain, 1976) but was not included in the list in de Ley & Blaxter (2002). The Ascaridomorpha and its families as well as the Onchocercidae are all paraphyletic and not well resolved in most analyses. It is well documented that the internal relationships of the Ascaridomorpha cannot be reliable reconstructed using the SSU rRNA gene because it is very conserved in this group (Nadler, 1992; Nadler & Hudspeth, 1998). Since only two infraorders of the Spirurina are represented by more than a single taxon no strong conclusions about the overall relationships amongst this suborder can be made.

However, it is almost certain that the Spiruromorpha are monophyletic and possibly the sister taxon to the Ascaridomorpha. De Ley & Blaxter (2002) very recently revised the Spirurina and the current analysis concurs with it apart from the remarks above. Therefore no further discussion of this group is given here.

Brevibucca sp. and *Myolaimus* sp. of uncertain affinity in the phylogeny are found closely related to the Tylenchina and Spirurina respectively but only under those analyses including more ambiguous characters. Under potentially better analysis conditions the positions of both species are uncertain. This suggests that there is a high level of character conflict in the remaining 'good' characters. Since both taxa were only identified to genus level no morphological data were collected. Clearly, more sequences have to be added to discover the true phylogenetic position of these two species. Therefore, no real conclusion can be drawn about their closest relationships amongst the Nematoda. However, Andrássy (1976) and more recently Dolinski *et al.* (2001) classified the Brevibuccidae in the Panagrolaimoidea, one position it has been found in in the current analyses. On the other hand, the Myolaimina are usually identified as close relatives of the Rhabditina (Inglis, 1983) whilst the current analysis suggest a closer relationship to the Spirurina.

The Tylenchina *sensu* de Ley & Blaxter (2002) have recently been revised by Siddiqi (2000) mainly based on morphological characters and by the aforementioned authors with the additional information from molecular characters. The Tylenchina and its three infraorders, the Panagrolaimomorpha, Cephalobomorpha (here only Cephalobidae) and the Tylenchomorpha, are probably monophyletic.

In the combined analysis the Strongyloidoidea are often found in the Rhabditina but in the analyses of the molecular characters alone they are usually the sister taxon of the monophyletic Panagrolaimidae. *Rhabditophanes* sp. (Rhabditina) is the sister taxon to *Strongyloides ratti* and therefore probably part of the Strongyloidoidea and not part of the Rhabditina (also Dorris *et al.*, 2002). The position of *Steinernema carpocapsae* varies in the Tylenchina; it is usually found basal to the Panagrolaimomorpha or Tylenchina but in many analyses no closest relative could be found in the Rhabditida and its position remained unresolved. The Cephalobidae are always monophyletic. It is more likely that they are the sister taxon to the Tylenchomorpha than to the Panagrolaimomorpha. This is in opposition to older data when the Panagrolaimomorpha and Cephalobomorpha *sensu* de Ley & Blaxter (2002) were classified in the Rhabditina (Chitwood & Chitwood, 1950) or Rhabditida (Andrássy, 1976) as sister taxon to the Tylenchina or Tylenchida respectively.

Within the Tylenchomorpha the relationships of the two species of the Aphelenchoidea are uncertain. *Aphelenchoides fragariae* can be found either with the Panagrolaimomorpha or as sister taxon to the Cephalobidae and remaining Tylenchomorpha. *Aphelenchus avenae* can be found as the sister group to the remaining Tylenchomorpha, Cephalobidae or both of these groups. Only in the combined analysis when positions with gaps and positions with gaps and uninformative characters are excluded are the Aphelenchoidea monophyletic and form the sister group to the remaining Tylenchomorpha; the Tylenchomorpha are monophyletic only under those conditions. Based on the quality of the analyses the Aphelenchoidea are either monophyletic or at least *Aphelenchoidea fragariae* and *Aphelenchus avenae* are very closely related and found at the base of the remaining Tylenchomorpha. In the case where the Aphelenchoidea are not monophyletic *Aphelenchoidea fragariae* might be the sister taxon to *Aphelenchus avenae* plus the remaining Tylenchomorpha.

The uncertain position of the Aphelenchoidea reflects its past history. Since Chitwood & Chitwood (1950) placed the superfamily Aphelenchoidea into the Tylenchina this taxon has usually been given the rank of suborder Aphelenchina in the order Tylenchida (Andrássy, 1976; Nickle & Hooper, 1991) or order Aphelenchida as sister group to the Tylenchida (Siddiqi, 1980; Hunt, 1993). Only de Ley & Blaxter (2002) reduced it back to superfamily rank because of their initial DNA analyses. More data are needed to clarify the phylogenetic position of the Aphelenchoidea.

The relationships of the remaining Tylenchomorpha are very variable. De Ley & Blaxter (2002) missed the Paratylenchidae in their list; nevertheless, *Paratylenchus dianthus* always groups with the other three species of the Criconematoidea *sensu* de Ley & Blaxter (2002) (Criconematina *sensu* Siddiqi, 2000) rendering this superfamily monophyletic. Also, *Globodera pallida* (Heteroderidae *sensu* Siddiqi, 2000) is always found within or as sister taxon to the Hoplolaimidae (Siddiqi, 2000; de Ley & Blaxter, 2002). Both taxa are also supported by the work of Chitwood & Chitwood (1950), Andrássy (1976) and Maggenti *et al.* (1987). The two species of the Anguinidae (Sphaerularioidea *sensu* de Ley & Blaxter, 2002; Anguinoidea *sensu* Siddiqi, 2000) are usually found in one clade with *Deladenus* sp. (Phaenopsitylenchidae) and the unidentified marine Tylenchid in the analysis of the molecular characters alone. Therefore, the Phaenopsitylenchidae are probably a family in the Sphaerularioidea. In the combined analysis the Anguinidae are monophyletic.

The Tylenchoidea *sensu* de Ley & Blaxter (2002) (Hoplolaimina, Siddiqi, 2000) are probably paraphyletic; they are only monophyletic in the analysis of the molecular characters alone under LogDet distances when positions with gaps are excluded and in the combined analysis when all characters are included and when uninformative characters are excluded. The Belonolaimidae *sensu* de Ley & Blaxter (2002) (Telotylenchidae *sensu* Siddiqi, 2000) and the Pratylenchidae (Siddiqi, 2000; de Ley & Blaxter, 2002) are never monophyletic. *Meloidogyne incognita* (Meloidogynidae) is found in various relationships with species of the Tylenchoidea *sensu* de Ley & Blaxter (2002) (Hoplolaimina *sensu* Siddiqi, 2000). No further relationships can be established because different analyses resulted in widely different internal relationships for the Tylenchomorpha.

The Rhabditina may or may not be monophyletic. The Diplogasteroidea are always monophyletic and, based on the analyses of higher quality, *Bunonema franzi* is probably their sister taxon. This molecular relationship supports recent findings of morphological apomorphies between the Diplogasterids and Bunonematidae (von Lieven, 2002; Dolinski & Baldwin, 2003). Together with *Poikilolaimus regenfussi*, *Cuticularia* sp. (always sister taxa) and *Rhabditoides inermiformis Bunonema franzi* and the Diplogasteroidea form one clade that is either basal to the remaining Rhabditina or basal to the Rhabditina and Tylenchina. The remaining Rhabditomorpha form one clade that is divided into two distinct and well supported groups which are identical in all analyses. One of the two groups includes the Strongyloidea which are monophyletic. *Heterorhabditis bacteriophora* is the most basal taxon and the Trichostrongylidae are probably monophyletic. The relative relationships between the Trichostrongylidae, Strongylidae and Syngamidae remain uncertain.

The molecular phylogeny strongly supports the drastic changes undertaken by de Ley & Blaxter (2002) with regards to moving the former Strongylina (*sensu* Chitwood & Chitwood, 1950) or Strongylida (*sensu* Andrássy, 1976) into the Rhabditina. De Ley & Blaxter (2002) also suggested that the Bunonematidae, the genus *Rhabditoides* and Diplogasteroidea may be closely related at the base of the Rhabditina. De Ley & Blaxter (2002) based their revisions of the findings of Sudhaus & Fitch (2001) who inferred rhabditid phylogeny with SSU rRNA data. The paraphyly of the Rhabditoidea was already observed by Blaxter *et al.* (1998) when the former Strongylida were found to be derived from the Rhabditidae. Sudhaus & Fitch (2001) further discovered that certain *Rhabditoides* spp. and *Poikilolaimus* are more basal and closely related to Diplogasteroidea than

previously thought (Sudhaus, 1976). The current analysis with the extended data set results in exactly the same major relationships.

8.3 Conclusions

Molecular sequences have been proven to be efficient and high quality characters for the phylogenetic reconstructions of relationships amongst the Nematoda. The phylogenetic analysis of the Nematoda with molecular characters has recovered many traditional monophyletic groups and also introduced some new relationships. It has been shown that molecular characters can add significant amounts of new information on phylogenetic relationships that may never have been deduced from morphological characters (e.g. Mermithida and Mononchida). Molecular characters can also aid in finding the true relationships where morphological character result in contradictory solutions (e.g. Comesomatidae and *Cyartonema* in Monhysterida) and find solutions for suspected paraphyletic taxa. However, molecular characters can also be misleading if they produce long branches for very divergent taxa (e.g. Trichinelloidea and outgroups). Lastly, for some taxa neither morphological, nor molecular characters can resolve the phylogenetic relationships (e.g. *Brevibucca* sp.).

Ambiguous solution for reconstructions under different models and character sets are probably the result of character conflict in the data, both within and between character sets. As seen in the individual analysis chapters, the consistency indices increase but the bootstrap proportions and posterior probabilities decrease under the exclusion of positions with gaps whilst the exclusion of parsimony-uninformative characters hardly influences the bootstrap proportions. Since positions with gaps potentially contain a lot of homoplasy, the high bootstrap proportions probably provide a lot of consistent but random phylogenetic information. It is particularly striking that there is less loss of bootstrap support when characters have been weighted according to their rescaled consistency index. Therefore, it is most likely that the different relationships amongst some nematode taxa under different models and character sets are a result of significant character conflict among the positions without gaps. In a future analysis, the consistency indices of positions with gaps can be taken into account individually and only those positions with gaps that have significantly low consistency indices need to be excluded.

The small subunit of the ribosomal RNA gene provided relatively good resolution for most parts of the phylograms, especially for the levels between families and orders. However, it was not conservative enough to reveal the root relationships amongst the Enoplea and in places too conservative to distinguish between closely related genera. Other

genes and additional taxa are needed to resolve the remaining uncertainties. Genes with a lower rate of evolution may provide a chance of recovering the root relationships and genes with a higher rate of evolution may recover relationships within taxa of more recent evolutionary origin (e.g. families, genera).

No new results can be drawn from these analyses in terms of trophic relationships and the evolution of parasitism. The major clades remained the same as in Blaxter *et al.* (1998) and Dorris *et al.* (1999). Apart from the clade of the Spirurina there is an obvious free-living ancestor to each of the parasitic species. The major addition in these data came from marine, free-living nematodes. It was hoped that those taxa may provide a clue to the origin of the Nematoda. However, other than that the ancestral nematode was most likely an Enoplea-like species no further conclusions can be drawn with regards to the primitive nematode and its ancient habitat.

9. The Potential Evolutionary History of the Morphological Characters

9.1 Characters included in the morphological and combined analyses

Probably the most comprehensive analysis of the morphological evolution in the phylum Nematoda was conducted by Chitwood & Chitwood (1950). In a series of chapters, they discussed the morphological variations of all characters used in nematode taxonomy to that date and produced the only comprehensive classification of the Nematoda until Andrássy published his approach in 1976. Since Chitwood & Chitwood (1950) any study of morphological characters has focused only on distinct groups of nematodes (e.g. Osche, 1952 [Rhabditidae]; Dougherty, 1955 [Rhabditidae]; Andrássy, 1959, 1960 [Dorylaimida]; Clark, 1961 [Enoplida]; Ferris, 1971 [Dorylaimida]; Sudhaus, 1976 [Rhabditidae]; Siddiqi, 1980 [Tylenchida]; Siddiqi, 1983 [Enoplia]; Jairajpuri & Ahmad, 1992 [Dorylaimida]; Siddiqi, 2000 [Tylenchida]). Even though these studies are extremely valuable for the understanding of the phylogenetic relationships within a certain group, they do not lead to a better understanding of the relationships of distantly-related groups of nematodes. Two problems obstruct the comparison of a certain morphological character across the entire phylum: firstly, until recently no consensus between alternative phylogenetic systems of the Nematoda could be established and secondly, some characters, especially the anterior feeding apparatus, have undergone radically different evolutionary changes with numerous instances of convergence that make homology assessments incredibly difficult.

In this study, all 180 morphological characters included in any of the analyses were mapped on the potentially best consensus tree from the morphological analysis, including all 159 characters under weighted parsimony, and on the potentially best consensus tree from the combined analysis, excluding positions with gaps in the molecular characters and all uninformative characters under weighted parsimony. Full lists of all character state changes with the respective phylogenies for both analyses are given in Appendix V (files <TableA5_1.rtf>, <TableA5_2.rtf>, <FigureA5_1.rtf> and <FigureA5_2.rtf> on CD). Additionally, the overall consensus of the partially revised phylogeny of Figure 8.1 was also taken into account. A comparison of the proposed character state changes on all consensus trees of the morphological and combined analyses would have taken too much time. The characters could not be mapped on the trees from the molecular data set because a larger number of taxa were used rendering the data sets incompatible.

The consistency indices (CIs) and weights for each morphological character including those eliminated from the analysis are found in Table 9.1. The CIs and the weights were both higher when the characters were mapped on the tree constructed with the morphological characters alone compared to mapping them on the tree constructed using both sets of characters but this difference was not significant at the 0.05 significance level when tested with the Mann-Whitney U-Test. However, if the non-normality of the data was ignored, a Paired T-Test gave significantly higher means for the CIs and weights at the $p < 0.01$ significance level when the characters were mapped on the morphological tree. The values of the consistency index on the morphological tree are probably higher because when the characters were mapped onto the combined tree the influence of potentially conflicting molecular characters had to be accounted for.

Table 9.1: Consistency indices and weights for all characters mapped on the phylogenetic reconstructions based on all characters excluding positions with gaps in the molecular characters and all uninformative characters (Combined Data) and the phylogenetic reconstruction based on all morphological characters (Morphology); both phylogenies were reconstructed under weighted parsimony using the rescaled consistency index.

	Consistency Index		Weights	
Character	Combined Data	Morphology	Combined Data	Morphology
1	0.222	0.222	0.1358	0.1358
2	0.091	0.118	0.0486	0.0756
3	0.222	0.286	0.1971	0.2623
4	0.111	0.125	0.0794	0.0938
5	1	1	1	1
6	1	1	1	1
7	0.15	0.231	0.1125	0.1968
8	1	1	1	1
9	0.333	0.333	0.2	0.2
10	0.077	0.083	0.044	0.0506
11	0.167	0.286	0.1369	0.2597
12	0.2	0.2	0.1556	0.1543
13	0.5	0.5	0.375	0.375
14	0.462	0.4	0.4206	0.3544
15	0.111	0.115	0.0815	0.0856
16	0.333	0.333	0	0
17	0.091	0.083	0.0374	0.026
18	1	1	1	1
19	0.5	0.5	0.25	0.25
20	0.5	0.5	0.4706	0.4706
21	1	1	1	1
22	1	1	1	1
23	0.739	0.773	0.6486	0.6939
24	1	1	1	1
25	0.5	0.5	0	0
26	0.143	0.091	0.0952	0.0374
27	0.5	0.5	0	0
28	0.053	0.048	0.0075	0.0023
29	0.037	0.038	0.0095	0.011
30	1	1	1	1
31	1	0.5	1	0.375
32	1	1	1	1
33	0.3	0.429	0.2408	0.3796
34	0.364	0.4	0.1	0.0571

	Consistency Index		Weights	
Character	Combined Data	Morphology	Combined Data	Morphology
35	0.1	0.1	0.025	0.025
36	0.111	0.111	0	0
37	0.125	0.125	0	0
38	0.25	0.167	0.2373	0.1523
39	1	1	1	1
40	0.667	0.667	0.5926	0.5926
41	0.125	0.125	0.102	0.102
42	0.5	1	0.475	1
43	0.167	0.2	0.1463	0.1805
44	0.125	0.182	0.1039	0.1621
45	0.083	0.125	0.0642	0.1068
46	0.143	0.143	0.0769	0.1026
47	0.5	0.5	0.25	0.25
48	0.167	1	0.0741	1
49	0.143	0.167	0.1214	0.1458
50	0.5	1	0.25	1
51	0.5	0.333	0.4	0.2
52	1	1	1	1
53	0.333	0.333	0.2593	0.2593
54	0.25	0.5	0.1818	0.4545
55	0.538	0.5	0.4986	0.4568
56	0.545	0.353	0.51	0.3032
57	0.5	0.5	0.4583	0.4574
58	0.4	0.5	0.28	0.4
59	0.75	0.75	0.6563	0.6563
60	0.139	0.135	0.0473	0.039
61	0.25	0.25	0.0357	0.0357
62	0.067	0.083	0.0383	0.0494
63	0.038	0.048	0.0156	0.0249
64	0.231	0.214	0.1817	0.1641
65	0.5	0.5	0.375	0.375
66	0.333	0.333	0.2982	0.2982
67	0.2	0.2	0	0
68	0.5	0.5	0	0
69	1	1	1	1
70	0.333	0.167	0.2381	0.0476
71	0.25	0.25	0	0
72	0.5	0.5	0.1667	0.1667
73	0.75	1	0.675	1
74	0.125	0.136	0.0625	0.0747
75	0.13	0.13	0.0652	0.0652
76	0.269	0.269	0.188	0.1854
77	0.19	0.235	0.1461	0.1928
78	0.083	0.1	0.0324	0.0438
79	0.119	0.122	0.0561	0.0583
80	0.333	0.4	0.1852	0.2667
81	1	1	1	1
82	0.143	0.133	0.0526	0.0421
83	0.091	0.091	0.0455	0.0455
84	0.1	0.1	0.0719	0.071
85	0.167	0.143	0.1071	0.0816
86	1	1	1	1
87	0.5	0.5	0	0
88	1	1	1	1
89	0.143	0.167	0	0.0278
90	0.154	0.286	0.1171	0.2547
91	0.118	0.125	0.0607	0.0667
92	0.667	0.667	0	0
93	0.111	0.111	0.0688	0.0688
94	0.333	0.25	0.2778	0.1875
95	0.667	0.667	0.5333	0.5333
96	0.038	0.038	0.0126	0.0102
97	1	1	1	1
98	0.091	0.1	0.0663	0.0757
99	0.5	1	0.4545	1

	Consistency Index		Weights	
Character	Combined Data	Morphology	Combined Data	Morphology
100	0.091	0.083	0.0545	0.0467
101	0.043	0.053	0.0165	0.0239
102	0.05	0.062	0.0212	0.0332
103	0.5	1	0.3333	1
104	1	1	1	1
105	0.067	0.071	0.0503	0.0548
106	1	1	1	1
107	0.059	0.071	0.0461	0.0548
108	0.2	0.2	0	0
109	0.059	0.071	0.0441	0.0567
110	1	1	1	1
111	0.333	0.333	0	0
112	0.143	0.143	0.0357	0.0357
113	1	1	1	1
114	0.1	0.1	0	0
115	0.111	0.143	0	0.0357
116	0.125	0.125	0.0625	0.0625
117	0.333	0.333	0	0
118	0.125	0.125	0.0455	0.0455
119	0.5	0.5	0	0
120	0.5	0.5	0	0
121	0.5	1	0	1
122	0.292	0.318	0.1612	0.1926
123	0.056	0.059	0.0341	0.0374
124	0.5	0.5	0	0
125	0.091	0.083	0.0083	0.0069
126	1	1	1	1
127	0.25	0.5	0	0.3333
128	1	1	1	1
129	0.091	0.088	0.0525	0.0486
130	0.038	0.034	0.0253	0.0209
131	0.143	0.125	0.1039	0.0833
132	0.071	0.077	0.0292	0.033
133	0.053	0.062	0.0188	0.029
134	0.062	0.053	0.0179	0.0075
135	0.029	0.027	0.0099	0.0083
136	0.071	0.067	0.0168	0.0118
137	1	1	1	1
138	1	1	1	1
139	0.5	1	0	1
140	1	1	1	1
141	0.059	0.067	0.0074	0.0083
142	0.067	0.062	0.0293	0.025
143	0.143	0.143	0.0204	0.0204
144	1	1	1	1
145	1	1	1	1
146	1	1	1	1
147	1	1	1	1
148	0.333	0.333	0.2	0.2
149	0.5	1	0.25	1
150	1	1	1	1
151	1	1	1	1
152	1	1	1	1
153	0.067	0.071	0.0459	0.0503
154	0.091	0.091	0.0625	0.0616
155	0.167	0.25	0.0972	0.1818
156	0.125	0.111	0.0278	0.0278
157	1	1	1	1
158	0.067	0.071	0.0392	0.0441
159	0.333	0.333	0	0
160	0.333	0.333	0.2222	0.2222
161	0.133	0.182	0.0993	0.1491
162	1	1	1	1
163	0.25	0.333	0.1818	0.2727
164	0.167	0.25	0.1557	0.24

	Consistency Index		Weights	
Character	Combined Data	Morphology	Combined Data	Morphology
165	1	1	1	1
166	0.5	0.5	0	0
167	0.2	0.25	0.1385	0.1923
168	0.5	0.5	0.25	0.25
169	0.197	0.211	0.0628	0.0789
170	0.269	0.298	0.1165	0.1512
171	0.091	0.091	0.0476	0.0476
172	0.04	0.043	0.0259	0.0314
173	0.083	0.118	0.0523	0.0877
174	0.125	1	0.1104	1
175	1	1	1	1
176	0.222	0.333	0.1997	0.314
177	0.5	0.5	0.4545	0.45
178	0.091	0.1	0.0404	0.05
179	0.5	0.5	0.4375	0.4444
180	0.118	0.105	0.0374	0.0239

Below, the potential evolutionary histories of all 180 characters are discussed. For those characters for which the homology assessment proved dubious alternative coding strategies are suggested. Those characters for which an assessment of the evolutionary history was possible are mapped onto the topology of Figure 8.1. In order to display all data, the figure was divided into seven individual figures, Figures 9.1 to 9.7. The plesiomorphic state for the phylum Nematoda could only be determined for a selection of characters. These states are given below the discussion in Table 9.2. In some cases, synapomorphies for a specific clade are mapped onto the tree but the plesiomorphic state could not be determined because the relationships of the subclasses of the Nematoda remain uncertain. Therefore, not for every character that is mapped onto Figures 9.1 to 9.7 the plesiomorphic state is given in Table 9.2. However, the lack of information about the plesiomorphic state is mentioned in the discussion of the characters in the text below.

During the analysis of the characters the author discovered a few mistakes in the coding:

1. # 91, reflex of testes: *Enoplus brevis* Bastian, 1865, and *Adoncholaimus fuscus* should have been coded 'C' instead of 'B'; *Enoploides brunettii* should have been coded 'A' instead of 'B'; all Rhabditina apparently have reflexed testes, but the literature recorded straight testes for *Pelodera punctata*; *Pellioditis marina* should have been coded 'B' instead of 'A'.
2. All ambiguous character states were analysed as uncertainties. However, *Strongyloides stercoralis* possess alternating generations of free-living and parasitic life stages and *Heterorhabditis bacteriophora* possess alternative generations of dioecious and hermaphroditic life states. Some alternative character states in these two species are not uncertainties but they relate to the different generations. A distinction should have been

made in these cases and the character states of the alternative generations should have been coded as polymorphic states.

3. # 103, 104 and 106, preanal papillae in *Prismatolaimus intermedius*, in the Chromadoridae and in *Nudora bipapillata*: In all cases where an 'A' was given for the specific state, an 'A' should also have been given for character 102 because characters 103, 104 and 106 are specific variations of preanal, medioventral papillae.
4. # 26, presence of rugae: *Neochromadora poecilosoma* (de Man, 1893) and *Calomicrolaimus parahonestus* (Gerlach, 1950) both possess rugae but were coded as missing them, while the presence of rugae is dubious for *Molgolaimus demani*.
5. # 23, structure of lips: The lips in *Steinernema carpocapsae* should have been coded 'L' instead of 'M'.

Many of the characters only occur in selected species in several different taxa. In all cases, there must be some genetic potential for the development of such a structure in all nematodes because it is unlikely a completely new gene has evolved recently in only a selected group. Therefore, if the plesiomorphic state for the phylum Nematoda is the absence of a character, it is expected that some underlying genetic potential was always present but the responsible gene became active only when its product provided a certain group of species with a higher fitness under the prevailing environmental conditions.

1 – 3: *The structure of the anterior sensory organs*

The plesiomorphic state of the labial sensilla is most likely to be papilliform. The papillae have become elongated to setae several times over the evolutionary history at least in the Enoplida, Panagrolaimidae and Rhaditidae. The reduction to pits in the prestome may have occurred only once in the Tylenchoidea plus Criconematoidea if these form a single clade or several times if these clades are paraphyletic. Some reversals to papilliform structures may also have occurred in these clades. A complete reduction and potential loss of labial sensory sensilla may have occurred several times in the Spirurina and Rhabditomorpha and in some freeliving Chromadorea (e.g. *Cyartonema elegans*).

There is no clear pattern for the evolutionary history of the structure of the inner cephalic sensilla. Many reversals from papilliform to setiform structures and *vice versa* seem to have occurred. However, there seems to be some pattern in the evolution of the structures of the outer cephalic sensilla but as long as it is uncertain whether the Dorylaimia or Enoplia are the more primitive subclass of the Enoplea, no estimate of the plesiomorphic state can be made. At least two major changes have occurred between

papilliform and setiform states: all Dorylaimia and the majority of Rhabditida have papilliform outer cephalic sensilla whilst the majority of the Enoplia and the Chromadorea excluding the Rhabditida have setiform sensilla. Therefore, an extension from papilliform to setiform sensilla must have occurred from the Dorylaimia to the Chromadorea and a reversal to the papilliform state in the Rhabditida. Within all groups except the Dorylaimia reversals from the majority state have also occurred for individual taxa, e.g. in the Enoplia *Adoncholaimus fuscus, Syringolaimus striatocaudatus* and the Diphtherophoroidea have papilliform sensilla whilst in the Rhabditida *Panagrobellus stammeri, Bunonema franzi* and *Pellioditis marina* have setiform sensilla. Other reversals from setiform to papilliform structures are found in selected species in most families of the freeliving Chromadorea.

4: *The position of the three sensilla circles relative to each other*

There is no clear pattern for the presence of two or three separate circles of anterior sensory organs. Many reversals are found in each individual clade. Therefore, the position of the sensilla of the second circle either between the sensilla of the other two circles or amongst the third circle is not a good character for the estimation of the evolutionary history of the Nematoda. Additionally, for many taxa no data were available introducing a lot of uncertainty in the data set.

5 – 6: *The joint nature of the anterior sensory organs*

In this data set only the two species of the Tripyloididae have joint sensilla of the inner cephalic sensilla and only *Bathylaimus assimilis* de Man, 1922, has joint sensilla of the outer cephalic sensilla. Therefore, character 6 is uninformative for this data set. However, according to Lorenzen (1981/1994) joint sensilla have been observed in species of many other taxa including the Xyalidae, Comesomatidae, Cyatholaimidae, Tobrilina and Anoplostomatidae from which species are represented in this study. No evolutionary estimate can be made before species with joint sensilla of these taxa are included in a wider study but considering that the above listed taxa were found in many different clades in the present phylogenetic reconstructions, it is likely that joining or a reduction to unjointed sensilla has occurred several times over the evolutionary history.

7: *The relative lengths of the sensilla in each circle*

Estimating the evolutionary history of relative lengths is rather difficult. However, there are some patterns in this data set. The plesiomorphic state is the presence of papillae of equal size if the Dorylaimia are the most primitive group. If the Enoplia are the most primitive group, the sensilla of the second circle are originally the longest. In both cases,

one change has occurred to the Chromadorea where the sensilla of the third circle are the longest. At least two further changes occurred in the Chromadorea; in the Monhysterida the sensilla of the second circle are usually the longest but there are some reversals in this group, and in the Rhabditida the sensilla are usually all papilliform but some reversals have occurred again where the sensilla of the second circle are the longest. Therefore, there is as of yet no clear pattern and many reversals are likely to have taken place over time. Because the use of relative lengths is difficult, this character needs to be used with care.

8: *The lateral sensilla of the second circle are shorter than the submedian sensilla*

In the present data set this character has only been observed in *Sphaerolaimus gracilis* de Man, 1884, and it therefore uninformative. It contributes only to the difference between *S. gracilis* and *S. hirsutus*. Lorenzen (1981/1994) reports a few other occasions where the sensilla of the same circle are of unequal length, however, this is a state that occurs only sporadically in the Nematoda and therefore seems to have evolved several times and in different groups.

9: *The presence of additional cephalic setae*

Additional cephalic setae have evolved at least in two groups independently. They are found in *Prismatolaimus intermedius* as well as in the Xyalidae. Without detailed ontogenetic and ultrastructural data no assessment can be made whether the additional setae are of the same evolutionary history in these two groups.

10 – 11: *The presence of subcephalic and cervical setae*

Subcephalic setae evolved several times in many of the major taxa of the Nematoda except in the Rhabditida and they are only present in selected species for a family of even the same genus, i.e. in the genus *Daptonema* they are present only in *D. hirsuitum* and *D. setosum* (which might actually be the same species) but not in the other three species represented in this study. There must be a general underlying genetic potential for these setae to be developed but they probably only develop as phenotypic states when they are of significant advantage to the animals. The presence of cervical setae is probably plesiomorphic because most Enoplia and Chromadorea except the Rhabditida have cervical setae. Cervical setae have become lost in the Dorylaimia, Rhabditida and selected other taxa.

However, since there is a sliding scale between the presence of subcephalic setae, cervical setae and somatic setae, the present homology estimate might be false and the presence of these setae has different evolutionary origins. As for the presence of additional

cephalic setae, only detailed ontogenetic and ultrastructural studies can reveal the true homology of these setae.

12: *The presence of dereids or cervical papillae*

The evolutionary origin of the dereids is controversial. They are found in most Rhabditida and Plectida and selected species of the Dorylaimia (*Mononchus truncatus* Bastian, 1865, and *Aporcelaimellus obtusicaudatus* in this study). Ward *et al.* (1975) and Lorenzen (1981/1994) argued that they are displaced sensilla from the outer cephalic sensilla circle based on findings of similar underlying neuron structures by Chitwood & Wehr (1934). Lorenzen also found lone pairs of sensilla between the amphids and the nerve ring in a similar position to the dereids in some other Enoplia. Alternatively, the dereids are an organ of independent evolutionary origin (e.g. de Coninck, 1942; Gerlach, 1966). According to the current data set the plesiomorphic state for the Nematoda is most likely the absence of dereids but their presence in the Rhabditida plus Plectida is a synapomorphy for this group.

13: *The presence of cervical alae*

Cervical alae are only found in selected species of the Ascaridoidea and its presence is probably a synapomorphy for the Ascaridoidea or even for the Ascaridomorpha. Those species lacking cervical alae in this group have probably lost them secondarily. The plesiomorphic state for the existence of cervical alae in the phylum Nematoda is probably its absence.

14: *Amphid shape*

The plesiomorphic state of the amphids is probably pocket-shaped because this is the predominant type in all Enoplea. *Prismatolaimus intermedius* is one of two exceptions in this data set because its amphids are dorsally spiral rather than pocket shaped; the other exception is the Tripyloididae that have ventrally spiral amphids. Such ventrally spiral amphids are the most common type and probably the plesiomorphic state in the freeliving Chromadorea. In the genus *Axonolaimus* de Man, 1889, the amphids have become elongated to a long loop while in the Monhysterida the spiral is reduced to a simple circle. In the Chromadoridae the amphids of many species are reduced to slits but a resemblance to a ventral loop can still be found. There are some potential reversals where the spiral amphid is reduced to a round amphid in the Plectidae (*Plectus acuminatus* Bastian, 1865) and Monoposthiidae and the round amphid evolved back into a ventrally spiral amphid in *Diplolaimelloides meyli*. In all Rhabditida and in the Trichinelloidea and in *Isolaimium* sp.

the amphids are reduced to a simple pore. If the Plectidae are part of the Rhabditida a reversal occurred here back to ventrally spiral amphids. Overall, the amphid shape is rather consistent and a few reversals or multiple reductions are easily explained. Therefore, the amphid shape is probably a good diagnostic character for the evolution of nematodes.

15: *The position of the amphids*

The position of the amphids does not present a regular pattern. The most likely plesiomorphic state is behind the cephalic sensilla in the head region. The amphids moved backwards behind the head numerous times in the Enoplia, in a number of genera of the Monhysterida and other Chromadorea and in *Teratocephalus lirellus*. They moved into the labial region in the Rhabditida but there are a number of reversals where the amphids are found around the level of the third sensilla circle. The amphids are also found in the labial region in the Trichinelloidea and *Isolaimium* sp.; both taxa also have pore-shaped amphids. The amphids are found within the third sensilla circle in selected taxa in the Chromadoridae, Desmodoridae and Rhabditina as well as in *Mermis nigrescens*. More data are needed to draw detailed conclusions on the history of the position of the amphids. However, two points become clear: pore-shaped amphids are almost always found in the labial region and the amphid position behind the head and buccal cavity evolved several times from ancestors with amphids within or just posterior to the third sensilla circle.

16: *The lateral position of the amphids*

In almost all species the amphids are found strictly lateral on each side of the body. However, there are some exceptions when the amphids are slightly displaced to the dorsal side. This data set only included three such cases from three of the four infraorders of the Rhabditida. Here, the amphids are found either in the labial region or in the posterior part of the head. This data set does not provide sufficient information on the homology of the displacements. However, it can be concluded that the plesiomorphic state is most likely a strictly lateral position of the amphids.

17: *Sexual dimorphism of amphid size*

In most species the amphids of the male and female are equal in size, at least relative to they body diameter. However, in a number of freeliving nematodes from a diverse set of taxa one sex (usually the male) has larger amphids. Therefore, equally-sized amphids are probably plesiomorphic and an increase in size has occurred several times in the evolutionary history. Sexual dimorphism is most common in spiral or round amphids, rare in pocket-shaped amphids and absent in pore-shaped amphids. The position of the sexually

dimorphic amphids is usually posterior to the third sensilla circle or posterior to the head, rarely within the third sensilla circle and never in the labial region.

18 – 19: *The presence of a cephalic capsule*

There are two types of cephalic capsules that are of different evolutionary histories. Both types are probably apomorphic features. One type is found in the Enoploidea and here in all representatives of this superfamily, the other type is found in at least two genera of the Desmodoridae, *Desmodora* and *Laxus*. At least the second type evolved several times in the Desmodoridae because at least in this data set the three species with a cephalic capsule do not form a monophyletic group.

20: *The presence of a cephalic framework*

A cephalic framework is present in almost all Tylenchomorpha. In this selection of species only *Geocenamus quadrifer* (Andrássy, 1954) is lacking such a structure. A cephalic framework is probably a synapomorphy for the Tylenchomorpha. It is most likely that it evolved only once and the lack of the cephalic framework in some Tylenchomorpha is the result of secondary loss.

21: *Dorsal and ventral cuticular inflation on the head*

Selected taxa in the Plectidae, here *Tylocephalus auriculatus* (Bütschli, 1973) and *Wilsonema schuurmansstekhoveni* (de Man, 1880, de Coninck, 1931), possess extensive cuticular structures on either side of the head. This structure is unique to the Plectidae and not homologous to the cervical alae of the Ascaridomorpha (#13) or the cuticular inflations of the Strongylidae (#149). The structure has probably only evolved once but more data from other taxa are needed to verify this observation.

22: *The presence of cephalic organs*

Cephalic organs are found in the genera *Enoplus* and *Enoploides* Ssaweljev, 1912. Their function is unknown but it is possibly of chemosensory nature. The cephalic organs probably only evolved once and they are apomorphies for these two genera.

23: *The structure of the lips*

Both phylogenies suggest that the plesiomorphic state of the lips in the Nematoda is the presence of six open, equally shaped and clearly distinguishable lips. This is in accordance with previous suggestions that the basic anterior structure of the nematodes is hexapartite. However, as can be seen from the following discussion, the presence of three lips and subsequent division of each lip to form the hexapartite structure is evolutionarily

just as plausible. Most Enoplia possess only three lips, one dorsal and two sub-ventral lips. In the Tripyloididae each lip is partially divided forming three bilobed lips. In many Dorylaimia all lips have become merged and often none of the original lip structures remain visible. Some taxa of the Dorylaimida retain six lips. Most Chromadorea possess six, open lips, two sub-dorsal, two sub-ventral and two lateral ones; many species of the Desmodoridae have lost their lips entirely. In *Teratocephalus lirellus* the six lips are extended to form six deep incisures and six extensions. In numerous taxa of the Rhabditidae the six lips have become closed. Depending on the true phylogeny, reversals to the open state may also have occurred. In the Strongyloidea the lips have also become merged to form an oral membrane but this structure is different to that of the Dorylaimida. In *Strongylus equinus* the lips evolved into a leaf crown (corona radiata). The Cephalobidae and the majority of the Panagrolaimidae retain the six lips. However, in *Panagrobellus stammeri* the six lips have become displaced from the normal axis by 30° to form six pseudolips with extensions. The Strongyloidoidea also only possess three lips but these are slightly different from those found in the Enoplia. In the majority of Tylenchomorpha the lips are again merged to form an oral disc. This disc is very distinct as it can contain pseudolips, remains of the lateral lips or four distinct pseudolips in the males. The oral disc is often off-set from the cephalic annules. In an analysis of the Tylenchomorpha alone it is probably possible to allow further distinctions between different types of oral discs. However, for the analysis of the whole phylum too many states would have been needed to distinguish further between such sub-structures. The Spiruromorpha present the fourth type of merged lips and the state should have been separated from that of the Strongyloidea. In many Spiruromorpha, the lips have entirely disappeared. In *Ascarophis arctica,* two lateral pseudolips are still detectable. In *Gnathostoma turgidum* the six lips have become merged to two fleshy, trilobed lips. All Ascaridomorpha have three lips. These can be accompanied by three interlabia and/or by denticular ridges along the inside of each lip. However, after comparing the two reconstructions it is not clear how the four combinations, three plain lips, three lips plus three interlabia, three lips with denticles and three lips with denticles plus three interlabia, have evolved from each other. Considering that the hexapartite structure only occurred at the transition from the Enoplea to the Chromadorea there may be a case to argue that in fact three lips constitute the plesiomorphic state and that each lip became divided to form six lips in the Chromadorea.

24: *The presence of labial and cephalic probolae*

Labial and cephalic probolae are only present in the Cephalobidae. They take different forms in the respective species. However, if each differentiation was coded as a different state the states had become uninformative and parsimony would have distinguished each species of the Cephalobidae as unique compared to all other taxa. The probolae probably only evolved once in the ancestor of the Cephalobidae and within this group they either evolved into elaborate tentacle-like structures or they became reduced to or remained as just small, triangular extensions on the anterior end of the worm.

25: *The cuticular pattern on the inner surface of the lips*

Transverse striations only occur in *Enoploides brunettii* in the current data set and longitudinal striations only occur in three unrelated species (*Daptonema setosum, Desmolaimus zeelandicus* and *Gnathostoma turgidum*). The different types of striations probably evolved from different structures although it is possible that the longitudinal striations evolved once in the Monhysterida from the same structure and other species of this order suffered a secondary loss. More data are needed to make a clear judgement on this character and its history.

26: *The presence of longitudinal striations (rugae) on the inside of the vestibulum*

Longitudinal striations originate from the cheilorhabdia of the buccal cavity (Chitwood & Chitwood, 1950). According to Chitwood & Chitwood (1950) rugae are the equivalent to the 12 odontia and replace the lips. They occur at least in all Chromadorida. Rugae are also present in the Monoposthiidae. Their presence in *Molgolaimus demani* is debatable; the author saw some structure resembling rugae in *Molgolaimus* but it may have been an artefact of preservation. Micoletzky (1925) reported a folded vestibulum in *Daptonema setosum* but the presence may also be debatable. The six cuticularised tooth-like structures in *Ascolaimus elongatus* Bütschli, 1874, are probably of the same origin (Chitwood & Chitwood, 1950). Therefore, the rugae may have evolved once in an ancestor of the Chromadorea and are subsequently lost either with the reduction of the buccal cavity in the remaining freeliving Chromadorea or with the development of the more elaborate rhabditoid buccal cavities. Alternatively, they evolved separately in the different families. Only more ontogenetic and ultrastructural data can provide answers to the true evolutionary origin of the longitudinal structures in the freeliving Chromadorea.

27: *The presence of transverse striations on the inside of the vestibulum*

Only two species of the Mononchida, *Mononchus truncatus* and *Mylonchulus arenicolus* feature transverse striations of the vestibulum. Without more data it is impossible to decide whether this is a common feature of this order and whether it has evolved once in the Mononchida and has subsequently been lost in some of its members.

28 – 29: *The off-set nature of the lips and head*

Both these characters were treated with suspicion from the onset and were not included in the final analyses because they introduced considerable homoplasy into the data and provided dubious topologies when they were included in an analysis of otherwise reasonable characters. Mapping of these two characters onto the two selected phylogenies verifies this suspicion. The consistency indices are very low for both characters and a large number of repeated reversals are necessary to explain the data. Therefore, the off-set nature of the lips and/or head may have several evolutionary origins. It may be related to other features of the anterior part of the worm in which case different occurrences of these structures may not be homologous. Additionally, fixation and preservation may also play a significant role in the appearance of the outline shape of the anterior end and constrictions may actually be artefacts.

30: *The orientation of the opening of the buccal cavity*

In almost all nematodes the buccal cavity opens strictly terminally at the anterior end of the body and this is also the plesiomorphic state for all nematodes. In this data set the character is uninformative because the two alternative states are only found in one species each; in *Syngamus trachea* the buccal cavity is dorsally displaced and in *Bunonema franzi* it is found at the right hand side of the head. These two states probably have different evolutionary origins and could also be coded as two separate characters.

31: *The reduction in the development of the stylet in the Tylenchomorpha*

In both representatives of the Criconematoidea and in two representatives of the Pratylenchidae (*Pratylenchoides ritteri* Sher, 1970, and *Radopholus similis*) the stylet is reduced or absent in juveniles and/or males but becomes well developed in all females. Depending on the true phylogeny the reduction may have occurred once in the Tylenchomorpha or at least twice, once in the Criconematoidea and once in part of the Tylenchoidea. As long as conflicting hypotheses exist about the phylogeny of these groups no strong conclusions can be drawn.

32: *Buccal cavity dimorphism in the females of* Aduncospiculum halicti

Two different types of buccal cavity morphologies, stenostomatous and eurostomatous, are found in the females of *Aduncospiculum halicti*. As this structure is unique to one species in this data set this character is uninformative and only contributes to the branch length but not directly to the topology of the phylogenies. After the analysis was conducted the author became aware that this buccal cavity dimorphism is relatively common in the Diplogasteromorpha (von Lieven & Sudhaus, 2000). However, it most commonly occurs in both sexes, as it does for the other species in this data set, *Pristionchus pacificus*. Therefore, the general dimorphism is probably a synapomorphy for the Diplogasteromorpha but the sexual dimorphism only evolved once in the Diplogasteridae in *Aduncospiculum halicti* and possibly some of its relatives.

33: *The structure of the oesophageal tissue around the buccal cavity*

The oesophageal tissue reaches around the buccal cavity to different degrees in different taxa and species. The coding for this character only took into account the gross morphology of this feature. However, ultrastructural analyses of the oesophageal tissue and the buccal cavity lining of the Rhabditida (e.g. de Ley *et al.*, 1995) have shown that the homologies are more fine-structured. Because little data are available for species outside the Rhabditida the fine structure was not taken into account for this character. However, a more detailed discussion of the buccal cavity structure is provided in Section 9.2 below and therefore no interpretation of this character is given here.

34: *The cuticularisation of the inner oesophageal tube*

In the ancestral nematode the oesophageal tube was probably fully cuticularised and a partial or full reduction of the cuticularisation occurred several times over the evolutionary history. The multiple reductions, even if they occur in the same part of the oesophagus, may not be homologous. Additionally, for many species no data were available. Therefore, this is probably not a very strong candidate for the reconstruction of the phylogeny of the Nematoda, at least not until more data are available.

35 – 37: *The internal structure of the oesophageal tube*

In most nematodes the oesophageal tube is tripartite throughout the whole structure; in some nematodes it is entirely rounded. However, in some species, e.g. *Xiphinema, Zeldia punctata* (Thorne, 1925) and *Bunonema franzi*, the structure changes from tripartite to round and vice versa. The plesiomorphic state is most likely the tripartite structure and the oesophagus opened up into a rounded structure in many groups several times over the

evolutionary history. Such changes may be homologous or homoplastic and only more detailed ontogenetic and ultrastructural data can reveal the true evolutionary history.

37: *The presence of a distinct isthmus between the metacorpus and postcorpus*

A distinct isthmus between the metacorpus and postcorpus is the result of a narrowing of the oesophagus and its surrounding tissue. Its presence usually coincides with the existence of a bulb in either the metacorpus (#41) or the postcorpus (#44) or both (Plectidae, Rhabditida, *Tylolaimophorus minor*). In the Rhabditida an isthmus is only missing in the Strongyloidea except in *Heterorhabditis bacteriophora* and in some Filarioidea. In those species of the Filarioidea where an isthmus is present no bulb is developed. In the majority of species the presence of an isthmus may simply be the result of the existence of one of the bulbs but that does not account for the presence in the Filarioidea. Either the isthmus of the Filarioidea is not homologous to that of the remaining Rhabditida or it is an independent structure. Therefore, if the isthmus is an independent structure it is probably absent in the nematode ancestor and evolved in the ancestor of the Plectida plus Rhabditida. In the Rhabditida it has probably been lost secondarily in the Strongyloidea and Filarioidea. Its presence in *Tylolaimophorus minor* (Enoplea) may be of separate evolutionary origin.

39: *The presence of a muscular procorpus*

The procorpus is muscular in all nematodes except the Trichodoridae. The reduction of the oesophageal musculature probably occurred once in this group and the presence of a muscular procorpus is the plesiomorphic state for the Nematoda.

40: *The presence of a muscular metacorpus*

Similar to the previous character, the metacorpus is muscular in almost all nematodes and therefore this is the plesiomorphic state. A reduction of the oesophageal musculature occurred at least twice, once in the Trichodoridae and once in *Cyartonema elegans*. In the Filarioidea and *Setaria digitata* the oesophageal musculature has been replaced by the extension of the oesophageal glands.

41: *The presence of a bulb in the metacorpus*

A bulb in the metacorpus is commonly present in the Rhabditina and Tylenchina but only rarely present in the remaining nematoda (here in *Tylolaimophorus minor* and *Laxus* spp.). The presence of such a bulb is therefore probably a synapomorphy in the Rhabditina and Tylenchina and an apomorphy in *Tylolaimophorus minor* and *Laxus* spp. The bulbs in the Rhabditina and Tylenchina are probably homologous to each other but those in the

other taxa may be of different evolutionary origin. Not all taxa in the Rhabditina and Tylenchina possess such a bulb, which probably became secondarily reduced in Panagrolaimomorpha, Strongyloidea except *Heterorhabditis bacteriophora* and some species of the Cephalobidae and Rhabditidae. Depending on the true phylogenies of the latter two families, reversals from the absence back to the presence of a bulb may have occurred.

42: *The presence of a bulb with valve in the metacorpus*

Since the ancestral nematode probably did not have a bulb in the metacorpus (see previous character) it is also unlikely that it had a valve in this position. A valve in the bulb of the metacorpus is only found in the Diplogasteroidea and Tylenchomorpha in the present study. Its presence is probably an apomorphy of the Diplogasteroidea. As long as the true phylogeny of the Tylenchina is unknown, it is uncertain whether the valve is a synapomorphy of the Tylenchina with secondary losses in the Panagrolaimomorpha and Cephalobidae or whether it is a synapomorphy of the Tylenchomorpha if these are monophyletic. The valve is probably not of the same evolutionary origin in the Diplogasteroidea and Tylenchomorpha and evolved twice over the evolutionary history.

43: *The presence of a muscular or glandular postcorpus*

As with the structure of the procorpus (#39) and metacorpus (#40), the postcorpus is also muscular in most nematodes and this is probably the plesiomorphic state for the phylum. The musculature of the postcorpus is replaced by the oesophageal glands in several groups: Trichodoridae, Dorylaimida, *Isolaimium* sp., Filarioidea including *Setaria digitata*, Diplogasteroidea and Tylenchomorpha. Because of the different nature of the oesophageal glands in the different groups (see below) the different types of glandular postcorpi are probably of different evolutionary origin and synapomorphies for the respective taxa. In the current set of taxa no reversals have to be accounted for in any of the groups apart from the Tylenchomorpha. In the last group reversals have occurred if the Tylenchomorpha are not monophyletic within the Tylenchina.

44: *The presence of a bulb in the postcorpus*

A bulb in the postcorpus is found in many Chromadorea but not in the Enoplea (except *Syringolaimus striatocaudatus* and *Campydora demonstrans*). Instead, in the Dorylaimida excluding *Belondira apitica* and in *Isolaimium* sp. the postcorpus features an elongated swelling. Therefore, the plesiomorphic state in the Nematoda is probably the absence of a bulb. Independently from each other, the postcorpus either evolved into an

elongated swelling containing the oesophageal glands or into a bulb. The current reconstruction suggests that within the Chromadorea a bulb may have evolved in its ancestor and was secondarily lost in the Cyatholaimidae, in the Monhysterida with a reversal to the synapomorphic state for the Chromadorea in the Linhomoeidae and Comesomatidae, in the Axonolaimidae, in the Spirurina with a reversal to the synapomorphic state in *Cruzia americana*, in the Aphelenchoidea, Hoplolaimidae, *Meloidogyne incognita*, Pratylenchidae and in the Strongyloidea except *Heterorhabditis bacteriophora*. Alternatively, the bulb may have evolved independently in the Chromadoridae, Microlaimoidea, Desmodoridae, Linhomoeidae, Comesomatidae, Plectidae, *Cruzia americana*, Rhabditina, Anguinidae, Criconematoidea and Belonolaimidae with a reversal in the Strongyloidea. Whilst the first option is the more parsimonious one the second option is just as feasible. In the first case the presence of the bulb is a synapomorphy for the whole of the Chromadorea and of the same evolutionary origin, in the second case it is a synapomorphy for each individual clade whose species have developed a bulb and the different occurrences of a bulb may be of different evolutionary origins.

45: *The presence of a valve in the postcorpus*

Under almost all circumstances a valve is only present if a bulb is present in the postcorpus. In this data set a valve is present without the existence of a bulb in *Longidorus elongates* (de Man, 1876), *Xiphinema* spp. and *Axonolaimus spinosus*. In the former two species the postcorpus features an elongated swelling. No valve is present in the bulb in *Syringolaimus striatocaudatus, Cyartonema elegans*, the Desmodoridae except *Acanthopharynx micans* and *Desmodora communis*, Microlaimoidea, *Terschellingia longicaudata*, Comesomatidae, *Halicephalobus gingivalis*, Tylenchomorpha (where a bulb is present) and the Diplogasteroidea. The valve was probably absent in the ancestral nematode and evolved either once in the Chromadorea with subsequent losses in some groups or it evolved separately in the different taxa of the Chromadorea. In either case, it probably evolved separately in the taxa of the Enoplea. Therefore the valves of the Enoplea and Chromadorea are probably not homologous. Considering the proposed system from Figure 8.1, the valve in the postcorpus probably evolved several times in the Chromadorea where it constitutes synapomorphies for the respective groups. A few studies (Albertson & Thomson, 1976 [*Caenorhabditis elegans*]; Zhang & Baldwin, 1999 [*Diplenteron* sp.], 2000 [*Zeldia punctata*], 2001 [*Teratocephalus lirellus*]) have looked at the ultrastructure of the

postcorpus and when more data are available it can be determined whether the valves are of the same evolutionary origin in the different taxa.

46: *The presence of plasmic interruptions in the postcorpus bulb*

Plasmic interruptions are found in the postcorpus bulb in selected species of the Desmodoridae, Chromadoridae and Monoposthiidae. The ancestral nematode probably did not have plasmic interruptions especially if it did not have a bulb. The plasmic interruptions probably evolved separately in the three families and were lost secondarily in some of the species in which case they may not necessarily be homologous amongst the three families. Alternatively, the plasmic interruptions may have evolved in the ancestor of the Chromadorea for which they would be synapomorphic and secondarily lost in a number of families and at the point when the Monhysterida evolved.

47: *The presence of tooth-like cuticularisations in the postcorpus bulb*

Strong cuticularisations taking on the resemblance of teeth are found only in three species of two different orders in this data set (*Plectus acuminatus, P. aquatilis* Andrássy, 1985, and *Zeldia punctata*). Therefore, this structure probably evolved twice independently and the two occasions may not be homologous.

48: *The presence of a double haustrulum in the postcorpus bulb*

A strong cuticularisation forming a double haustrulum in the postcorpus bulb is only found in selected species of the Rhabditidae. In the reconstruction using only morphological characters this character unites all species with this feature while in the reconstruction using the combined analysis, species with the double haustrulum are found in both major clades of the Rhabditina and several reversals are necessary to explain the data. Because this feature is very conservative in all species it probably only evolved once and was subsequently lost numerous times but more ultrastructural data will be able to provide a clarification of the homology of the haustrulum.

49: *The presence of oesophageal tubes*

Because of the lack of data no clear assessment can be made of the evolutionary history of the existence of three oesophageal tubes at the end of each of the arms of the tripartite oesophageal lumen. They either evolved somewhere in the Chromadorea and were lost secondarily in many taxa or they evolved independently in the different taxa of the Chromadorea where they are present. In any case they do not seem to be present in the Enoplea; however, this may be a false observations based on insufficient data.

50: *The presence of a stichosome*

A stichosome is present in the Mermithidae and Trichinelloidea. As long as the relationship between these two groups of the Dorylaimia are not certain no estimate can be made of the evolutionary history of the stichosome. However, since this is such a unique structure it suggests a unique origin and possible reversals or losses in those related taxa that do not posses a stichosome.

51: *The reduction in the development of the postcorpus bulb in the Tylenchomorpha*

The bulb in the postcorpus is weakly developed in the males and juveniles in the same five species of the Criconematoidea and Tylenchoidea in which the stylet is reduced or absent as well as in the infective larval stage of *Steinernema carpocapsae*. The reduction probably occurred at least twice, once in the Strongyloidoidea and once in the Tylenchomorpha, but as long as conflicting hypotheses exist about the phylogeny of the Tylenchomorpha no strong conclusions can be drawn about the evolutionary origin of this feature.

52: *The sexual dimorphism of the buccal cavity and oesophagus of* Calyptronema maxweberi

Calyptronema maxweberi is a peculiar species in that the buccal cavity and oesophagus of the male are markedly different from that of the female. The buccal cavity and oesophageal musculature is markedly reduced in the male. The reduction probably occurred uniquely in this species. This character is uninformative for the phylogenetic reconstruction of this data set because it only occurs once.

53: *The number of oesophageal glands*

All nematodes possess at least three oesophageal glands but the majority of the Dorylaimia and Triplonchida possess five glands. In the Enoplida, dinuclear subventral glands are found in the Oncholaimidae and there are suggestions that there are possibly five glands in the genus *Enoplus*. However, few data are available for the majority of the Chromadorea. It is only known that the Ascaridomorpha, Tylenchina and Rhabditina possess three glands and that the glands of the Spiruromorpha are polynuclear. Without a clear knowledge of the relationships between the Dorylaimia and Enoplia and the lack of data for the freeliving Chromadorea it is difficult to estimate whether the plesiomorphic state is the presence of three or five glands. Because the Oncholaimidae (Enoplia) possess dinuclear subventral glands it was suggested by Chitwood & Chitwood (1950) that the primitive state was the presence of three glands and that the subventral glands became

doubled in selected groups of the Enoplea. If, however, the plesiomorphic state is the presence of five glands two subventral glands have become lost secondarily or merged in several groups of the Enoplia and possibly once in the Chromadorea. In either case, the multiplication of nuclei in the oesophageal glands of the Spiruromorpha must be of separate evolutionary origin compared to that of the five glands.

54: *The position of the oesophageal glands in the oesophagus or body cavity*

In most nematodes the oesophageal glands are found entirely in the posterior part of the oesophagus and this is probably also the plesiomorphic state. In the Aphelenchoidea and Tylenchoidea excluding the Belonolaimidae the oesophageal glands extend over the cardia into the body cavity parallel to the intestine. This extension either evolved once in the ancestor of the Tylenchomorpha and was reduced subsequently in some taxa, or it evolved separately in the two superfamilies and was reduced in the Belonolaimidae. In either case it is likely that this character has a common evolutionary origin because the feature is very similar in all groups.

55 – 56: *The position of the opening of the oesophageal glands*

The openings of the oesophageal glands are found in a variety of positions along the entire length of the oesophagus. There are only two clear patterns: firstly, the dorsal gland always opens anterior to the subventral glands and secondly, that if a second pair of subventral glands exists they open posterior to the first pair. There is no clear pattern as to how the openings of the glands have moved over the evolutionary history. However, it is possible that the opening of the dorsal gland at least was originally found in the anterior part of the oesophagus and subsequently moved further posterior because it is found in the anterior part in all Enoplia and freeliving Chromadorea. The subventral gland openings are more likely to have evolved further posterior in the oesophagus and moved anterior and into the teeth only in the Enoplida. All glands probably became displaced to open through the teeth secondarily but early in the evolutionary history because this position is found in *Enoploides brunettii* and in the Oncholaimoidea. Overall, more data are needed to clarify the evolutionary history of the oesophageal gland openings.

57: *The structure of the cervical gland*

The plesiomorphic state of the cervical gland structure is probably the presence of a unicellular and uninuclear gland cell with a duct opening to the ventral side of the animal. This type is present in all nematodes except the Plectidae and Rhabditida. In the Plectidae three glands are found that are connected to the terminal duct that is cuticularised at its

end. The three glands possibly evolved from the single gland by multiplication. In the evolution of the Rhabditida the single gland became double and two lateral canals developed that can reach almost the entire length of the body. In some groups the glands were finally lost and in others one of the canals was partially or completely reduced. There are many different degrees of development of the lateral canals that have been discussed extensively by Chitwood & Chitwood (1950). The rhabditoid types either evolved straight from the plesiomorphic type or from the plectoid type depending on whether the Plectidae are part of the Rhabditida or their sister taxon. Contrary to Chitwood & Chitwood (1950) and other popular opinion the rhabditoid system of elaborate canals is not the plesiomorphic state of the cervical gland system but a specialised feature that evolved later from the simple gland found in the more ancestral nematodes. Alternatively, the two systems have entirely different evolutionary origins. In this case the unicellular gland is unique to aquatic taxa and was lost along those lineages that became fully terrestrial or parasitic and the system with lateral canals only evolved in the ancestor of the Rhabditida.

58 – 59: *The position of the unicellular cervical gland*

When only a single gland is present this may extend into the intestinal region and may lie either ventrally or (sub-) laterally to the digestive tract. In most species it is observed on the ventral side of the body but in the genus *Enoplus* it is on the left side, in *Viscosia viscosa* on the right and the glands in the Plectidae are found ventrally and subdorsally. Without more data the plesiomorphic state cannot be determined. The gland is found restricted to the oesophageal region in the genus *Enoplus* and in *Pontonema vulgare* in this data set. Its presence and position are uncertain in most Triplonchida and Dorylaimia and it is found in the intestinal region in those Chromadorea where a unicellular gland has been observed. Because there are so much missing data it is uncertain where the gland was situated in the primitive nematode.

60 – 61: *The position of the excretory pore*

In most species the excretory pore is found in the oesophageal region and this is probably the plesiomorphic state. Over evolutionary time it moved into the head region to open between the lips or further back into the intestinal region. The respective changes probably occurred several times independently because the anterior position is found in several different orders and the posterior position is found in at least three different suborders of the Rhabditida, the Tylenchomorpha, Rhabditomorpha and Diplogasteromorpha. There is, however, no clear indication of a pattern with respect to the position in the oesophageal region anterior or posterior to the nerve ring.

62: *The presence of a hemizonoid*

The hemizonoid was probably absent in the primitive nematode. In this data set it is almost exclusively found in the Rhabditida with the exception of the Longidoridae and *Tylocephalus auriculatus* (Plectidae). In the Rhabditida it is found at least in parts of the Cephalobidae, Tylenchomorpha and Rhabditidae but no data were available for a number of taxa. However, its presence is probably a symplesiomorphy for the Tylenchina and possibly for the Rhabditina. It has been lost secondarily in those species of the Tylenchina were it is absent. Only more data can determine the real extent of the presence of a hemizonoid in the Nematoda. In this data set this character has a very low CI and is therefore probably not a very good character for the reconstruction of the phylogeny of the phylum.

63: *The presence of a cardia*

In most nematodes some structure resembling a cardia can be found. However, in many species there does not seem to be such a structure or it has been obscured, e.g. by the overlapping oesophageal glands. Originally the author coded this character with a number of states for different shapes but it proved very difficult to find any coherent homology assessment. The simple coding for the presence or absence of a cardia is probably not a very good alternative because the character has a very low CI.

64: *The shape of the cardial lumen*

In most nematodes the lumen of the cardia is triradiate or triangular as is the lumen of the oesophagus. This probably represents the plesiomorphic state. However, the cardial lumen opens up to take on a simple round cross-section in the Chromadoridae, Tylenchomorpha and several other species. In other groups, such as the Diphtherophoroidea, Tripyloididae, Monhysterida, Plectidae, Spiruromorpha, part of the Dorylaimida and Mononchida, and in *Mermis nigrescens* and *Teratocephalus lirellus* it became dorsoventrally flattened and at least in *Mononchus truncatus* it is vertically flattened. A flattening of the lumen, therefore, probably occurred several times during the evolutionary history of the Nematoda and is of independent origin in the different taxa.

65: *The presence of a posterior appendage on the cardia*

In a number of species of the Ascaridoidea the cardia possess a posterior appendage that can reach far back in the body cavity along the intestine. This feature may have developed once and has subsequently been lost in other taxa or it arose several times

independently in the Ascaridoidea. Only more data and a clearer knowledge of the true phylogeny of this group can provide a sound understanding of the evolution of this feature.

66: *The presence of a cardial valve*

In this data set a valve in the cardia has been observed in the Mononchidae, in most Tylenchomorpha and *Haemonchus contortus*. However, it is possible that a valve has been overlooked in many other species because it tends to be very small. A cardial valve does not occur in conjunction with a valve in the postcorpus of the oesophagus but probably has the same function of controlling the entrance of food into the intestine and of preventing gut contents being regurgitated. Judging by the current presence of such a valve it probably developed several times in the Nematoda and was absent in the primitive nematode.

67: *The presence of cardial glands*

Cardial glands have only been observed in a few species so far but may have been overlooked in many species, as they can be very small and difficult to observe. Without more data, no estimate of the evolutionary history can be made.

68: *The presence of symbiotic bacteria in the gut*

Symbiotic bacteria are present in the guts of many animals and in nematodes they are found in diverse groups, in this case in the two species of entomopathogenic nematodes, *Steinernema carpocapsae* and *Heterorhabditis bacteriophora*. Other nematodes contain bacteria in other parts of the body (see #148, on surface). The bacteria have probably been acquired at separate occasions during evolution of this phylum. Internal bacteria in parasitic species are closely related to the host-parasite interaction.

69: *The presence of a trophosome*

A trophosome is only present in the Mermithida. It has probably developed once in evolutionary history and was probably absent in the primitive nematode. This character is uninformative in this data set.

70: *The presence of an intestinal caecum or diverticulum anterior to the cardia*

An intestinal caecum or diverticulum that reaches anterior to the cardia is present in many Ascaridomorpha. It was probably absent in the primitive nematode but until the true phylogeny of the Ascaridomorpha is known no estimate can be made as to whether this feature developed once in this group or had several independent origins.

71: *The presence of an intestinal caecum posterior to the anus*

An extension of the intestine into the caudal region is found in a number of species of the Tylenchomorpha and Diphtherophoridae (here *Tylolaimophorus minor*). This feature developed at least twice, once in each taxon, but until the true phylogeny of the Tylenchomorpha is known it is unclear whether it has multiple origins in the latter group. There was probably no caecum of the intestine posterior to the gut in the primitive nematode.

72: *The reduction of the function of the rectum and anus*

The anus has been reduced in the adult nematode at least twice over the evolutionary history, in the Mermithida and Criconematoidea. In the two entomopathogenic nematodes the infective juvenile state has no functional rectum or anus. In all instances the reduction is related to the parasitic habitat of the animals. The reduction probably has several independent origins in the phylum Nematoda. The primitive nematode probably had a fully functional anus, especially if it was freeliving and parasitism arose from freeliving species as it is widely suggested (Blaxter *et al.*, 1998, 2000; Dorris *et al.*, 1999).

73: *The presence of a pre-rectum*

A true pre-rectum is only present in the Dorylaimida and Mononchidae, where it might be secondarily reduced (here in *Xiphinema theresiae* Stocker & Kruger, 1987). In the Trichodoridae a different type of pre-rectum can be present but judging by the closer relationship of the latter taxon to the Enoplida the two types of pre-recta are probably of separate evolutionary origin. The primitive nematode probably did not have a pre-rectum.

74 – 75: *The position of the gonads relative to the intestine*

Lorenzen (1981/1994) introduced this character. Previously, the positions of the gonad relative to the intestine had not been recorded for many species. The feature is also difficult to determine in preserved specimens because the inner organs become difficult to distinguish after long preservation times. Therefore, the author was hardy ever able to add additional data and almost solely relied on the data that were available in the literature of which there was little available for the non-aquatic species. The position of the ovaries is generally consistent within families and even orders but there are many changes from right to left and vice versa and also to subventral positions between major taxa. Without knowledge of the states for the Dorylaimia and Rhabditida no estimates can be made of the evolutionary history of this character.

76: *The number and vertical position of the ovaries*

The presence of two opposed, fully developed ovaries is most probably the plesiomorphic state for the phylum. Investigations into the development of germ cells found that even in such species that only develop one ovary genital primordia are present for both ovaries (Chitwood & Chitwood, 1950). From this state several different reductions have occurred. The anterior or the posterior ovary has either been lost completely or it has become reduced to form a pre- or post-vulval sac. The Spirurina are united by the synapomorphy of two posterior ovaries; only *Ascarophis arctica* retains the plesiomorphic state of two opposed ovaries. The move of the anterior ovary to face posterior is connected to the anterior position of the vulva (see #79A and #80A). In the Tylenchina the posterior ovary has either been reduced to a post-vulval sac or lost in many taxa. The most parsimonious solution for the ancestor of this suborder under the current data set is the presence of only one anterior ovary with a post-vulval sac. However, considering the more likely evolutionary scenario of a reduction of one ovary rather than the re-development of a previously reduced ovary suggests that two opposed ovaries may be the ancestral state for the Tylenchina. At least in *Globodera pallida* and *Meloidogyne incognita* the vulva is found terminally (see #79F and #80C) and the two ovaries are effectively placed anteriorly. In the Rhabditina a reduction of the posterior ovary has also taken place many times and at least in *Syngamus trachea* the vulva moved forward to provide space for two backward-facing ovaries. A multiple reduction of one of the ovaries has therefore occurred many times in the evolutionary history and in some cases a previously reduced ovary has become re-developed.

77: *The shape of the ovaries*

The majority of the Enoplea and the Chromadorida have antidromously-reflexed ovaries, which are probably the plesiomorphic state. In some groups, such as the Trichinelloidea, *Mermis nigrescens, Isolaimium papillatum*, the Microlaimidae, Axonolaimidae and the majority of the Monhysterida the ovaries have become outstretched. In the Rhabditina and Panagrolaimomorpha the flexure changed to the homodromous state. In a few species the flexure turned back to the antidromous shape (e.g. *Bunonema franzi*, Diplogasteroidea, *Panagrellus redivivus, Heterorhabditis bacteriophora*). In the Cephalobidae the flexure is found at a novel position resembling the antidromous state; this state probably only developed once in their ancestor. In the majority of the Tylenchomorpha the ovaries are again outstretched. In saccate females (see #127) and in the majority of Spirurina the ovaries lie in multiple coils in the body cavity of the

females. It is quite plausible that the ovaries may change from a reflexed position to an outstretched position depending on how much space is available parallel to the intestine. Apart from several reversals in the Rhabditidae between the two reflexed states this character seems rather consistent in the phylum despite its low CI. Nevertheless, this character was also only introduced by Lorenzen (1981/1994) and might still prove suitable for the phylogenetic reconstruction of the Nematoda.

78: *The shape of the tip of the ovary*

The tip of the ovary can be folded over in a variety of species in most major taxa of the Nematoda. There is no clear pattern but there are also no reversals in this data. The fold is probably related to the available spaces and it can occur in all types of ovaries. The primitive nematode probably did not have ovaries with folded tips.

#79 – 80: *The position of the vulva along the body*

In the majority of nematodes, the vulva lies in the intestinal region somewhere roughly between the middle of the body to the beginning of the last third of the body length. This is probably the plesiomorphic state. In a few Tylenchomorpha (here *Globodera pallida* and *Meloidogyne incognita*), the vulva is terminal. In *Trichinella spiralis* and in the majority of the Spiruromorpha and some Ascaridomorpha, the vulva lies in the oesophageal region. Since the vulva is found in the anterior intestinal region in the majority of the other Spirurina the forward move of the vulva probably occurred once over evolutionary time in this subclass but it moved past the cardial region only in selected species. The anterior position in the Trichinellida is probably of separate evolutionary origin. The posterior movement probably occurred several times in different groups. In the Tylenchomorpha, there is an overall tendency of a posterior position of the vulva but it probably moved to the terminal position at several times in the evolutionary history. However, a finer analysis with separate characters for the relative position of the vulva along the intestine proved very inconsistent.

81: *The Demanian organ*

The Demanian organ is only found in certain species of the Oncholaimidae, here in *Adoncholaimus fuscus* and *Viscosia viscosa*. It probably is a synapomorphy for this taxon and was lost secondarily in other species (here *Pontonema vulgare*).

82 – 83: *Reproduction: oviparity versus ovoviviparity*

Most nematodes lay eggs from which the larvae hatch in the environment. However, in all major clades except the Tylenchomorpha some species release live young or lay eggs

containing well-developed larvae. The change from oviparity to ovoviviparity probably occurred several times and depending on the true phylogeny of the Rhabditidae some reversals may also have occurred. It is possible that in the Spiruromorpha a succession occurred from species laying eggs to those laying eggs containing well-developed larvae to species that release live young. In any case, oviparous reproduction is probably the plesiomorphic state for the Nematoda.

84: *The presence of spermathecae*

Spermathecae are found in a few groups in the Enoplea (*Trichodorus primitivus* [de Man, 1880], *Mermis nigrescens* and *Xiphinema* partim) and in many Rhabditida, especially in the Cephalobidae and Tylenchomorpha. The structure of the spermathecae varies and therefore it is possible that they are of different evolutionary origin, in which case this character should have been divided into individual characters. However, the primitive nematode probably did not possess any spermathecae.

85: *The presence of an ovijector*

An ovijector is found in a variety of families (*Xiphinema*, *Tylocephalus auriculatus, Gnathostoma turgidum*, Filarioidea, *Cruzia americana*, Trichostrongylidae) but many of their relatives do not possess such a structure. The primitive nematode probably did not possess this organ and it is likely that it developed independently in the different families.

86: *The presence of the Z-organ*

A so-called Z-organ is only found in the genus *Xiphinema* but not in all species and when it has been observed it can be more or less well developed. This organ is probably a synapomorphy for the *Xiphinema* and was secondarily reduced in some species.

87: *The presence of vulval papillae*

In this data set vulval papillae are only found in two species of the Dorylaimida, *Aporcelaimellus obtusicaudatus* and *Paractinolaimus macrolaimus*. Without more data it is difficult to say if the presence of vulval papillae is a synapomorphy for all Dorylaimida or only a subgroup thereof or whether they are of different origins altogether.

88: *The presence of vulval flaps*

Elaborate flaps covering the vulva opening are only found in selected taxa of the Trichostrongylidae. Without more data it is not possible to determine whether the flaps are a synapomorphy of the Trichostrongylidae or whether they evolved several times in the selected species.

89: *The presence of vulval glands*

Vulval glands have been found in a variety of species from different major taxa. There are still too few data to determine whether they developed once in the primitive nematode or whether they are of separate evolutionary origin. As with the presence of the cardial glands (see #67) it is possible that vulval glands have been overlooked in many species.

90: *The number and position of testes*

Nematodes either possess two opposed testes or one anterior or, more rarely, one posterior testis. As discussed for the number of ovaries (# 76), even those species with only one testis possess two genital primordia leading to the conclusion that a single testis is lost and the ancestral nematode possessed two testes. This is in accordance with the current phylogeny. Most Enoplea and many freeliving aquatic Chromadorea possess two opposed testes but some Enoplia, Chromadorida, Desmodoridae and all Rhabditida only have one anterior testis. Some species of the Cyatholaimidae, Microlaimoidea and Plectidae have two opposed testis forcing some reversals. The Trichinelloidea only possess one posterior testis and *Meloidogyne incognita* (and also other *Meloidogyne* species) may have two anterior testes when the juvenile originally developed as a female and then underwent a late sex change to become a male. Despite the relatively recent introduction of this character by Lorenzen (1981/1994) its evolution seems to be quite consistent in the phylum.

91: *The structure of the testes*

There is no real pattern as to when the testis or its tip is reflexed. A reflex is found in most orders except the Chromadorida, Microlaimoidea and Desmodoridae. The primitive nematode probably had no flexure but this cannot be taken as certain. However, no distinction was made between a testis that was only reflexed at the tip of the germinal zone and one that was reflexed in the growth zone because in many cases no figure was given and the descriptions only state 'testis reflexed'.

92: *The number of spicules*

Almost all nematodes possess two spicules and that is probably also the plesiomorphic state. At least twice the spicules were lost independently, once in *Monoposthia costata* and in the Trichinelloidea. In the latter case, there may be one spicule present (*Trichuris muris*) or none (*Trichinella spiralis*).

93 – 94: *The size and shape of the spicules*

In most nematodes both spicules are of the same size and shape and these are probably the plesiomorphic states. However, in some species of the genus *Plectus*, in *Calyptronema maxweberi*, and in *Teratorhabditis synpapillata* the two spicules are of different size. In some species of the genus *Enoplus* and in many species of the Spirurina, the two spicules are not only of different sizes but also of different shapes. In the Ascaridomorpha a number of reversals from equal to unequal spicules and vice versa are found. However, this applies more to the size than the shape. The change in spicule size and shape probably occurred several times in the evolutionary history of this character.

95: *The fusion of the spicules*

In most nematodes the spicules are separate and this is probably the plesiomorphic state. However, at least in *Haemonchus contortus* the spicules are fused along almost their entire length while they are partially fused distally in the Rhabditidae clade 1 and in *Nematodirus battus*. The fusion of the spicules either occurred once in the Rhabditomorpha with several reversals back to the plesiomorphic state or at least twice, once in the Rhabditidae clade 1 and once in the Trichostrongylidae. The full fusion in *Nematodirus battus* is probably an extension to the partial fusion seen in the other species.

96: *The presence of ventral alae on the spicules*

Ventral alae are only present in the Chromadorea and in *Syringolaimus striatocaudatus*, at least in this data set. Within the Chromadorea, ventral alae developed many times in the different orders and families and at least in the Panagrolaimomorpha and Spirurina reversals seem to have occurred. The primitive nematode probably did not possess ventral alae on the spicules as they are absent in almost all Enoplea.

97: *The presence of a sheath around the spicules*

A (thorny) sheath around the spicules is found only in *Trichuris muris* in this data set. The primitive nematode probably did not possess such a sheath and it only developed in the ancestor of the genus *Trichuris*. This character is uninformative for this data set.

98: *The presence of a gubernaculum*

A gubernaculum is present in most major taxa with the exception of the Dorylaimia (excluding the Mononchida). It has been lost several times over the evolutionary time, especially in many species of the Spirurina. Therefore, the gubernaculum was probably present in the primitive nematode. However, the structure of the gubernaculum differs in

the different taxa. It may be a simple guiding piece or it can be constructed from several different pieces. It is difficult to construct true homologies between the different types but all gubernacula may actually have developed from the same original structure.

99: *The presence of two lateral accessory pieces*

Instead of a true gubernaculum, the Dorylaimida possess two lateral accessory pieces that take the function of the gubernaculum to guide the spicules out of the cloaca. Such accessory pieces are also found in the Mononchida which therefore possess both structures. The lateral accessory pieces probably developed in the ancestor of the Dorylaimia and have subsequently been lost in the Mermithida and Trichinellida.

100: *The presence of apophyses on the gubernacula*

One or two apophyses on the gubernaculum are not commonly observed in the Enoplea (here only *Enoplus anisospiculum* Nelson *et al.*, 1972) but they are common in the Monhysterida, Axonolaimidae, Plectidae and some Desmodoridae. They are almost entirely absent in the Rhabditida, which usually possess very simple gubernacula but an apophysis can be found in some species of the Panagrolaimidae. The apophyses are lost in some species of the Monhysterida and Plectidae. The primitive nematode probably did not possess an apophysis and this character may have arisen several times independently.

101: *The presence of rectal glands*

Rectal glands are found in many species of the Rhabditida and some other species. As with the other glands, rectal glands may have been overlooked when species were examined. Within the Rhabditida several reversals may have occurred between the presence and absence of rectal glands. As long as it is not clear whether rectal glands are really rare in the non-Rhabditida no assessment can be made of its plesiomorphic state. However, they may be a synapomorphy for the Rhabditida.

102: *The presence of a ventral row of preanal papillae*

Preanal papillae are found in a ventral row in most groups of the Nematoda except in the Rhabditida. In the Rhabditida they are only found in *Panagrellus redivivus*, at least in this data set. In the Enoplea they are present in all Dorylaimia and most Triplonchida (except *Prismatolaimus intermedius* where they reach the oesophageal region) but absent in all Enoplida. Therefore, as long as the true relationship between the two subclasses of the Enoplea is uncertain no estimate can be made of the plesiomorphic state for the Nematoda. In the Chromadorea several reversals from a lack of such papillae to its presence and vice versa have occurred.

103: *The presence of a ventral row of cup-shaped preanal papillae*

Cup-shaped ventral papillae are only present in the Chromadoridae and probably constitute a synapomorphy for this family. They are probably derived from the papillae described in # 102 but the papillae in the Chromadoridae are all of exactly the same type whilst the other papillae summarized in # 102 are variable in their precise structure.

104: *The presence of a ventral row of preanal papillae that reaches all the way into the oesophageal region*

Ventral papillae usually do not exceed the level of the end of the anterior testis but in *Prismatolaimus intermedius* these papillae reach all the way into the oesophageal region. They are probably of the same origin as those described in # 102 and the extended range is an apomorphy for this species. This character is uninformative and its inclusion therefore does not alter the tree topology.

105: *The presence of two subventral rows of preanal papillae*

Two rows of preanal papillae are found in a number of Dorylaimida, in the Mononchida and *Mermis nigrescens* as well as many Rhabditida. They are absent in most Tylenchomorpha and they have been lost secondarily in some Spirurina. In this data set they are also found in *Metachromadora remanei* and *Trischistoma monohystera* of the freeliving, non-Rhabditida nematodes. Amongst the Rhabditida, especially in the Strongyloidea, these papillae form the bursal rays when a bursa is present (see also # 109). As with # 102, no estimate of the plesiomorphic state can be made before the true phylogenetic relationship between the Enoplia and Dorylaimia has been established.

106: *The presence of two ventral preanal papillae with conical tips and a pore*

In *Nudora bipapillata* the ventral papillae are specifically modified. They are rather large compared to other papillae and possess a conical cone as well as a clearly visible pore. These papillae are probably also of the same evolutionary origin as those described in # 102. This character is uninformative and its inclusion therefore does not alter the tree topology.

107: *The presence of adanal papillae*

Adanal papillae are mainly present in the Rhabditida and here in most Spirurina and Rhabditina and some species of the Panagrolaimomorpha, Cephalobidae and Aphelenchidae. In some species of the Spirurina and Rhabditina they have been lost secondarily. They are also commonly observed in the Dorylaimida and Trichinelloidea and

some other species of aquatic nematodes. It is uncertain whether adanal papillae were present in the primitive nematode as their presence in the Enoplea is scattered.

108: *The presence of a ventral row of postanal papillae*

One row of postanal ventral papillae is found in a few taxa of freeliving nematodes, including the Mononchida, Mermithidae, Plectidae, Chromadoridae and Monhysteridae, however, its presence is inconsistent. Therefore, no estimate about the plesiomorphic state can be made.

109: *The presence of two subventral rows of postanal papillae*

Two rows of postanal ventral papillae are found in a variety of nematodes. They are present in some Enoplia, in the Mononchidae, *Isolaimium papillatum* and *Mermis nigrescens* suggesting that this may also be the plesiomorphic state. They are also found in some Monhysteridae, Plectidae and in almost all Rhabditida except the Tylenchomorpha but they are found in the Aphelenchoidea. Because they are absent in almost all aquatic Chromadorea it is not possible to determine whether the papillae of the Enoplea and Rhabditida are of the same evolutionary origin. If postanal papillae occur in conjunction with a bursa they usually form the bursal rays (see also # 105).

110: *The presence of ventral and dorsal postanal papillae*

In *Panagrellus redivivus* postanal papillae are found not only on the ventral surface but also on the dorsal surface of the tail. The ventral papillae are probably homologous to those of # 108 but the dorsal papillae may be of separate evolutionary origin to the other types of papillae. This character is uninformative and its inclusion therefore does not alter the tree topology.

111: *The position of the papillae in the cuticle*

The papillae are almost always found rising straight out of the cuticle apart from a few species of the Spiruromorpha where the papillae are raised above the surface of the cuticle on stalks. The most plesiomorphic solution is that this state developed in the ancestor of the Spiruromorpha and was lost subsequently in some species. Alternatively, it may have evolved separately in *Ascarophis arctica* and some species of the Onchocercidae. It is most likely that the papillae in the primitive nematode came straight out of the cuticle providing they were present.

112: *The presence of preanal tubules*

Preanal tubules are present in most orders of the freeliving Chromadorea as well as in the Mononchidae. Because they can take on a variety of shapes they may not all be of the same evolutionary origin. They are probably absent in the primitive nematode and have developed several times.

113: *The presence of a single, trumpet-shaped preanal tubule*

A specific, trumpet-shaped preanal tubule is only present in the genera *Enoplus* and *Enoploides*. It is probably of a common evolutionary origin but was absent in the primitive nematode. It is therefore a synapomorphy for the two taxa.

114: *The presence of glands in the supplements*

Glands have been observed in the supplements of a variety of species but there is no clear pattern as to when they occur. Additionally, many supplements are very small and it is difficult to discern whether glands are present or not. Therefore, it is likely that glands have been overlooked in some species and no estimate can be given as to whether they were present in the primitive nematode.

115: *The presence of one ventral row of preanal setae*

A single row of preanal setae is found in the Chromadorida, Xyalidae, Comesomatidae and Axonolaimidae. Because they are absent in all Enoplea they were probably also absent in the primitive nematode. The most parsimonious solution suggests that they have developed in the ancestor of the Chromadorea and have been lost secondarily in the Rhabditida and other Chromadorea. More ultrastructural data are needed to determine whether the different occurrences of ventral preanal setae are indeed of the same evolutionary origin.

116: *The presence of two subventral rows of preanal setae*

Two rows of preanal setae are found in the Enoplida, Desmodoridae and Plectidae. They were also present in the ancestral nematode if the Enoplia are older than the Dorylaimia. They have been lost secondarily in the Rhabditida. It is possible that the setae in one (see #115) and two rows (this character) are of the same evolutionary origin. In this case, two subventral rows have been reduced to one ventral row at several occasions. More ultrastructural data are needed to determine the evolutionary origin of the preanal setae.

117: *The presence of two subventral rows of adanal setae*

In some species with two rows of preanal setae the setae extend to the level of the cloaca. These setae are probably of the same evolutionary origin but its range along the body has been extended. The primitive nematode may also have had setae extending alongside the cloaca.

118: *The presence of two subventral rows of postanal setae*

In most species with preanal subventral setae, postanal subventral setae are also present. A few other species from the same families and *Chromadora nudicapitata* also feature setae in the caudal region. These setae may also be of the same evolutionary origin as those in characters 116 and 117 but extending their range along the body. The primitive nematode may also have featured postanal subventral setae. More ultrastructural data are needed to determine whether the different occurrences of subventral preanal, adanal and postanal setae are of the same evolutionary origin.

119 – 120: *The presence of ventral preanal and postanal thorns instead of setae*

At least in *Monoposthia costata* and *Molgolaimus demani* strong thorn-like cuticular structures are found on the ventral surface anterior and posterior to the cloaca. These thorns may be strengthened setae or of independent origin. In any case, the ancestral nematode probably did not possess such thorns but just simple setae (see above # 115 – 118).

121: *The presence of preanal and postanal setae arranged in groups*

In *Gnathostoma turgidum* and *Goezia pelagia* the pre-, ad- and postanal papillae are arranged in groups rather than in distinct rows. They may still be of the same evolutionary origin as those in subventral rows but they have, at least in part, been displaced. The primitive nematode probably did not possess such an arrangement of setae but it developed within the Spirurina.

122 – 123: *The presence of a bursa and the type of bursa*

Most non-Rhabditida do not have a bursa and this is probably the plesiomorphic state for the phylum. The bursae of *Anoplostoma viviparum*, *Paratrichodorus pachydermus* (Seinhorst, 1954), *Campydora demonstrans* and *Diplolaimelloides meyli* are probably all of separate evolutionary origin as is the ancestral bursa of the Rhabditida. Within the Rhabditida, the simple alate bursa of the Spirurina probably developed separately from the more elaborate bursae found in the Rhabditina and Tylenchina. The most parsimonious

solution for this data set suggests that the simple alate bursa evolved several times in the Spirurina; however, it may also have evolved in the ancestor of the Spirurina and been lost secondarily in some species. If the Rhabditina and Tylenchina have a common origin their ancestor probably had an open peloderan bursa. This bursa subsequently developed into the leptoderan or simple adanal bursa of the Tylenchina and the leptoderan or the closed peloderan bursa of the Rhabditina. In this data set species of the Rhabditina clade 1 only possess a leptoderan bursa and those of the Rhabditidae clade 1 only an open peloderan bursa. The remaining groups of the Rhabditina clade 2 include species with a leptoderan bursa and species with both types of peloderan bursae. The bursa of the Strongyloidea (except *Heterorhabditis bacteriophora*) with its elaborate lobes and rays probably evolved from the peloderan bursa. Within all three suborders of the Rhabditida the bursa has also been lost secondarily in many species. It also suggests that the leptoderan bursa has at least two separate evolutionary origins.

124: *The presence of a genital cone on the bursa*

A genital cone is only found in a connection with a strongyloid bursa in *Strongylus equinus* and *Haemonchus contortus* in this data set. The cone probably developed once in the Strongyloidea and was subsequently lost in some of its species. It is unlikely the primitive nematode possessed a genital cone as it probably did not possess a strongyloid bursa either.

125: *The presence of roughened cuticular patches and pads in the anal region*

Roughened cuticular patches are found in a number of species of the Spirurina and in some freeliving nematodes. These patches are probably of separate evolutionary origin and may not be homologous amongst the different taxa. The ancestral nematode probably did not possess roughened cuticular patches in the anal region.

#126: *The presence of denticular plates posterior to the cloaca*

At least in this data set denticular plates were only found behind the cloaca in *Pseudoterranova decipiens* (Krabbe, 1878). This character probably only developed in its direct ancestor and was absent in the primitive nematode. This character is uninformative for this data set.

127: *The overall body shape*

In several species of the Tylenchomorpha the mature female becomes enlarged to take on an obese or saccate body form on account of the large number of eggs she carries. This enlargement, however, seems to have developed independently in different families of

the Tylenchomorpha and the ancestral nematode was most likely of the simple thread-like shape as most extent nematodes are.

128: *The presence of biocrystals in the body cavity*

In some species crystalline bodies are found in the body cavity. Their function is uncertain and in this data set they were only found in *Ironus dentifurcatus*. The primitive nematode probably did not possess biocrystals. This character is uninformative for this data set.

129 – 133: *The cuticular ornamentation*

The homology assessment of the cuticular ornamentation proves rather difficult. Judging by the effect on the overall topology and the CIs the binary characters were chosen for the phylogenetic reconstruction. However, mapping both, the binary character and the multi-state character, on the resulting tree showed that the individual CIs for all five characters are rather low. Many reversals have to be included to explain the data. However, using the mapping of the multi-state characters provides a good estimation of the possible evolution from one state to another state.

Most species of the Enoplea have either a smooth cuticle or the cuticle is only faintly striated. Clear striations are only found in *Enoploides brunettii*, *Prismatolaimus intermedius, Campydora demonstrans, Belondira apitica* and the Trichinelloidea in the Enoplea in this data set. Therefore, either the presence of faint striations or a smooth cuticle is the plesiomorphic state for the phylum. The faint striations became much more obvious in the Chromadorea. They developed into strong articulations in the Monoposthiidae, *Ascarophis arctica* and Cephalobidae and in many species of the Criconematoidea and Tylenchoidea. The striations became faint or were entirely lost again in many Monhysterida including the Comesomatidae as well as in some Panagrolaimomorpha, Tylenchoidea, Rhabditidae and Spirurina. Until more data on the ultrastructure and ontogenesis of the cuticle are available it is uncertain whether the different types of striations have the same or different evolutionary origins.

134: *The presence of longitudinal striations on the cuticle*

Longitudinal striations are found in many groups of nematodes and they are particularly common in the Rhabditomorpha. They occur occasionally in species of the Dorylaimia and in this data set have also been found in *Geocenamus quadrifer* (Tylenchina). They were probably absent in the ancestral nematode but this cannot be certain because the true root relationships are still unresolved.

135 – 138: *The distribution of the cuticular striations along the body*

In many species the cuticular pattern disappears either in the head or tail region. However, by mapping the character changes onto the phylogenetic trees it appears that the cuticular ornamentations have been lost in the head region many times and also quite often on the tail. Many reversals are needed to explain the loss of pattern on the head. In *Gnathostoma turgidum* the cuticular pattern is only present from the anterior end to the middle of the body length. In *Nematodirus battus* and *Nippostrongylus brasiliensis* the cuticular pattern is only present on the cervical cuticular inflations. It is uncertain whether the presence or absence of the cuticular pattern in certain body regions is homologous amongst different taxa. Especially in the case of the loss in the head region many reversals are needed to explain the data. Therefore it is also difficult to assert whether the cuticular pattern was continuous in the primitive nematode.

139: *The presence of retrorse scales and spines on the cuticle*

In *Gnathostoma turgidum* and *Goezia pelagia* retrorse scales and spines are found on the cuticle. Such ornamentation was probably absent in the primitive nematode but it is uncertain whether it evolved once or several times in the Spirurina.

140: *The presence of subcuticular longitudinal incisures*

In this data set subcuticular longitudinal incisures were only observed for *Aphelenchus avenae*. Therefore, this character is uninformative and no estimate of the plesiomorphic state can be made. However, it is more likely that subcuticular incisures were absent in the primitive nematode because this structure has not been found in any species of the Enoplea.

141: *The presence of subcuticular annulation*

Subcuticular annulations are present in species of the Dorylaimida, Campydoridae, Xyalidae, Comesomatidae, Triplonchida, Plectidae, Filarioidea, Cephalobidae, Aphelenchoidea, Criconematoidea and Rhabditidae. Because they are found only in selected species of a family or even genus (e.g. *Xiphinema* partim) but in a variety of taxa there is no clear pattern of the evolutionary history of subcuticular annulations and it is uncertain whether they were present in the ancestral nematode.

142: *The presence of cuticular punctations*

Cuticular punctations are found in selected species in all major groups of the nematodes except in the Spirurina. Only in the Chromadoridae are they commonly

observed. It is therefore difficult to estimate whether the primitive nematode also possessed cuticular punctations and whether the different occurrences of punctuations are homologous at all. However, the punctation probably forms a synapomorphy at least for the Chromadorida.

143: *The presence of subcuticular punctations*

Subcuticular punctations are found in a few nematode species from a variety of orders. Without more data or more detailed knowledge of their structure no assessment of their evolutionary origin and homology can be made.

144 – 146: *The presence of bosses on the cuticle*

Bosses are found on different parts of the body in a few vertebrate parasites (*Loa loa, Wuchereria bancrofti* and *Haemonchus contortus*). However, it is not certain whether these bosses are of the same evolutionary origin or whether they have arisen at least twice, once in the Filarioidea and once in the Trichostrongylidae. Therefore, the occurrence of bosses was coded as distinct characters, which made them uninformative for this data set. The bosses were, however, most likely absent in the primitive nematode.

147: *The presence of warts and "Kriechleisten" on the right side of the body*

Only in the Bunonematidae the right side of the animal is adorned with a complex structure of a double row of warts and so called "Kriechleisten". They probably only developed in this family and were absent in the ancestral nematode. This character is uninformative in this data set and only contributes to the branch length of *Bunonema franzi*.

148: *The presence of a bacterial coat on the cuticle*

Eubacteria or cyanobacteria were probably acquired by nematodes at least twice in the evolutionary history, once in the Rhabditidae (*Cruznema tripartitum*) and in the Desmodoridae. In the latter family they were acquired once and subsequently lost in some species or they were acquired several times, depending on the true phylogeny if this family.

149: *The presence of cervical inflations of the cuticle*

Some species of the Strongyloidea possess cervical inflations of the cuticle. These probably developed only in this group. They are either of a single origin and have been lost secondarily in some species or they have been developed several times in this group. In any case, they were probably not present in the primitive nematode.

150: *The presence of a bacillary band on the cuticle*

Only the Trichuridae possess an elaborate bacillary band on the ventral side of the body in the oesophageal region. Gibbons (1986) suggested that this band is in fact homologous to the lateral field and alae and includes the hypodermal glands (see #154). In any case, this modification of the cuticle probably only developed once and was absent in the ancestral nematode. In this data set only *Trichuris muris* possesses this feature rendering this character uninformative.

151: *The presence of V-like projections on the annules*

The Monoposthiidae possess V-like ornamentations on the strong annules of the cuticle. This character was probably absent in the primitive nematode and constitutes a synapomorphy for the Monoposthiidae.

152: *The presence of spiral thickenings of the cuticle*

Onchocerca cervicalis Railliet & Henry, 1910, possess a special type of annulation composed of spiral thickenings instead of a regular annulation. Since this character is only found once in the data set it is uninformative. It is also not possible to make an assessment of the evolutionary history of this character although it is likely that it was absent in the primitive nematode.

153: *The presence of strong incisures in the lateral field*

Especially strong incisures in the lateral field are mainly found in the Rhabditida. Here, they are absent in most Spirurina and Panagrolaimomorpha. Many reversals occurred between the two states in the Rhabditina; there are no incisures in the lateral fields of the Strongyloidea. In the remaining nematodes such a structure is rare; in this data set it is only found in the Plectidae and in *Paractinolaimus macrolaimus* and *Neochromadora* sp. Because lateral incisures are rare in the Enoplea they were probably absent in the ancestral nematode. Without more data and information of its ultrastructure it is not possible to make an assessment of its homology in the phylum.

154: *The presence of alae in the lateral field*

Lateral alae are only found in the Plectidae and Rhabditida. Within the Rhabditida they are common in the Cephalobidae, Aphelenchoidea and Tylenchoidea and in some species of the Rhabditidae and Ascaridoidea (*Hysterothylacium fortalezae* [Klein, 1973]). Lateral alae were therefore probably absent in the ancestral nematode. Within the Plectidae

plus Rhabditidae they either developed several times in the different groups or once in their common ancestor or were lost secondarily in other species.

155: *The presence of an areolated lateral field*

The lateral fields in many species of the Cephalobidae and Tylenchoidea are areolated. Such areolation either developed once in the ancestor of the Tylenchina or at least twice, once in each group. In either case, the areolation was probably lost secondarily in some species in both groups. The primitive nematode probably did not possess any areolations.

156: *The presence of modified punctation in the lateral field*

In *Enoplus anisospiculum*, *Distolabrellus veechi*, in the genus *Sabatieria* and in many species of the Chromadorida the punctation of the cuticle is modified in the lateral field. It is uncertain, however, whether this modification is homologous in the phylum or whether it is of independent evolutionary origin. More ultrastructural and ontogenetic data are needed to assess its homologies. Since punctation is rare in the Enoplea it was probable absent in the primitive nematode.

157: *The presence of a granular lateral field*

In some Dorylaimida, here in *Mesodorylaimus bastiani* and *Paractinolaimus macrolaimus*, the lateral field is not especially adorned but granular. This character probably developed once in this group and was lost subsequently in some species, depending on the true phylogeny of the Dorylaimida. The lateral field of the primitive nematode was probably not granular.

158: *The presence of hypodermal glands*

Hypodermal glands are present in most groups of nematodes apart from the Rhabditida (except *Pristionchus pacificus* in this data set). The glands may occur in different parts of the body but are probably all of the same evolutionary origin. They are probably plesiomorphic for the phylum because they have been found in most Dorylaimia and many Enoplia. It has been suggested that they are probably present in many more species than they have been recorded in. Therefore, it is not possible to assess the degree of reversals and losses found in the data.

159: *The presence of ocelli*

In this data set ocelli are only found in three distantly related species (*Syringolaimus striatocaudatus, Mermis nigrescens* and *Diplolaimelloides meyli*). Therefore, it is not

possible to determine whether they may have been present in the ancestral nematode and whether they are homologous at all.

160: *The presence of pigment spots*

Pigment spots developed at least twice in the phylum, once in the genus *Enoplus* and once in the Chromadorida. Because these two groups are only distantly related the character may have two independent evolutionary origins. In the Chromadorida, the pigment spots have become lost secondarily in many species. Without more data it is not possible to determine whether the primitive nematode also possessed pigment spots.

161: *The presence of somatic setae*

Somatic setae are entirely absent in all Dorylaimia and Rhabditida. In the remaining orders, they are also absent in the Diphtherophoroidea and Microlaimidae, in some species of the genus *Enoplus*, in some species of the Monhysterida and in *Calyptronema maxweberi*, *Acanthopharynx micans* and *Odontophora rectangular* Lorenzen, 1972. Although the extent of the setae cover may differ in different taxa the somatic setae are probably of the same evolutionary origin in all nematodes. Whether they were already present in the primitive nematodes depends on the true phylogenetic relationship between the setose Enoplia and the setae-less Dorylaimia.

162: *The presence of glandular somatic setae*

At least in the genus *Laxus* the somatic setae are accompanied by glands that secrete a sticky substance. These glands may well be modified hypodermal glands or they are of individual evolutionary origin. In any case, this modification probably developed in the ancestor of this genus (and related genera that may also have glandular setae) and the primitive nematode was lacking such glands.

163: *The presence of metanemes*

Metanemes are only present in the Enoplia. They probably developed once in the ancestor of this group and were lost secondarily in some species. Lorenzen (1981/1994) also suggested that they may have been overlooked in many species because they are difficult to detect. Depending on the relationships between the Enoplia and Dorylaimia they may have been present or absent in the ancestral nematode.

164: *The presence of phasmids*

Phasmids are present only in the Plectidae and Rhabditida. In both orders they have been lost secondarily in a few species. However, they were probably absent in the

primitive nematode and only developed in the ancestor of the Plectidae and Rhabditida for which they are a synapomorphy. In most species the phasmids are situated laterally on the tail but in selected taxa they are displaced anterior to the anus/cloaca. However, the displacement can be variable within a single species, either between sexes or ecotypes. Therefore, the evolutionary history of the precise position could not be analysed.

165: *The presence of prophasmids*

Prophasmids have only been observed in a few species of the Anguinidae. They are probably a synapomorphy for this family and were absent in the primitive nematode.

166: *The presence of postdereids*

Prodereids are only observed in the Spiruromorpha (*Setaria digitata*) and Rhabditidae (*Caenorhabditis elegans*) in this data set. Without more data it is difficult to assess whether they are of the same evolutionary origin in either taxon or whether they developed independently. In any case, they were probably absent in the primitive nematode.

167: *The presence of caudal pores*

Caudal pores are only present in the Dorylaimia. They are probably of the same evolutionary origin and have been lost secondarily in some species. Until it is known whether the Enoplia or Dorylaimia are the oldest group, it cannot be decided if they were present in the ancestral nematode.

168: *The presence of terminal caudal papillae*

Terminal caudal papillae are only found in the Spirurina and only in a few species. Until the ultrastructure of these characters is known it remains uncertain whether they are of common evolutionary origin in the different species. However, it is most likely that the primitive nematode did not possess this structure and that it developed in the ancestor of the Spirurina.

169 – 170: *The shape of the female and male tail*

In most nematodes, the tails of the male and female are very similar and the character state changes are also very similar in the tail shapes in the two sexes. The most parsimonious reconstruction suggested that the simple conical tail is the plesiomorphic state and that almost all other tail shapes are derived from this form. In some groups a succession of different forms is suggested: within the Dorylaimida a hemispherical shape may have acted as the intermediate for a long elongate tail. In the Enoploidea, the clavate

shape may have derived from the conical-filiform type and the swollen tip was probably lost in some species of the genus *Enoplus*. In the Oncholaimoidea, at least three shapes are present, the conico-cylindrical type with or without a swollen tip and a simple cylindrical tail. The cylindrical tail probably derived from the conico-cylindrical type and the clavate shape may have derived from the cylindrical tail. In the Comesomatidae, the clavate tail may also have lost the swollen tip in *Setosabatieria hilarula* (de Man, 1922). In the Spirurina a conical tail with a blunt mucron probably derived several times from either a cylindrical tail or conical tail. In the Tylenchoidea, the hemispherical tail probably derived from the cylindrical tail by reduction; this process is taken to the extreme in females of those species with a very short tail and subterminal vulva and those with a terminal vulva.

In both sexes reversals are proposed for the shapes in the Ascaridomorpha, Tylenchoidea and Rhabditidae. Surprisingly, the inclusion of the male character resulted in a lower total CI than the inclusion of the female character in the tests but the individual CIs are higher for the male character. Overall, each tail shape has probably been derived several times introducing a considerable amount of homoplasy in the data. Without more ontogenetic data the true evolutionary history of the tail shapes remains uncertain.

171: *The presence of terminal setae*

Terminal setae probably developed several times in the evolutionary history. They exist in some species of the Enoploidea, Oncholaimidae, Tripyloididae and Desmodoridae. They are very common in the Monhysterida including the Comesomatidae and in the Axonolaimidae. They may have evolved in the ancestor of the Desmodoridae plus Monhysterida and then have become lost secondarily in some species of the Desmodoridae, Monhysterida and Axonolaimidae and in the ancestor of the Plectida plus Rhabditida. If the Enoplia are more primitive than the Dorylaimia the ancestral nematode probably possessed terminal setae, otherwise it did not.

172: *The presence of sexual dimorphism in the tail shape*

Sexual dimorphism exists in at least some species of most major groups in the Nematoda. However, in some cases the tails differ in form and in other cases in relative size. The tails may also be different when the males possess a bursa. Therefore, the different changes are probably not homologous to each other and need to be divided into separate characters. The CI is also very low for this character as many reversals have to be included to account for the data.

173 – 174: *The presence of caudal setae*

Caudal setae are absent in all Dorylaimia and Rhabditida in concordance with the absence of any somatic setae. Caudal setae are present in most species of the remaining orders. They have evolved either once in the Enoplia and lost secondarily in some taxa or they have evolved several times in the Enoplia. Until it is known whether the Enoplia or Dorylaimia are more primitive, no assessment can be made of the presence of caudal setae in the ancestral nematode. If caudal setae were not present in the primitive nematode they evolved separately in the Enoplia and in the ancestor of the Chromadorea. In the Chromadorea they have been lost secondarily in those species lacking caudal setae including the ancestor of the Rhabditida. Caudal setae are present in all Plectidae which means a reversal has occurred in their ancestor if the Plectidae are part of the Rhabditida.

If there are more caudal setae on the male tail than on the female tail it is difficult to determine whether the additional setae on the male are homologous to those on the female or whether they are supplementary copulatory setae. More caudal setae are found on the male tail in a variety of species of all groups that possess caudal setae. There is no clear pattern as to when caudal setae became more numerous in the male. Therefore, it could be concluded that those additional setae are in fact copulatory setae and should be included in those respective characters.

175: *The existence of terminal flaps*

Terminal flaps on the tail are only recorded for *Trichinella spiralis*. Therefore, this character is uninformative in this data set and no estimate can be made about its evolutionary history. However, because this character is a specialised copulatory structure it is most likely that it only developed in the ancestor of the Trichinellidae and that it was absent in the primitive nematode.

176: *The presence of caudal glands*

Caudal glands are absent in all Dorylaimia except the Mononchida and in all Rhabditida. In the Triplonchida they may either be present or absent and in some species only the duct has been observed. In all other nematodes either two or three well-developed glands can be found. Therefore, these glands were probably also present in the ancestral nematode and have been lost at least twice, once in the majority of the Dorylaimia and once in the ancestor of the Rhabditida. If the Plectidae form one group with the Rhabditida a reversal must have occurred in the ancestor of the Plectidae.

177: *The number of caudal gland openings*

In the Sphaerolaimidae and in *Desmolaimus zeelandicus* the three glands open through separate ducts and in the genera *Eubostrichus* and *Laxus* the opening of the glands developed into a more complicated outlet system. These two structures are probably of independent evolutionary origin. In all other nematodes all three (or two) glands open through a single duct and this is probably also the plesiomorphic state.

178: *The type of the spinneret*

The spinneret is well set-off from the tip of the tail in a number of families, here in the Chromadorida, Desmodoridae, Monhysteridae and Plectidae and in *Syringolaimus striatocaudatus*, *Trischistoma monohystera* and *Cyartonema elegans*. Because the spinneret developed the set-off type in a variety of families but only in very few species of the Enoplia this feature has probably several evolutionary origins and was absent in the primitive nematode.

179: *The position of the caudal glands*

In most nematodes the caudal glands are restricted to the tail but in the Enoploidea, Oncholaimidae and in *Desmolaimus zeelandicus* they extend into the precaudal region, at least in the male. Since this enlarged feature appears to be rather common in the Enoplia it may also have been present in the primitive nematode if it possessed caudal glands. The extension of the caudal glands into the precaudal region in *Desmolaimus zeelandicus* is therefore either a reversal to the plesiomorphic state or of independent evolutionary origin.

180: *The mode of reproduction*

Most nematodes are dioecious (or amphimictic) and both sexes and copulation are required for the fertilisation of the eggs and the production of offspring. This is probably also the plesiomorphic state. Some species of the Rhabditidae are hermaphroditic. They are either obligate hermaphrodites (e.g. *Caenorhabditis elegans, Pristionchus pacificus* and *Rhabditis myriophila*) or they exhibit a generation change between amphimictic and hermaphroditic life stages (e.g. *Heterorhabditis bacteriophora*) or between amphimictic and pathogenic generations (e.g. *Strongyloides stercoralis*). In many nematodes, male individuals are rare and the females reproduce by pathogenesis in their absence. This strategy is particularly common in the Plectidae and Tylenchomorpha and also occurs in *Ironus dentifurcatus* and species of the Triplonchida, Dorylaimida, Mononchidae, Cephalobidae and Rhabditidae. Hermaphroditism and pathogenesis probably derived from

the amphimictic mode of reproduction several times in the evolutionary history of the Nematoda.

Table 9.2: Most likely plesiomorphic states of the morphological characters for the phylum Nematoda

2b	24b	45b	64a	81b	95a	115b	148b	164b
3b	26a	46b	65b	82a	96b	119b	149b	165b
7a	31a	47b	70b	84b	98a	120b	151a	166b
13b	38a	48b	71b	85b	99b	121b	153b	168b
18b	39a	51a	72a	86b	100b	122a	154b	176a
19b	40a	54a	73d	90a	103b	123b	155b	177a
20b	41b	57a	76a	92b	111b	127b	156b	178b
21b	42b	61b	77b	93a	112b	134b	157b	180a
22b	43a	62b	80b	94a	113b	139a	162b	

9.2 Buccal cavity morphology

The morphology of the buccal cavity is probably the most difficult and controversial character to find the true evolutionary history for. Traditionally, the Plectidae or Rhabditidae were thought to be the most primitive group of nematodes. This hypothesis was based on the morphological similarities and apparent simplicity of the buccal cavities of the two groups of nematodes. The rhabditoid buccal cavity was thought to be the most primitive one because its walls consist of clearly defined sections whilst in other groups these sections could not always be recognised easily. It was assumed that the more derived buccal cavities evolved by either reduction or amalgamation of the individual sections. This interpretation of the buccal cavity evolution has not changed since the ancestral nematode was thought to be found somewhere between the Enoplea and Chromadorea with the Rhabditidae and Plectidae being more derived groups. However, the evolutionary history of the buccal cavity of nematodes is in need of revision since the phylogeny of the Nematoda has become more stable. It can no longer be assumed that the rhabditoid-plectoid type is the most primitive one because the Plectida and Rhabditida have evolved from a chromadorid ancestor, which in turn probably evolved from an enoplid ancestor.

De Ley *et al.* (1995) produced a very valuable comparison of the buccal cavities of the Panagrolaimidae, Cephalobidae and Rhabditina using ultrastructural methods. They discovered that the buccal cavities of each of these three groups where made up of six distinct cellular sections instead of five as proposed by Steiner (1933) and revised by Goodey (1963a) for the Cephalobidae, Andrássy (1962) and Paramonov (1968) for the Panagrolaimidae, Sachs (1950) and Andrássy (1962) for the Rhabditina and Andrássy (1962) and Goodey (1963b) for the Tylenchida. De Ley *et al.* (1995) were also able to

homologise these sections between the three groups. Further ultrastructural studies on the buccal cavity of nematodes were conducted on the following species: *Caenorhabditis elegans* (Rhabditida, Wright & Thomson, 1981), *Rhabdodemania minima* (Triplonchida, Hope, 1988), *Monhystera* sp., *Gammarinema gammari*, *Theristus flevensis* and *Daptonema setosum, Pseudosteineria horrida, Rhynchonema lyngei* (Monhysterida, Tchesunov, 1990a, 1990b), *Geomonhystera disjuncta* and *Diplolaimella dievengatensis* (Monhysterida, van der Velde & Coomans, 1991), Cephalobidae (Rhabditida, van der Velde *et al.*, 1994), *Tylopharynx* sp. (Rhabditida, de Ley *et al.*, 1993), *Ceramonema carinatus* (Plectida, Stewart & Nicholas, 1994), *Heterorhabditis bacteriophora* (Rhabditida, Endo & Nickle, 1994), *Steinernema feltiae* (Rhabditida, Endo & Nickle, 1995), *Zeldia punctata* (Rhabditida, Baldwin & Edelman, 1995), *Aduncospiculum halicti* (Rhabditida, Baldwin *et al.*, 1997), *Gonionchus australis* (Monhysterida, Nicholas & Stewart, 1997), *Heth mauriesi* (Rhabditida, Spiridonov & Yushin, 2000), Diplogasterina (Rhabditida, von Lieven & Sudhaus, 2000), and *Bunonema* sp. and *Teratorhabditis palmarum* (Rhabditida, Dolinski & Baldwin, 2003). The above list shows that the majority of work into the ultrastructure of the buccal cavities of the nematodes had been carried out in the order Rhabditida. Outside the Rhabditida the ultrastructure of the buccal cavity of several species of the Monhysterida has been studied but no other freeliving Chromadorea and only two species of the Enoplea have been investigated.

The revised structure of the rhabditoid buccal cavity *sensu* de Ley *et al.* (1995) is as follows: the buccal cavity is divided into three main regions and the posterior region is subdivided into a further four sections. All boundaries are defined by the underlying cellular structures rather than the cuticular structure as in Steiner (1933). Under the light microscope the buccal cavity walls appear to be divided into transversely broken platelets, called rhabdia. However, with the aid of the electron microscope it became apparent that the only true cuticular discontinuity is found between the anterior two sections and any other divisions are more clearly defined by the cell processes. In the new definition, Steiner's (1933) cheilostom (Greek χειλος = lip) is retained. It is the anterior-most part of the buccal cavity that is lined by labial (or endo-) cuticle. The middle section is newly defined as that part of the buccal cavity that is not covered by oesophageal tissue but surrounded by two arcade epidermis rings. It is given the new name gymnostom (Greek γνμνός = naked). The posterior edge of the gymnostom corresponds to the anterior rim of the perioesophageal basal lamina layer. The third and most posterior part is now called the stegostom (Greek στεγω = to cover, to enclose) and is defined by the area coved with oesophageal tissue (also called the pharyngeal sleeve) and enveloped by the

perioesophageal basal lamina layer. The stegostom is further divided into the pro-, meso-, meta-, and telostegostom preserving Steiner's (1933) original prefixes. All divisions are defined by the underlying inter- and adradial cell processes of the pharyngeal sleeve. In terms of the cuticular lining, de Ley *et al.* (1995) suggest the use of the wording "walls of the gymnostom" etc. instead of using the old term "rhabdia".

Although de Ley *et al.*'s (1995) new terminology appears to be homologous across the Rhabditina and Tylenchina (the stomatostyle originates from the buccal cavity lining) no one has yet attempted to transfer it to the other groups of the Nematoda. Only Nicholas & Stewart (1997) have homologised the three parts of the buccal cavity of some species of the Monhysterida with the new system but they were unable to define the different interradial, adradial and marginal cell processes of the stegostom according to de Ley *et al.* (1995). However, in almost all species quoted above the three main regions, cheilo-, gymno- and stegostom, can be identified. This shows that the underlying buccal cavity structure is indeed homologous across the entire phylum of the Nematoda and that all buccal cavities are derived from one original type. Variation occurs in the precise structure, e.g. in many species the cheilostom is reduced or completely absent (e.g. *Gonionchus australis*, Nicholas & Stewart, 1997) and the lengths of the gymnostom and stegostom can vary significantly, even within closely related species (e.g. in the Diplogasterina, von Lieven & Sudhaus, 2000).

Whilst it is potentially possible to identify the three main sections for all nematodes the homology assessment for the subsections of the stegostom poses greater problems. As long as four sections can be identified, the homology assessment according to de Ley *et al.* (1995) is possible, e.g. in *Caenorhabditis elegans*, the underlying interradial cells of the anterior two sections (pro- and mesostegostom) are non-muscular, epithelial cells forming the 'pharyngeal sleeve' while the underlying interradial cells of the stegostom of other Rhaditidae, Cephalobidae, Panagrolaimidae and Monhysterida are fully muscular. However, in other groups, the number of sections of the stegostom varies; e.g. *Aduncospiculum halicti* (Baldwin *et al.*, 1997) has got two sets of epithelial and three sets of muscle cells underlying the stegostom but only four cuticular sections; the second and third sets of muscle cells are both underlying the telostegostom; the cuticular lining is homologous to other nematodes but the third set of interradial cells cannot be homologised (von Lieven & Sudhaus, 2000). The stegostom of the infective juveniles of *Steinernema feltiae* has five underlying sets of muscles (Endo & Nickle, 1995). The anterior end of the oesophageal tissue of the stegostom of the Monhysterida is often preceded by an electron

dense layer (van der Velde & Coomans, 1991). It is uncertain whether it is an independent structure or part of the stegostom (van der Velde *et al.* 1994). The stegostom of *Heth mauriesi* is so small, no subdivisions were given (Spiridonov & Yushin, 2000).

Despite the ability to homologise most of the underlying tissue types of all buccal cavity types it remains uncertain how the ancestral buccal cavity was structured. More ultrastructural data are needed to determine the structure of the buccal cavity of the Enoplea and to homologise the finer structure of the stegostom across the whole phylum.

9.3 The relationship between morphological and molecular evolution in nematodes

Two clades in the revised phylogeny (Figures 8.1, 9.1 to 9.7) are not supported by any of the 180 morphological characters: the Enoplida and Strongyloidea. The Enoplida *sensu* de Ley & Blaxter (2002) correspond to Lorenzen's 'Enoplina' but with an extended content. According to Lorenzen (1981/1994) the synapomorphy for the 'Enoplina' is the extension of the caudal glands into the precaudal region. However, in the revised phylogeny the Enoplida include more taxa all of which retain the caudal glands in the tail. Therefore, Lorenzen's synapomorphy cannot be retained and no new synapomorphies are found amongst the current characters. The author is not aware of any synapomorphies combining the Heterorhabditidae with the remaining Strongyloidea and the current data also does not reveal one either. It is suggested that the connection is based entirely on molecular data (D. Gibson, pers. comm.).

On the other hand, four synapomorphies were found for the novel clade of *Syringolaimus striatocaudatus* and *Campydora demonstrans*: both taxa possess a bulb in the postcorpus which is otherwise unusual for the Enoplida (# 44), both have only one anterior testis while most other Enoplida posses two opposed testes (# 90), both species possess metanemes (# 163) which Lorenzen (1981/1994) interpreted as a synapomorphy of the 'Enoplina', and their caudal glands are restricted to the tail region (# 179) which Lorenzen (1981/1994) described as an apomorphy of certain species of the 'Enoplida' but which is more common in the extended definition of the Enoplida (see above). The posterior testis may have become reduced independently and the presence of metanemes and the position of the caudal glands might be the plesiomorphic state for the Enoplia and Nematoda respectively. However, the presence of a muscular bulb in the postcorpus is unique and provides strong support for their sister relationship.

Two synapomorphies were found for the clade of Mermithida and Mononchida: the presence of two subventral rows of preanal (# 105) and postanal (# 109) copulatory

papillae. The first character is also found in many Dorylaimida and may be a symplesiomorphy for a larger clade but this can only be determined when the true relationships within the Dorylaimia is established. The second character is unique in the Dorylaimia for these groups and therefore provides strong support for their sister relationship.

Five synapomorphies were found for the Monhysterida in its new extent including the Comesomatidae but two of those, amphid shape (# 14) and postcorpus structure (# 44), find further changes in the evolution of the Comesomatidae. The other three characters, the presence of outstretched ovaries (# 77), a gubernaculum with caudal apophyses (# 100) and a spinneret that is continuous with the tail (# 178) did not undergo further evolution in the Comesomatidae. The first two character states only occur sporadically outside the Monhysterida and are therefore probably true synapomorphies for this clade. However, more often than not in the Nematoda the spinneret is continuous with the tail. As stated above, the offset nature of the spinneret is probably an apomorphy for those taxa with this state and the state for most Monhysterida is most likely to be a plesiomorphy for the Nematoda. Therefore, two morphological characters support the position of the Comesomatidae within the Monhysterida.

Three synapomorphies were found for the robust position of *Cyartonema elegans* as sister taxon of *Terschellingia longicaudata*: the existence of a bulb in the postcorpus (# 44) and the loss of all somatic and terminal setae (# 161, 171). However, all three character states are also found in related species and therefore these states are not exclusive for the two species but represent potential synapomorphies for a more comprehensive clade. Therefore, this sister relationship remains most strongly supported by the molecular characters alone.

The Setariidae are supported as part of the Filarioidea by three synapomorphies: a glandular metacorpus (# 44) and postcorpus (# 43) and the vulva in the oesophageal region (# 80). A glandular metacorpus and a vulva in the oesophageal region occur rarely outside the Filarioidea. These two characters therefore provide strong support for this clade. A glandular postcorpus in found in a variety of different clades but as discussed above, the different occurrences may actually be of independent evolutionary origin. Therefore, all three synapomorphies are probably homologies providing strong support for the position of *Setaria digitata* amongst the Filarioidea.

Even though the revised phylogeny is mainly based on molecular characters, the above analysis shows that it is also supported by many morphological characters. All but

two clades, the Enoplida and Strongyloidea, are supported by at least one synapomorphy using the current 180 characters. Apart from classic monophyletic taxa, morphological characters also provide evidence for five newly established clades. Whilst this study shows the support of some morphological characters for the molecular phylogeny, the molecular phylogeny itself provides a template for the analysis of the morphological evolution. As this study has shown, homology assessment for morphological characters is very difficult and if an alternative data set can assist in this process it can provide valuable information. Providing the molecular phylogeny is correct – or at least close to the true phylogeny – its ability to provide an insight into morphological evolution (Baldwin *et al.*, 1997; Fitch, 1997; Nadler & Hudspeth, 1998; Dorris *et al.*, 1999) is so important particularly because of the lack of a fossil record in nematodes.

Legend to Figures 9.1 to 9.7:

- ■ Synapomorphy
- ▨ Synapomorphy but clade also includes reversals
- □ Apomorphy for selected taxa in clade

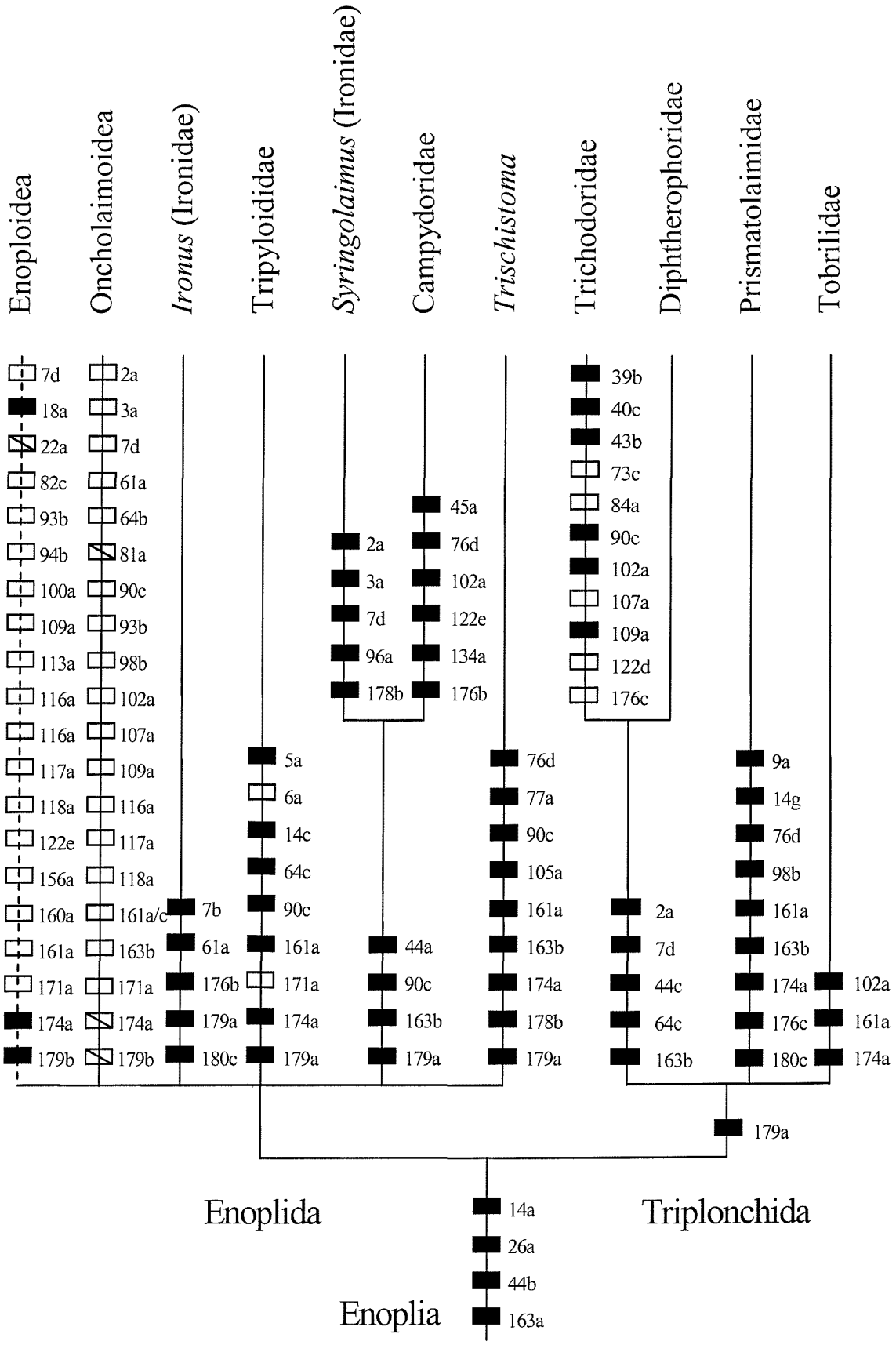


Figure 9.1: Character changes in the Enoplia

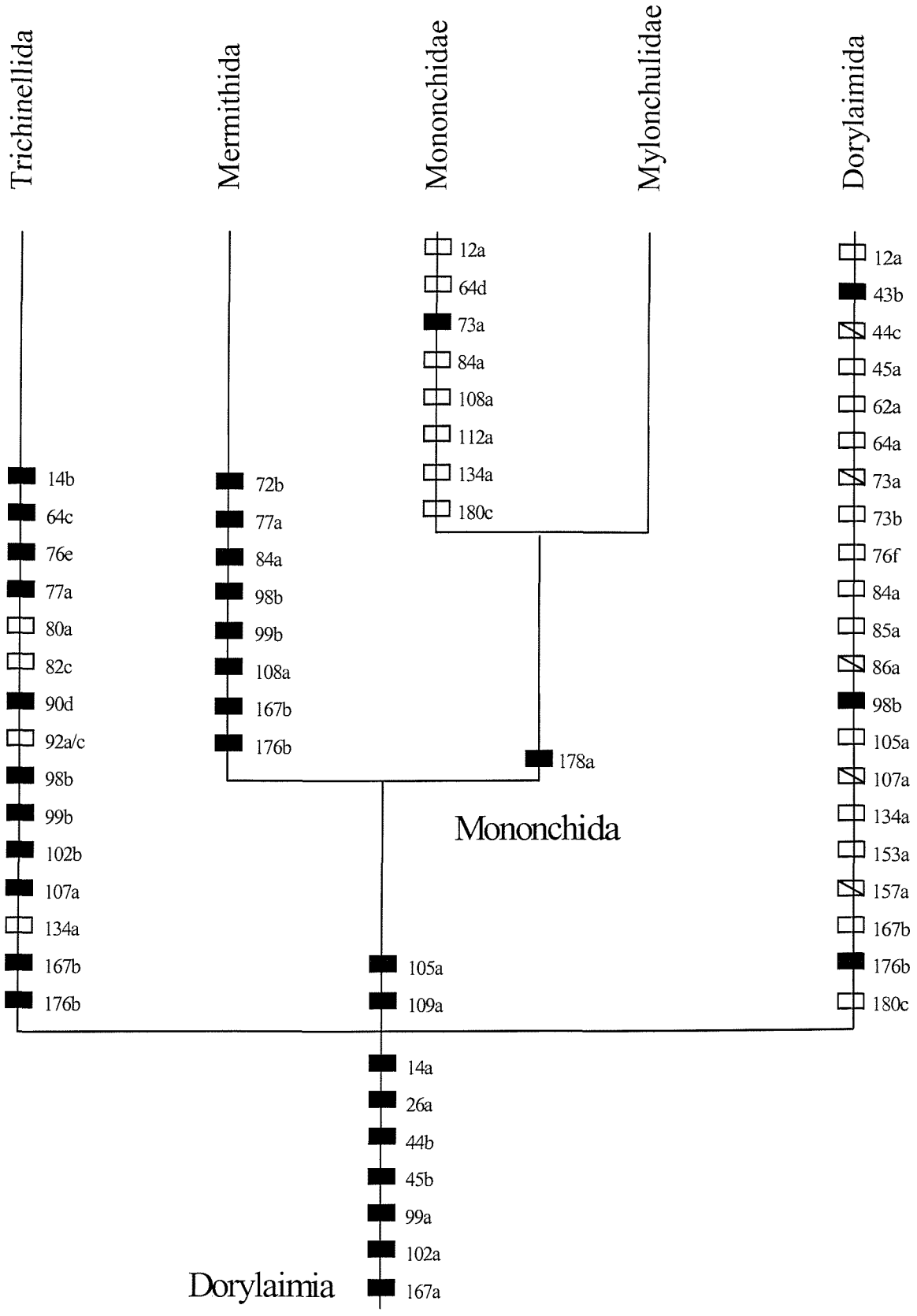


Figure 9.2: Character changes in the Dorylaimia

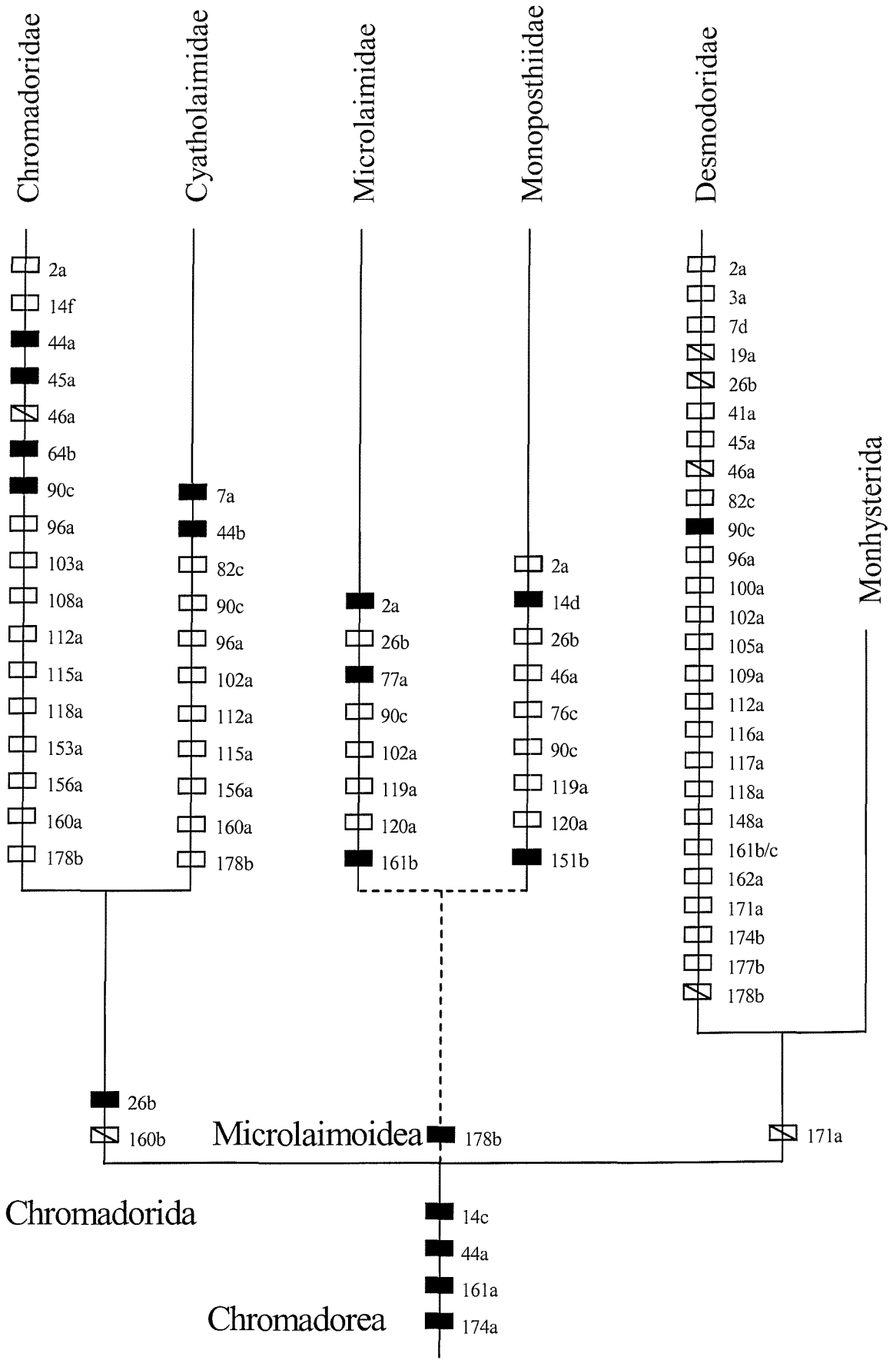


Figure 9.3: Character changes in the Chromadorea

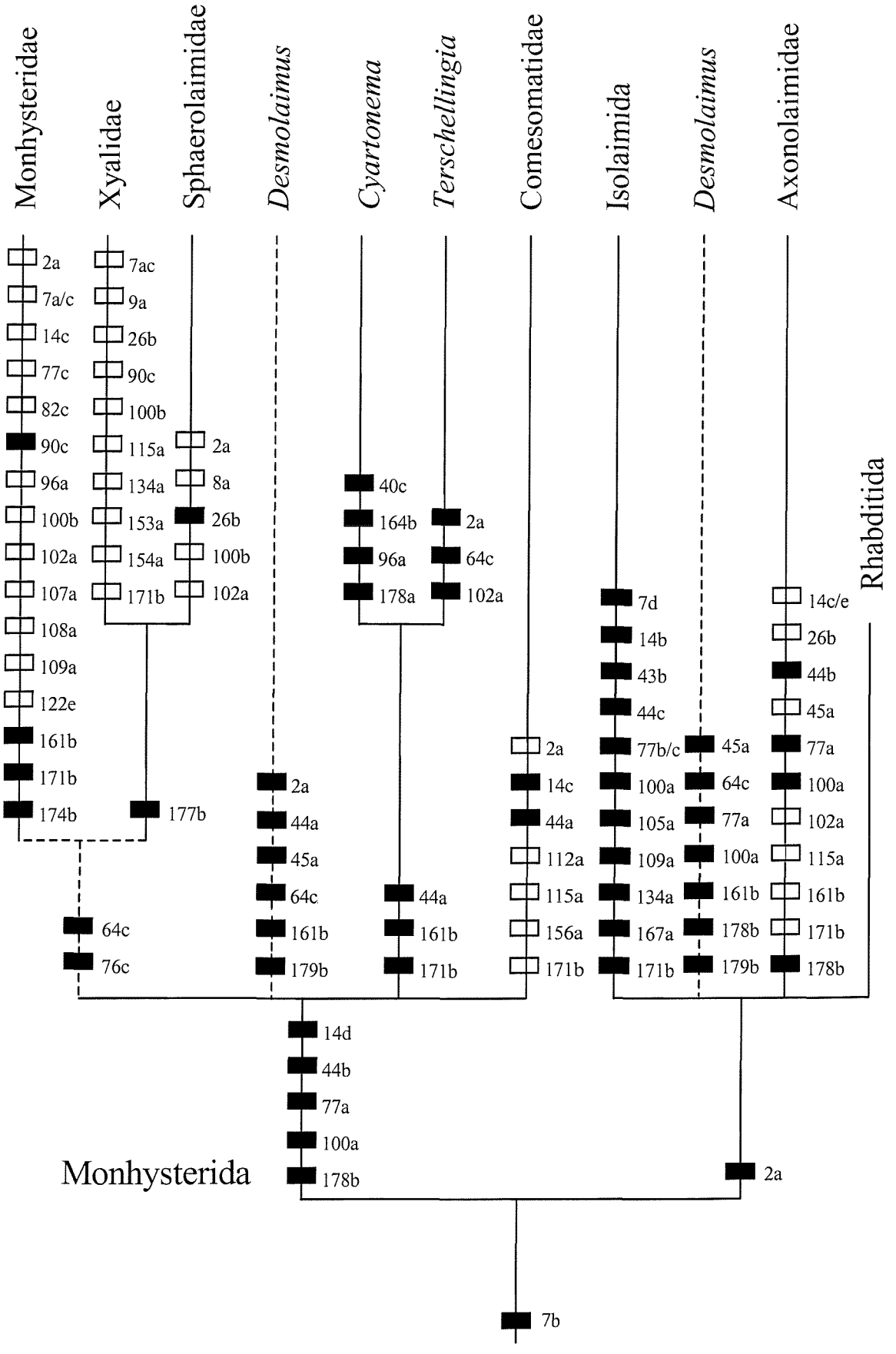


Figure 9.4: Character changes in the Monhysterida and Axonolaimidae

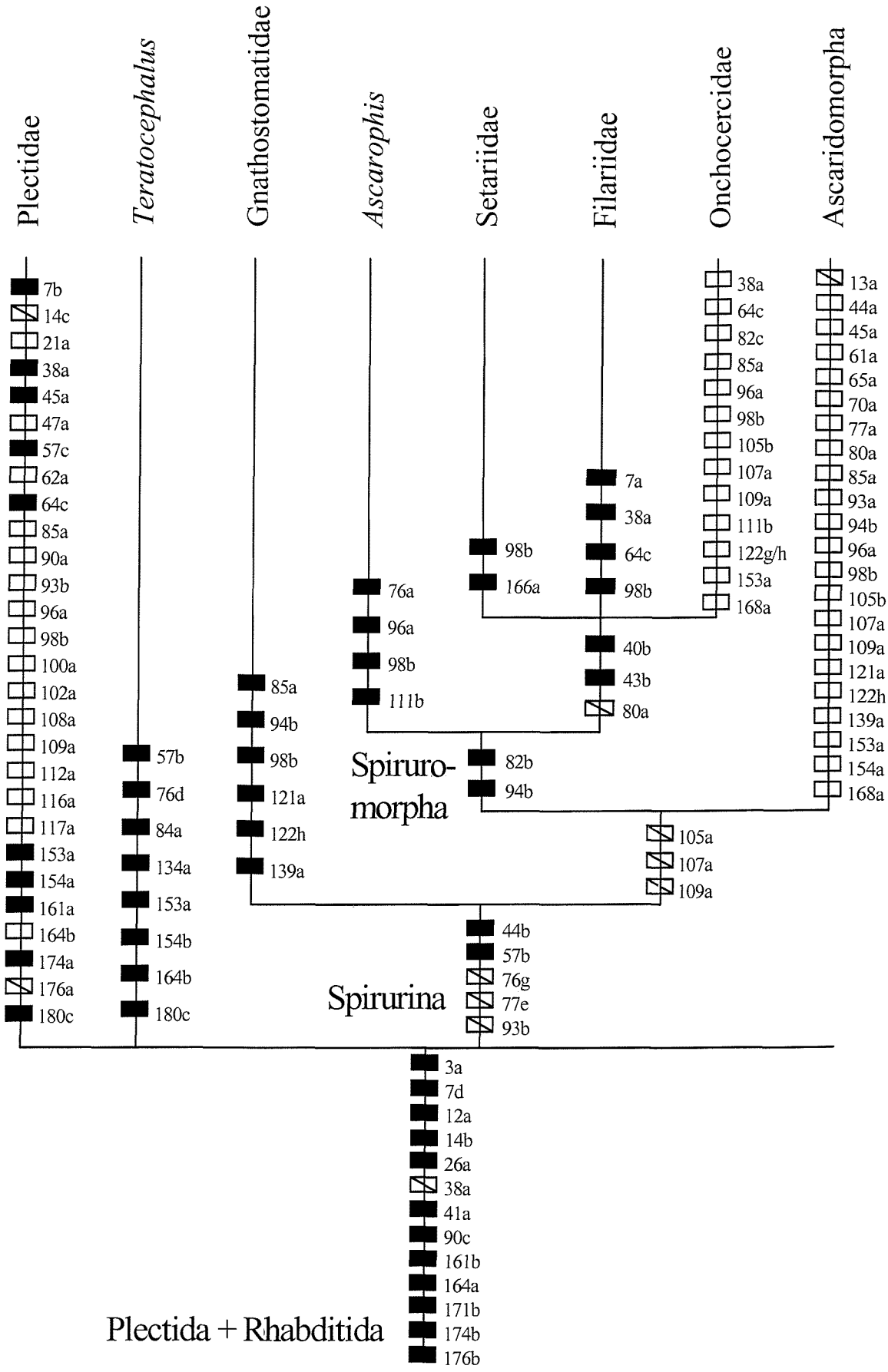


Figure 9.5: Character changes in the Plectidae and Spirurina

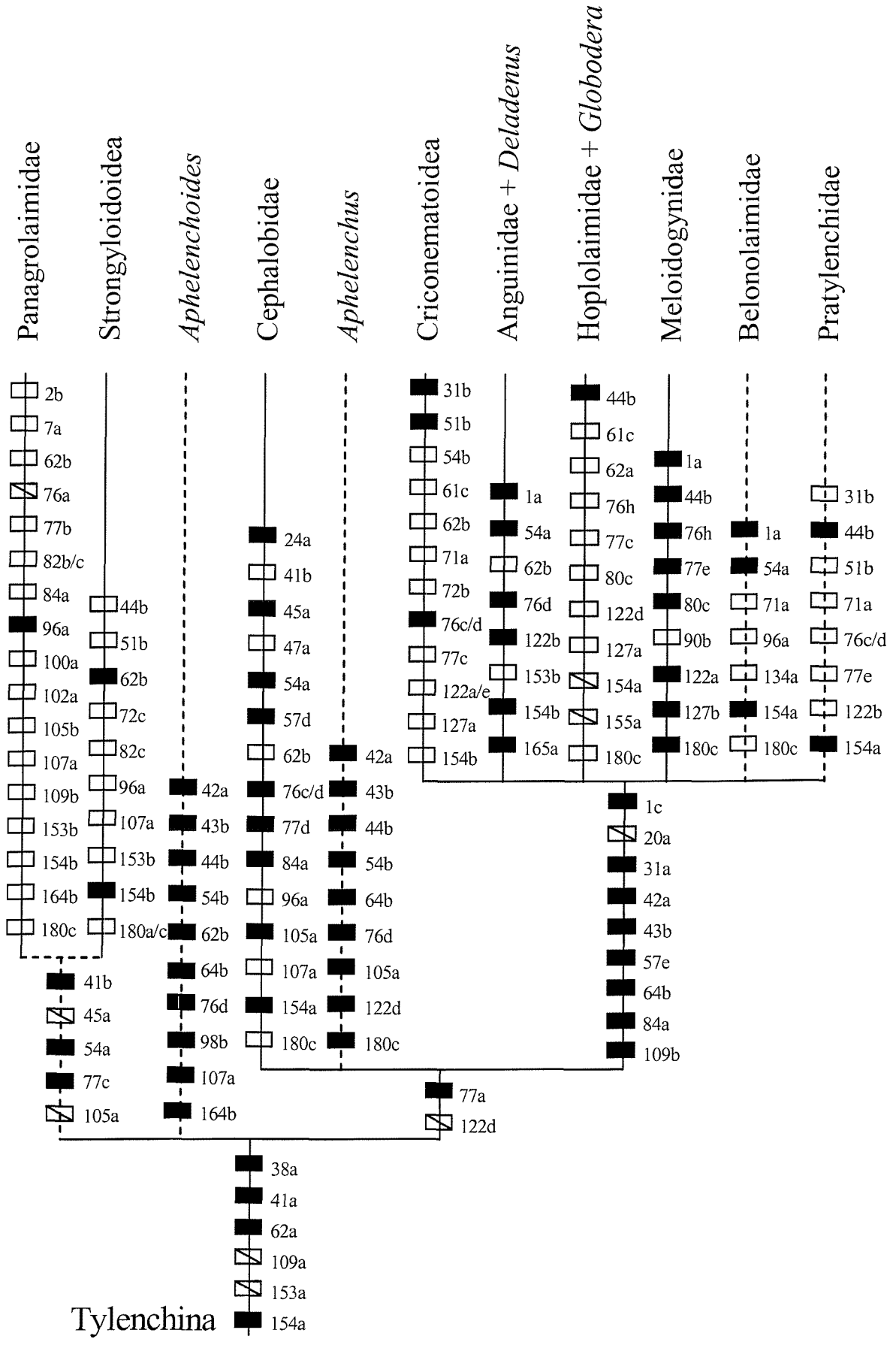


Figure 9.6: Character changes in the Tylenchina

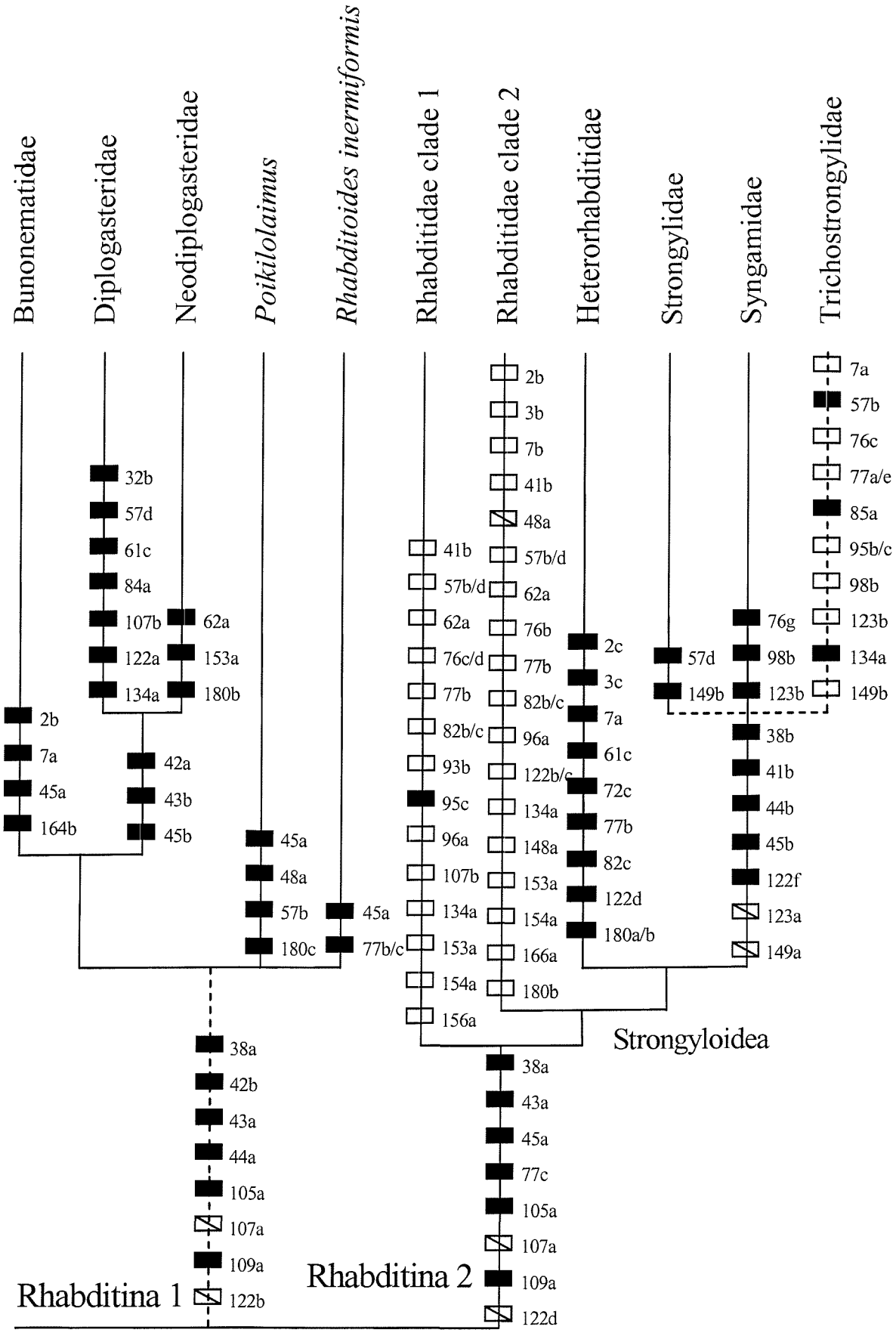
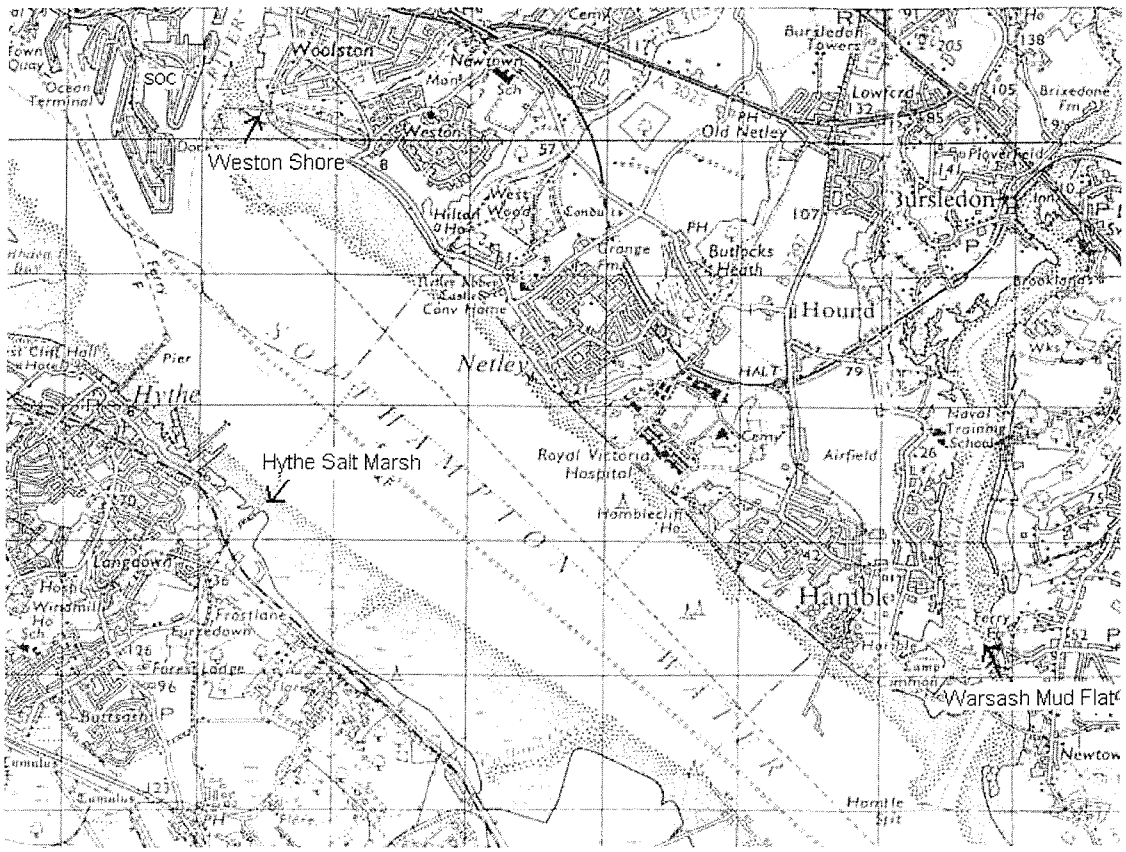


Figure 9.7: Character change in the Rhabditina

Appendix I

Figure A1: Map of sampling locations in Southampton Water (UK).



Appendix II

Table A2.1: GENBANK accession numbers and Principal Investigators for each sequence origin for all species used in this study (on CD)

Table A2.2: Inclusion of sequences/species in the outgroup analysis (OG), preliminary molecular analysis (Prelim), final molecular analyses (Molecular), morphological analysis (Morphology) and combined analyses (Combined) (on CD)

Table A2.3: Classification of species used in this study according to de Ley & Blaxter (2002)

Species Name	Family	Superfamily	Infraorder	Suborder	Order	Subclass	Class
Enoplus anisospiculus	Enoplidae	Enoploidea		Enoplina	Enoplida	Enoplia	Enoplea
Enoplus brevis	Enoplidae	Enoploidea		Enoplina	Enoplida	Enoplia	Enoplea
Enoplus communis	Enoplidae	Enoploidea		Enoplina	Enoplida	Enoplia	Enoplea
Enoplus meridionalis	Enoplidae	Enoploidea		Enoplina	Enoplida	Enoplia	Enoplea
Enoploides brunettii	Thoracomopsidae	Enoploidea		Enoplina	Enoplida	Enoplia	Enoplea
Anoplostoma viviparum	Anoplostomatidae	Enoploidea		Enoplina	Enoplida	Enoplia	Enoplea
Adoncholaimus fuscus	Oncholaimidae	Oncholaimoidea		Oncholaimina	Enoplida	Enoplia	Enoplea
Oncholaimus sp.	Oncholaimidae	Oncholaimoidea		Oncholaimina	Enoplida	Enoplia	Enoplea
Pontonema vulgare	Oncholaimidae	Oncholaimoidea		Oncholaimina	Enoplida	Enoplia	Enoplea
Viscosia viscosa	Oncholaimidae	Oncholaimoidea		Oncholaimina	Enoplida	Enoplia	Enoplea
Calyptronema maxweberi	Enschelidiidae	Oncholaimoidea		Oncholaimina	Enoplida	Enoplia	Enoplea
Ironus dentifurcatus	Ironidae	Ironoidea		Ironina	Enoplida	Enoplia	Enoplea
Syringolaimus striatocaudatus	Ironidae	Ironoidea		Ironina	Enoplida	Enoplia	Enoplea
Bathylaimus assimilis	Tripyloididae	Tripyloidoidea		Tripyloidina	Enoplida	Enoplia	Enoplea
Tripyloides gracilis	Tripyloididae	Tripyloidoidea		Tripyloidina	Enoplida	Enoplia	Enoplea
Alaimus sp.	Alaimidae	Alaimoidea		Alaimina	Enoplida	Enoplia	Enoplea
Tylolaimophorus minor	Diphtherophoridae	Diphtherophoroidea		Diphtherophorina	Triplonchida	Enoplia	Enoplea
Paratrichodorus anemones	Trichodoridae	Diphtherophoroidea		Diphtherophorina	Triplonchida	Enoplia	Enoplea
Paratrichodorus pachydermus	Trichodoridae	Diphtherophoroidea		Diphtherophorina	Triplonchida	Enoplia	Enoplea
Trichodorus primitivus	Trichodoridae	Diphtherophoroidea		Diphtherophorina	Triplonchida	Enoplia	Enoplea
Tobrilus gracilis	Tobrilidae	Tobriloidea		Tobrilina	Triplonchida	Enoplia	Enoplea
Prismatolaimus intermedius	Prismatolaimidae	Prismatolaimoidea		Tobrilina	Triplonchida	Enoplia	Enoplea
Tripyla sp.	Tripylidae	Tripyloidea		Tripylina	Triplonchida	Enoplia	Enoplea
Trischistoma monohystera	Tripylidae	Tripyloidea		Tripylina	Triplonchida	Enoplia	Enoplea
Mesodorylaimus bastiani	Dorylaimidae	Dorylaimoidea		Dorylaimina	Dorylaimida	Dorylaimia	Enoplea
Mesodorylaimus japonicus	Dorylaimidae	Dorylaimoidea		Dorylaimina	Dorylaimida	Dorylaimia	Enoplea
Mesodorylaimus nigritulus	Dorylaimidae	Dorylaimoidea		Dorylaimina	Dorylaimida	Dorylaimia	Enoplea
Aporcelaimellus obtusicaudatus	Aporcelaimidae	Dorylaimoidea		Dorylaimina	Dorylaimida	Dorylaimia	Enoplea
Allodorylaimus sp.	Qudsianematidae	Dorylaimoidea		Dorylaimina	Dorylaimida	Dorylaimia	Enoplea
Discolaimium gracile	Qudsianematidae	Dorylaimoidea		Dorylaimina	Dorylaimida	Dorylaimia	Enoplea
Eudorylaimus carteri	Qudsianematidae	Dorylaimoidea		Dorylaimina	Dorylaimida	Dorylaimia	Enoplea
Microdorylaimus sp.	Qudsianematidae	Dorylaimoidea		Dorylaimina	Dorylaimida	Dorylaimia	Enoplea

Species Name	Family	Superfamily	Infraorder	Suborder	Order	Subclass	Class
Pungentus sp.	Nordiidae	Dorylaimoidea		Dorylaimina	Dorylaimida	Dorylaimia	Enoplea
Longidorus breviannulatus	Longidoridae	Dorylaimoidea		Dorylaimina	Dorylaimida	Dorylaimia	Enoplea
Longidorus elongatus	Longidoridae	Dorylaimoidea		Dorylaimina	Dorylaimida	Dorylaimia	Enoplea
Paractinolaimus macrolaimus	Actinolaimidae	Dorylaimoidea		Dorylaimina	Dorylaimida	Dorylaimia	Enoplea
Belondira apitica	Belondiridae	Belondiroidea		Dorylaimina	Dorylaimida	Dorylaimia	Enoplea
Leptonchus microdens	Leptonchidae	Tylencholaimoidea		Dorylaimina	Dorylaimida	Dorylaimia	Enoplea
Tylencholaimus sp.	Tylencholaimidae	Tylencholaimoidea		Dorylaimina	Dorylaimida	Dorylaimia	Enoplea
Xiphinema dentatum	Xiphinematidae				Dorylaimida	Dorylaimia	Enoplea
Xiphinema flagellicaudatum	Xiphinematidae				Dorylaimida	Dorylaimia	Enoplea
Xiphinema rivesi	Xiphinematidae				Dorylaimida	Dorylaimia	Enoplea
Xiphinema theresiae	Xiphinematidae				Dorylaimida	Dorylaimia	Enoplea
Campydora demonstrans	Campydoridae	Campydoroidea		Campydorina	Dorylaimida	Dorylaimia	Enoplea
Anatonchus sp.	Anatonchidae	Anatonchoidea		Mononchina	Mononchida	Dorylaimia	Enoplea
Clarkus sp.	Mononchidae	Mononchoidea		Mononchina	Mononchida	Dorylaimia	Enoplea
Mononchus truncatus	Mononchidae	Mononchoidea		Mononchina	Mononchida	Dorylaimia	Enoplea
Prionchulus muscorum	Mononchidae	Mononchoidea		Mononchina	Mononchida	Dorylaimia	Enoplea
Mylonchulus arenicolus	Mylonchulidae	Mononchoidea		Mononchina	Mononchida	Dorylaimia	Enoplea
Isolaimium sp.	Isolaimiidae	Isolaimoidea			Isolaimida	Dorylaimia	Enoplea
Mermis nigrescens	Mermithidae				Mermithida	Dorylaimia	Enoplea
Trichinella spiralis	Trichinellidae	Trichinelloidea			Trichinellida	Dorylaimia	Enoplea
Trichuris muris	Trichuridae	Trichinelloidea			Trichinellida	Dorylaimia	Enoplea
Cyartonema elegans	Cyartonematidae	Desmoscolecoidea		Desmoscolecina	Desmoscolecida	Chromadoria	Chromadorea
Atrochromadora microlaima	Chromadoridae	Chromadoroidea		Chromadorina	Chromadorida	Chromadoria	Chromadorea
Chromadora nudicapitata	Chromadoridae	Chromadoroidea		Chromadorina	Chromadorida	Chromadoria	Chromadorea
Chromadora sp.	Chromadoridae	Chromadoroidea		Chromadorina	Chromadorida	Chromadoria	Chromadorea
Chromadorina germanica	Chromadoridae	Chromadoroidea		Chromadorina	Chromadorida	Chromadoria	Chromadorea
Chromadorita tentabunda	Chromadoridae	Chromadoroidea		Chromadorina	Chromadorida	Chromadoria	Chromadorea
Dichromadora geophila	Chromadoridae	Chromadoroidea		Chromadorina	Chromadorida	Chromadoria	Chromadorea
Neochromadora izhorica	Chromadoridae	Chromadoroidea		Chromadorina	Chromadorida	Chromadoria	Chromadorea
Neochromadora poecilosoma	Chromadoridae	Chromadoroidea		Chromadorina	Chromadorida	Chromadoria	Chromadorea
Spilophorella paradoxa	Chromadoridae	Chromadoroidea		Chromadorina	Chromadorida	Chromadoria	Chromadorea
Paracanthonchus caecus	Cyatholaimidae	Chromadoroidea		Chromadorina	Chromadorida	Chromadoria	Chromadorea
Paracyatholaimus intermedius	Cyatholaimidae	Chromadoroidea		Chromadorina	Chromadorida	Chromadoria	Chromadorea
Praeacanthonchus punctatus	Cyatholaimidae	Chromadoroidea		Chromadorina	Chromadorida	Chromadoria	Chromadorea

Species Name	Family	Superfamily	Infraorder	Suborder	Order	Subclass	Class
Acanthopharynx micans	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Catanema sp.	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Chromadoropsis vivipara	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Desmodora communis	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Desmodora ovigera	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Eubostrichus dianeae	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Eubostrichus parasitiferus	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Eubostrichus topiarus	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Laxus cosmopolitus	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Laxus oneistus	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Leptonemella sp.	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Metachromadora remanei	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Robbea hypermnestra	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Spirinia parasitifera	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Stilbonema majum	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Xyzzors sp.	Desmodoridae	Desmodoroidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Calomicrolaimus parahonestus	Microlaimidae	Microlaimoidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Molgolaimus demani	Microlaimidae	Microlaimoidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Monoposthia costata	Monoposthiidae	Microlaimoidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Nudora bipapillata	Monoposthiidae	Microlaimoidea		Desmodorina	Desmodorida	Chromadoria	Chromadorea
Diplolaimella stagnosa	Monhysteridae	Monhysteroidea		Monhysterina	Monhysterida	Chromadoria	Chromadorea
Diplolaimelloides meyli	Monhysteridae	Monhysteroidea		Monhysterina	Monhysterida	Chromadoria	Chromadorea
Geomonhystera disjuncta	Monhysteridae	Monhysteroidea		Monhysterina	Monhysterida	Chromadoria	Chromadorea
Tridentulus sp.	Monhysteridae	Monhysteroidea		Monhysterina	Monhysterida	Chromadoria	Chromadorea
Daptonema hirsuitum	Xyalidae	Sphaerolaimoidea		Monhysterina	Monhysterida	Chromadoria	Chromadorea
Daptonema normandicum	Xyalidae	Sphaerolaimoidea		Monhysterina	Monhysterida	Chromadoria	Chromadorea
Daptonema oxycerca	Xyalidae	Sphaerolaimoidea		Monhysterina	Monhysterida	Chromadoria	Chromadorea
Daptonema procerus	Xyalidae	Sphaerolaimoidea		Monhysterina	Monhysterida	Chromadoria	Chromadorea
Daptonema setosum	Xyalidae	Sphaerolaimoidea		Monhysterina	Monhysterida	Chromadoria	Chromadorea
Metadesmolaimus sp.	Xyalidae	Sphaerolaimoidea		Monhysterina	Monhysterida	Chromadoria	Chromadorea
Theristus acer	Xyalidae	Sphaerolaimoidea		Monhysterina	Monhysterida	Chromadoria	Chromadorea
Sphaerolaimus gracilis	Sphaerolaimidae	Sphaerolaimoidea		Monhysterina	Monhysterida	Chromadoria	Chromadorea
Sphaerolaimus hirsutus	Sphaerolaimidae	Sphaerolaimoidea		Monhysterina	Monhysterida	Chromadoria	Chromadorea
Desmolaimus zeelandicus	Linhomoeidae	Siphonolaimoidea		Linhomoeina	Monhysterida	Chromadoria	Chromadorea

Species Name	Family	Superfamily	Infraorder	Suborder	Order	Subclass	Class
Terschellingia longicaudata	Linhomoeidae	Siphonolaimoidea		Linhomoeina	Monhysterida	Chromadoria	Chromadorea
Ascolaimus elongatus	Axonolaimidae	Axonolaimoidea			Araeolaimida	Chromadoria	Chromadorea
Axonolaimus helgolandicus	Axonolaimidae	Axonolaimoidea			Araeolaimida	Chromadoria	Chromadorea
Axonolaimus spinosus	Axonolaimidae	Axonolaimoidea			Araeolaimida	Chromadoria	Chromadorea
Cylindrolaimus sp.	Axonolaimidae	Axonolaimoidea			Araeolaimida	Chromadoria	Chromadorea
Odontophora rectangula	Axonolaimidae	Axonolaimoidea			Araeolaimida	Chromadoria	Chromadorea
Sabatieria celtica	Comesomatidae	Axonolaimoidea			Araeolaimida	Chromadoria	Chromadorea
Sabatieria punctata	Comesomatidae	Axonolaimoidea			Araeolaimida	Chromadoria	Chromadorea
Setosabatieria hilarula	Comesomatidae	Axonolaimoidea			Araeolaimida	Chromadoria	Chromadorea
Anaplectus sp.	Plectidae	Plectoidea			Plectida	Chromadoria	Chromadorea
Chiloplectus sp.	Plectidae	Plectoidea			Plectida	Chromadoria	Chromadorea
Plectus acuminatus	Plectidae	Plectoidea			Plectida	Chromadoria	Chromadorea
Plectus aquatilis	Plectidae	Plectoidea			Plectida	Chromadoria	Chromadorea
Plectus minimus	Plectidae	Plectoidea			Plectida	Chromadoria	Chromadorea
Tylocephalus auriculatus	Plectidae	Plectoidea			Plectida	Chromadoria	Chromadorea
Wilsonema schuurmansstekhoveni	Plectidae	Plectoidea			Plectida	Chromadoria	Chromadorea
Teratocephalus lirellus	Teratocephalidae			Incertae sedis	Rhabditida	Chromadoria	Chromadorea
Brevibucca sp.	Brevibuccidae			Incertae sedis	Rhabditida	Chromadoria	Chromadorea
Anguillicola crassa	Anguillicoldiae		Incertae sedis	Spirurina	Rhabditida	Chromadoria	Chromadorea
Philonema sp.	Philometridae	Dracunculoidea	Incertae sedis	Spirurina	Rhabditida	Chromadoria	Chromadorea
Gnathostoma binucleatum	Gnathostomatidae	Gnathostomatoidea	Gnathostomatomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Gnathostoma neoprocyonis	Gnathostomatidae	Gnathostomatoidea	Gnathostomatomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Gnathostoma turgidum	Gnathostomatidae	Gnathostomatoidea	Gnathostomatomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Dentostomella sp.	Heteroxynematidae	Oxyuroidea	Oxyuridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Brumptaemilius justini	Carnoyidae	Ransomnematoidea	Rhigonematomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Ascarophis arctica	Cystidicolidae	Habronematoidea	Spiruromorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Loa loa	Filariidae	Filarioidea	Spiruromorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Acanthocheilonema viteae	Onchocercidae	Filarioidea	Spiruromorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Brugia malayi	Onchocercidae	Filarioidea	Spiruromorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Dirofilaria immitis	Onchocercidae	Filarioidea	Spiruromorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Litomosoides sigmodontis	Onchocercidae	Filarioidea	Spiruromorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Onchocerca cervicalis	Onchocercidae	Filarioidea	Spiruromorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Onchocerca gibsoni	Onchocercidae	Filarioidea	Spiruromorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Onchocerca volvulus	Onchocercidae	Filarioidea	Spiruromorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea

Species Name	Family	Superfamily	Infraorder	Suborder	Order	Subclass	Class
Wuchereria bancrofti	Onchocercidae	Filarioidea	Spiruromorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Heterocheilus tunicatus	Heterocheilidae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Ascaris lumbricoides	Ascarididae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Ascaris suum	Ascarididae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Baylisascaris procyonis	Ascarididae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Baylisascaris transfuga	Ascarididae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Parascaris equorum	Ascarididae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Porrocaecum depressum	Ascarididae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Toxascaris leonina	Ascarididae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Toxocara canis	Ascarididae or Toxocaridae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Goezia pelagia	Raphidascarididae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Hysterothylacium fortalezae	Raphidascarididae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Hysterothylacium pelagicum	Raphidascarididae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Hysterothylacium reliquens	Raphidascarididae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Iheringascaris inquies	Raphidascarididae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Anisakis sp.	Anisakidae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Contracaecum multipapillatum	Anisakidae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Pseudoterranova decipiens	Anisakidae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Terranova caballeroi	Anisakidae	Ascaridoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Heterakis sp.	Heterakidae	Heterakoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Paraspidodera sp.	Aspidoderidae	Heterakoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Cruzia americana	Kathlaniidae	Cosmocercoidea	Ascaridomorpha	Spirurina	Rhabditida	Chromadoria	Chromadorea
Setaria digitata	Setariidae			Spirurina?	Rhabditida	Chromadoria	Chromadorea
Myolaimus sp.	Myolaimidae	Myolaimoidea		Myolaimina	Rhabditida	Chromadoria	Chromadorea
Halicephalobus gingivalis	Panagrolaimidae	Panagrolaimoidea	Panagrolaimomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Panagrellus redivivus	Panagrolaimidae	Panagrolaimoidea	Panagrolaimomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Panagrobelus stammeri	Panagrolaimidae	Panagrolaimoidea	Panagrolaimomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Panagrolaimus sp.	Panagrolaimidae	Panagrolaimoidea	Panagrolaimomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Plectonchus sp.	Panagrolaimidae	Panagrolaimoidea	Panagrolaimomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Turbatrix aceti	Panagrolaimidae	Panagrolaimoidea	Panagrolaimomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Steinernema carpocapsae	Steinernematidae	Strongyloidoidea	Panagrolaimomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Strongyloides ratti	Strongyloididae	Strongyloidoidea	Panagrolaimomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Strongyloides stercoralis	Strongyloididae	Strongyloidoidea	Panagrolaimomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Acrobeles ciliatus	Cephalobidae	Cephaloboidea	Cephalobomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea

Species Name	Family	Superfamily	Infraorder	Suborder	Order	Subclass	Class
Acrobeles complexus	Cephalobidae	Cephaloboidea	Cephalobomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Acrobeloides bodenheimeri	Cephalobidae	Cephaloboidea	Cephalobomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Acrobeloides nanus	Cephalobidae	Cephaloboidea	Cephalobomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Cephalobus cubaensis	Cephalobidae	Cephaloboidea	Cephalobomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Cephalobus oryzae	Cephalobidae	Cephaloboidea	Cephalobomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Cervidellus alutus	Cephalobidae	Cephaloboidea	Cephalobomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Pseudacrobeles variabilis	Cephalobidae	Cephaloboidea	Cephalobomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Seleborca complexa	Cephalobidae	Cephaloboidea	Cephalobomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Triligulla aluta	Cephalobidae	Cephaloboidea	Cephalobomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Zeldia punctata	Cephalobidae	Cephaloboidea	Cephalobomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Aphelenchus avenae	Aphelenchidae	Aphelenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Aphelenchoides fragariae	Aphelenchoididae	Aphelenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Bursaphelenchus sp.	Aphelenchoididae	Aphelenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Criconema sp.	Criconematidae	Criconematoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Hemicycliophora conida	Hemicycliophoridae	Criconematoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Tylenchulus semipenetrans	Tylenchulidae	Criconematoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Paratylenchus dianthus	Incertae sedis	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Ditylenchus angustus	Anguinidae	Sphaerularioidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Subanguina radicicola	Anguinidae	Sphaerularioidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Globodera pallida	Heteroderidae or Hoplolaimidae?	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Helicotylenchus dihystera	Hoplolaimidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Rotylenchus robustus	Hoplolaimidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Scutellonema bradys	Hoplolaimidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Meloidogyne arenaria	Meloidogynidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Meloidogyne artiellia	Meloidogynidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Meloidogyne duytsi	Meloidogynidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Meloidogyne exigua	Meloidogynidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Meloidogyne hapla	Meloidogynidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Meloidogyne ichinohei	Meloidogynidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Meloidogyne incognita	Meloidogynidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Meloidogyne javanica	Meloidogynidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Meloidogyne maritima	Meloidogynidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Meloidogyne microtyla	Meloidogynidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea

Species Name	Family	Superfamily	Infraorder	Suborder	Order	Subclass	Class
Boleodorus sp.	Tylenchidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Geocenamus quadrifer	Belonolaimidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Tylenchorhynchus dubius	Belonolaimidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Tylenchorhynchus maximus	Belonolaimidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Nacobbus aberrans	Pratylenchidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Pratylenchoides magnicauda	Pratylenchidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Pratylenchoides ritteri	Pratylenchidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Pratylenchus goodeyi	Pratylenchidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Pratylenchus thornei	Pratylenchidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Radopholus similis	Pratylenchidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Zygotylenchus guevarai	Pratylenchidae	Tylenchoidea	Tylenchomorpha	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Deladenus sp.	Phaenopsitylenchidae		Incertae sedis	Tylenchina?	Rhabditida	Chromadoria	Chromadorea
marine Tylenchid	[on seaweed]		Incertae sedis	Tylenchina	Rhabditida	Chromadoria	Chromadorea
Bunonema franzi	Bunonematidae	Bunonematoidea	Bunonematomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Aduncospiculum halicti	Diplogasteridae	Diplogasteroidea	Diplogasteromorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Diplogaster lheritieri	Diplogasteridae	Diplogasteroidea	Diplogasteromorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Pristionchus lheritieri	Neodiplogasteridae	Diplogasteroidea	Diplogasteromorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Pristionchus pacificus	Neodiplogasteridae	Diplogasteroidea	Diplogasteromorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Diploscapter sp.	Diploscapteridae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Caenorhabditis briggsae	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Caenorhabditis elegans	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Caenorhabditis remanei	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Choriorhabditis dudichi	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Crustorhabditis scanica	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Crustorhabditis sp.	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Cruznema tripartitum	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Cuticularia sp.	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Distolabrellus veechi	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Dolichorhabditis sp.	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Mesorhabditis anisomorpha	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Mesorhabditis spiculigera	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Oscheius insectivora	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Parasitorhabditis sp.	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Pellioditis marina	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea

Species Name	Family	Superfamily	Infraorder	Suborder	Order	Subclass	Class
Pellioditis mediterranea	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Pellioditis typical	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Pelodera punctata	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Pelodera strongyloides	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Pelodera teres	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Poikilolaimus oxycerca	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Poikilolaimus regenfussi	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Protorhabditis sp.	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Rhabditella axei	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Rhabditis blumi	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Rhabditis inermiformis	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Rhabditis myriophila	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Rhabditoides inermiformis	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Rhabditoides inermis	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Rhabditoides regina	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Teratorhabditis palmarum	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Teratorhabditis synpapillata	Rhabditidae	Rhabditoidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Heterorhabditis bacteriophora	Heterorhabditidae	Strongyloidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Strongylus equinus	Strongylidae	Strongyloidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Haemonchus contortus	Trichostrongylidae	Strongyloidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Haemonchus placei	Trichostrongylidae	Strongyloidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Haemonchus similis	Trichostrongylidae	Strongyloidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Ostertagia ostertagi	Trichostrongylidae	Strongyloidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Nippostrongylus brasiliensis	Trichostrongylidae	Strongyloidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Nematodirus battus	Trichostrongylidae	Strongyloidea	Rhabditomorpha	Rhabditina	Rhabditida	Chromadoria	Chromadorea
Otostrongylus sp.	Trichostrongylidae?	Strongyloidea?	Rhabditomorpha?	Rhabditina?	Rhabditida	Chromadoria	Chromadorea
Parafilaroides sp.	Trichostrongylidae?	Strongyloidea?	Rhabditomorpha?	Rhabditina?	Rhabditida	Chromadoria	Chromadorea
Syngamus trachea	Syngamidae	Strongyloidea	Rhabditomorpha?	Rhabditina?	Rhabditida	Chromadoria	Chromadorea
Rhabditophanes sp.	Alloinematidae			Incertae sedis	Rhabditida	Chromadoria	Chromadorea

Appendix III

Table A3: Combinations of phyla in outgroup analysis and numbers of most parsimonious trees recovered. MP = most parsimonious, IG = ingroup only, ALL OGs = all in- and outgroups (40 + 22 sequences), C = Chaetognatha, G = Gastrotricha, K = Kinorhyncha, N = Nematomorpha, P = Priapulida, T = Tardigrada

Combination	No. MP Trees recovered
IG	1
ALL OGs	4
C	1
G	11
K	7
N	4
P	41
T	3
CG	2
CK	13
CN	5
CP	2
CT	8
GK	3
GN	3
GP	8
GT	7
KN	24
KP	3
KT	6
NP	1
NT	1
PT	22
CGK	23
CGN	3
CGP	1
CGT	24
CKN	1
CKP	2
CKT	4
CNP	19
CNT	6
CPT	1

Combination	No. MP Trees recovered
GKN	7
GKP	6
GKT	6
GNP	5
GNT	5
GPT	26
KNP	3
KNT	2
KPT	4
NPT	3
CGKN	2
CGKP	2
CGKT	2
CGNP	3
CGNT	6
CGPT	13
CKNP	2
CKNT	6
CKPT	6
CNPT	6
GKNP	3
GKNT	18
GKPT	32
GNPT	4
KNPT	1
CGKNP	5
CGKNT	2
CGKPT	6
CGNPT	1
CKNPT	2
GKNPT	1
CGKNPT	1

Appendix IV

Table A4.1: List of all morphological characters and their states as tested in this study. C = character number, S = state number

C	S	Character and state description
1	*A*	Labial sensilla papilliform
	B	Labial sensilla setiform
	C	Labial sensilla absent but pits in prestome present
2	*A*	Inner cephalic sensilla papilliform
	B	Inner cephalic sensilla setiform
	C	Inner cephalic sensilla only nerve ends observed (*Heterorhabditis bacteriophora* only)
3	*A*	Outer cephalic sensilla papilliform
	B	Outer cephalic sensilla setiform
	C	Outer cephalic sensilla only nerve ends observed (*Heterorhabditis bacteriophora* only)
4	*A*	Sensilla in 3 separate circles
	B	Labial and inner cephalic sensilla circles at same level (*Panagrobellus stammeri* only)
	C	Inner and outer cephalic sensilla circles at same level
5	*A*	Sensilla of inner cephalic sensilla jointed
	B	Sensilla of inner cephalic sensilla unjointed
6	*A*	Sensilla of outer cephalic sensilla jointed (*Bathylaimus assimilis* only)
	B	Sensilla of outer cephalic sensilla unjointed
7	*A*	Inner cephalic sensilla longest
	B	Outer cephalic sensilla longest
	C	Inner and outer cephalic sensilla equally long
	D	All equally long (papillae)
8	*A*	Lateral sensilla of 2nd circle shorter than submedian sensilla (*Sphaerolaimus gracilis* only)
	B	Lateral sensilla of 2nd circle equal to submedian sensilla
9	*A*	Additional cephalic setae present
	B	Additional cephalic setae absent
10	*A*	Subcephalic setae present
	B	Subcephalic setae absent
11	*A*	Cervical setae present
	B	Cervical setae absent
	C	Cervical setae papilliform (*Adoncholaimus fuscus* only)
12	*A*	Cervical papillae/dereids present
	B	More than 1 ventromedian cephalic papilla (*Trichodorus primitivus* only)
	C	Cervical papillae/dereids absent
13	*A*	Cervical alae present
	B	Cervical alae absent
14	*A*	Amphids pocket-shaped
	B	Amphids pore-shaped
	C	Amphids ventrally spiral
	D	Amphids round
	E	Amphids elongated loop
	F	Amphids slit-like
	G	Amphids dorsally spiral (*Prismatolaimus intermedius* only)
15	*A*	Amphids in labial region
	B	Amphids in head region within outer cephalic sensilla circle
	C	Amphids in head region posterior to outer cephalic sensilla circle
	D	Amphids behind head (cephalic capsule/buccal cavity)
16	*A*	Amphids laterally
	B	Amphids displaced (sub-)dorsally
17	*A*	Amphids same size in both sexes
	B	Amphids larger in male or female
18	*A*	Enoplid-type cephalic capsule present
	B	Cephalic capsule absent
19	*A*	Desmodorid-type cephalic capsule present
	B	Cephalic capsule absent

C	S	Character and state description
20	*A*	Cephalic framework sclerotised
	B	Cephalic framework absent
21	*A*	Head with dorsal and ventral cuticular inflations
	B	Head without dorsal and ventral cuticular inflations
22	*A*	Cephalic organs present
	B	Cephalic organs absent
23	*A*	6 lips open
	B	6 lips closed
	C	6 lips & 6 deep cuticularised incisions & 6 extensions
	D	3 lips (Ascarid-type)
	E	3 lips & 3 interlabia (Ascarid-type)
	F	Female with labial disc, male with 4 lips, markedly set-off
	G	3 bilobed lips (Tripyloididae only)
	H	2 fleshy trilobed lips (pseudolabia) (*Gnathostoma turgidum* only)
	I	Lips amalgamated/merged (Dorylaimid-type)
	J	Lips merged & fused to labial disc (Tylenchid-type)
	K	3 lips (Enoplid-type)
	L	Circumoral membrane instead of lips
	M	3 lips (Strongyloid-type)
	N	Pair of lateral pseudolabia connected via latero-median plate to buccal capsule (*Ascarophis arctica* only)
	O	Leaf crown (corona radiata) present instead of lips (*Strongylus equinus* only)
	P	3 lips with denticular ridges (Ascarid-type)
	Q	3 lips & 3 interlabia with denticular ridges (Ascarid-type)
	R	Pseudolips with extensions 30 degrees off normal lip axes (*Panagrobellus redivivus* only)
24	*A*	3 labial probolae & 6 cephalic probolae present
	B	Labial & cephalic probolae absent
25	*A*	Lips smooth
	B	Lips with longitudinal ridges (from cheilorhabdia)
	C	Lips with traverse ridges (*Enoploides brunettii* only)
26	*A*	Walls of vestibule smooth
	B	Walls of vestibule longitudinally striated (pleated, rugae; from Cheilorhabdia)
27	*A*	Walls of vestibule smooth
	B	Walls of vestibule transversely striated
28	*A*	Head continuous with neck and cephalic region
	B	Head set off in neck region
29	*A*	Lips continuous with head
	B	Lips set off from head
30	*A*	Buccal cavity terminally
	B	Buccal cavity dorsally displaced (*Syngamus trachea* only)
	C	Buccal cavity on right side of head (*Bunonema franzi* only)
31	*A*	Sexual dimorphism of buccal cavity absent
	B	Males and juveniles with less developed or no stylet
32	*A*	No buccal cavity dimorphism
	B	Males all stenostomatous, females both stenostomatous & eurostomatous (*Aduncospiculum halicti* only)
33	*A*	Oesophageal tissue reaches buccal cavity & might surround it
	B	Oesophageal tissue commences at end of stylet with different tissue surrounding stylet
	C	Buccal cavity absent
	D	Buccal cavity with rhabditoid pharyngeal sleeve
	E	Buccal cavity free as in Diphtherophoroidea (little oesophageal musculature)
	F	Buccal cavity with cephaloboid pharyngeal sleeve
	G	Oesophageal tissue commences behind spear
34	*A*	Inner oesophageal wall cuticularised throughout
	B	Inner oesophageal wall cuticularised only anteriorly
	C	Inner oesophageal wall cuticularised only posteriorly (*Monoposthia costata* only)
	D	Inner oesophageal wall not cuticularised at all
	E	Inner oesophageal wall weakly cuticularised
35	*A*	Lumen of procorpus tripartite

C	S	Character and state description
	B	Lumen of procorpus tubular
36	*A*	Lumen of metacorpus tripartite
	B	Lumen of metacorpus tubular
37	*A*	Lumen of postcorpus tripartite
	B	Lumen of postcorpus tubular
38	*A*	Distinct constriction (isthmus) between 2 parts of oesophagus present
	B	Distinct constriction (isthmus) between 2 parts of oesophagus absent
39	*A*	Procorpus muscular
	B	Procorpus non-muscular (Trichodoridae only)
40	*A*	Metacorpus muscular
	B	Metacorpus glandular
	C	Metacorpus weakly or non-muscular
41	*A*	Metacorpus with bulb
	B	Metacorpus without bulb
42	*A*	Metacorpus with valve
	B	Metacorpus without valve
43	*A*	Postcorpus muscular
	B	Postcorpus glandular
44	*A*	Postcorpus with bulb
	B	Postcorpus without bulb
	C	Postcorpus with elongated swelling
45	*A*	Postcorpus with valve
	B	Postcorpus without valve
46	*A*	Postcorpus with plasmic interruptions
	B	Postcorpus without plasmic interruptions
47	*A*	Postcorpus with teeth
	B	Postcorpus without teeth
48	*A*	Postcorpus with double haustrulum
	B	Postcorpus without double haustrulum
49	*A*	Oesophageal tubes present
	B	Oesophageal tubes absent
50	*A*	Stichosome present
	B	Stichosome absent
51	*A*	Sexual dimorphism of oesophagus absent
	B	Bulb in males, juveniles or infective juvenile stage 3 less developed
52	*A*	Sexual dimorphism of buccal cavity & oesophagus absent
	B	Buccal cavity and oesophagus very different between male & female (*Calyptronema maxweberi* only)
53	*A*	3 uninuclear oesophageal glands present
	B	5 uninuclear oesophageal glands present
	C	3 oesophageal glands, 2 subventral glands dinuclear
	D	3 oesophageal glands polynuclear
54	*A*	Oesophageal gland nuclei in posterior part of pharyngeal lining or in bulb (if present)
	B	Oesophageal gland nuclei in lobe lying freely in body cavity
55	*A*	Dorsal oesophageal gland opens at base of stoma (or near base)
	B	Dorsal oesophageal gland opens in stoma through tooth
	C	Dorsal oesophageal gland opens in median bulb anterior to valve
	D	Dorsal oesophageal gland opens just behind spear knobs
	E	Dorsal oesophageal gland opens through gutter in buccal cavity (*Strongylus equinus* only)
	F	Dorsal oesophageal gland opens in anterior oesophageal region
	G	Dorsal oesophageal gland opens in posterior/glandular part of oesophageal region
	H	Dorsal oesophageal gland opens in mid/posterior oesophagus and gland extends further back into body cavity
56	*A*	First pair of subventral glands (SV1) open at or near base of stoma
	B	SV1 open in stoma through teeth
	C	SV1 open in mid-oesophageal region/metacorpus/posterior to median bulb valve/near nerve ring
	D	SV1 open in oesophagus posterior to nerve ring, anterior to postcorpus
	E	SV1 open in posterior, glandular part just behind dorsal gland opening

C	S	Character and state description
	F	SV1 open behind nerve ring at the base of the corpus/anterior to bulb
	G	SV1 glands extend posteriorly
57	*A*	Unicellular & uninuclear
	B	H-type & derivates, i.e. lateral canales without glands
	C	V and SD glands, without lateral canales
	D	Two parallel SV glands and canals (Rhabditoid-type)
	E	Tylenchoid-type, asymmetric, possibly with gland cell
58	*A*	Cervical gland entirely in oesophageal body section
	B	Cervical gland extends post-pharyngeal
	C	Cervical gland in cardial region (*Monoposthia costata* only)
59	*A*	Cervical gland ventrally of oesophagus/intestine
	B	Cervical gland ventrally and subdorsally of oesophagus/intestine
	C	Cervical gland right of oesophagus/intestine
	D	Cervical gland left of oesophagus/intestine
60	*A*	Excretory pore in head region
	B	Excretory pore in oesophageal region anterior to nerve ring
	C	Excretory pore in oesophageal region at level with nerve ring
	D	Excretory pore in oesophageal region posterior to nerve ring
	E	Excretory pore in anterior part of intestine
	F	Excretory pore in posterior part of intestine (*Tylenchulus semipenetrans* only)
61	*A*	Excretory pore in head region
	B	Excretory pore in oesophageal region
	C	Excretory pore in intestinal region
62	*A*	Hemizonoid present
	B	Hemizonoid absent
63	*A*	Cardia clearly visible as distinct organ
	B	Cardia not visible
64	*A*	Cardia lumen triradiate/triangular
	B	Cardia lumen round
	C	Cardia lumen dorsoventrally flattened
	D	Cardia lumen vertically flattened
65	*A*	Cardia with posterior appendage
	B	Cardia without posterior appendage
66	*A*	Cardia with valve
	B	Cardia without valve
67	*A*	Cardial glands present
	B	Cardial glands absent
68	*A*	Intestine without symbiotic bacteria
	B	Intestine with symbiotic bacteria
69	*A*	Trophosome absent
	B	Trophosome present (*Mermis nigrescens* only)
70	*A*	Intestinal caecum/diverticulum anterior to cardia present
	B	Intestinal caecum/diverticulum anterior to cardia absent
71	*A*	Intestinal caecum posterior to anus present
	B	Intestinal caecum posterior to anus absent
72	*A*	Rectum & anus fully developed
	B	Rectum & anus rudimentary or absent
	C	Rectum & anus closed in infective larva stage 3
73	*A*	Anus with pre-rectum
	B	Anus with indistinct pre-rectum (*Xiphinema theresiae* only)
	C	Anus with Trichodorid-type pre-rectum (*Trichodorus primitivus* only)
	D	Anus without pre-rectum
74	*A*	Anterior gonad to right of intestine
	B	Anterior gonad to left of intestine
	C	Anterior gonad ventrally to intestine
	D	Anterior gonad position not constant
75	*A*	Posterior gonad to right of intestine
	B	Posterior gonad to left of intestine
	C	Posterior gonad ventrally to intestine

C	S	Character and state description
	D	Posterior gonad position not constant
76	*A*	2 opposed ovaries (didelphic-amphidelphic)
	B	2 opposed ovaries (didelphic-amphidelphic) but posterior weakly developed (*Cruznema tripartitum* only)
	C	1 anterior ovary (monodelphic-prodelphic)
	D	1 anterior ovary (monodelphic-prodelphic) with post-vulval sac
	E	1 posterior ovary (monodelphic-opistodelphic)
	F	1 posterior ovary (monodelphic-opistodelphic) with pre-vulval sac (*Belondira apitica* only)
	G	2 posterior ovaries (didelphic-opistodelphic)
	H	2 anterior ovaries (didelphic-prodelphic)
77	*A*	Ovaries outstretched
	B	Ovaries antidromously reflexed
	C	Ovaries homodromously reflexed
	D	Ovaries antidromously reflexed at oviduct
	E	Ovaries irregularly coiled
78	*A*	Tip of ovary straight
	B	Tip of ovary folded over
79	*A*	Vulva in oesophageal region
	B	Vulva in intestinal region anteriorly to mid-body (approx. <40%)
	C	Vulva in intestinal region approx. at mid-body (approx. 40-60%)
	D	Vulva in intestinal region posteriorly to mid-body (approx. >60%)
	E	Vulva in intestinal region far posteriorly to mid-body (approx. >80%)
	F	Vulva terminal
80	*A*	Vulva in oesophageal region
	B	Vulva in intestinal region
	C	Vulva terminal
81	*A*	Demanian organ present
	B	Demanian organ absent
82	*A*	Reproduction oviparous
	B	Reproduction oviparous/ovoviviparous - larvae or embryos develop in eggs but hatch after release from female
	C	Reproduction (ovo)viviparous - larvae/juveniles hatch from eggs in uterus
83	*A*	Reproduction oviparous
	B	Reproduction (ovo)viviparous
84	*A*	Spermathecae present
	B	Spermathecae absent
85	*A*	Ovijector present
	B	Ovijector absent
86	*A*	Z-organ present
	B	Z-organ absent
87	*A*	Vulval papillae present
	B	Vulval papillae absent
88	*A*	Vulval flaps present
	B	Vulval flaps absent
89	*A*	Vulva glands present
	B	Vulva glands absent
90	*A*	2 opposed testes (di-orchic, amphi-orchic)
	B	2 anterior testes (di-orchic, pro-orchic)
	C	1 anterior testis (mono-orchic, pro-orchic)
	D	1 posterior testis (mono-orchic, opisto-orchic)
91	*A*	Testes straight
	B	Testes reflexed or tip folded over
	C	Posterior testis continues anteriorly at first and then bends backwards
92	*A*	1 spicule present (*Trichuris muris* only)
	B	2 spicules present
	C	Spicules absent
93	*A*	Spicules equal in size
	B	Spicules unequal in size
94	*A*	Spicules equal in shape

C	S	Character and state description
	B	Spicules unequal in shape
95	*A*	Spicules separate
	B	Spicules fused (*Haemonchus contortus* only)
	C	Spicules partially fused distally
96	*A*	Spicules with ventral alae
	B	Spicules without ventral alae
97	*A*	Spicules with sheath (*Trichuris muris* only)
	B	Spicules without sheath
98	*A*	Gubernaculum present
	B	Gubernaculum absent
99	*A*	2 lateral accessory pieces present
	B	2 lateral accessory pieces absent
100	*A*	Gubernaculum with caudal process(es) (apophysis)
	B	Gubernaculum without caudal processes (apophysis)
101	*A*	Rectal glands present
	B	Rectal glands absent
102	*A*	Preanal papillae in ventromedian row present
	B	Preanal papillae in ventromedian row absent
103	*A*	Cup-shaped preanal papillae in ventromedian row present
	B	Cup-shaped preanal papillae in ventromedian row absent
104	*A*	Preanal papillae in ventromedian row present & reach all the way into oesophageal region (*Prismatolaimus intermedius* only)
	B	Preanal papillae in ventromedian row do not reach all the way into oesophageal region
105	*A*	Preanal papillae in 2 subventral rows present
	B	Preanal papillae in 2 subventral rows absent
106	*A*	2 preanal papillae with conical tip & pore present (*Nudora bipapillata* only)
	B	2 preanal papillae with conical tip & pore absent
107	*A*	Peri/ad-anal papillae present
	B	Peri/ad-anal papillae absent
108	*A*	Postanal papillae in ventromedian row present
	B	Postanal papillae in ventromedian row absent
109	*A*	Postanal papillae in 2 subventral rows present
	B	Postanal papillae in 2 subventral rows absent
110	*A*	Postanal papillae ventromedian and dorsally present (*Panagrellus redivivus* only)
	B	Postanal papillae ventromedian and dorsally absent
111	*A*	Papillae sessile
	B	Papillae pedunculate
112	*A*	Preanal tubules present (always medioventral)
	B	Preanal tubules absent
113	*A*	Single, trumpet-shaped preanal tubule present
	B	Single, trumpet-shaped preanal tubule absent
114	*A*	Supplements with glands
	B	Supplements without glands
115	*A*	Preanal setae in 1 ventromedian row present
	B	Preanal setae in 1 ventromedian row absent
116	*A*	Preanal setae in 2 subventral rows present
	B	Preanal setae in 2 subventral rows absent
117	*A*	Peri/ad-anal setae in 2 subventral rows present
	B	Peri/ad-anal setae absent
118	*A*	Postanal setae in 2 subventral rows present
	B	Postanal setae in 2 subventral rows absent
119	*A*	Preanal thorns present
	B	Preanal thorns absent
120	*A*	Postanal thorns present
	B	Postanal thorns absent
121	*A*	SCO papillae pre- and postanal in several groups or pairs present
	B	SCO papillae pre- and postanal in several groups or pairs absent
122	*A*	Bursa absent
	B	Bursa leptoderan (with tail, open)

C	S	Character and state description
	C	Bursa peloderan (without tail, closed) (*Caenorhabditis elegans* only)
	D	Bursa peloderan (without tail, open)
	E	Bursa adanal, open
	F	Bursa with larges lobes & rays
	G	Bursa reduced to narrow velum either side (*Dirofilaria immitis* only)
	H	Bursa as simple caudal, lateral alae
123	*A*	Bursa or caudal alae present
	B	Bursa absent
124	*A*	Bursa with genital cone
	B	Bursa without genital cone
125	*A*	Roughened cuticular patched/pads present
	B	Roughened cuticular patched/pads absent
126	*A*	Denticulate caudal plates present (*Pseudoterranova decipiens* only)
	B	Denticulate caudal plates absent
127	*A*	Body with simple curve or more or less straight
	B	Female saccate (spindle-shaped)
128	*A*	Body cavity filled by biocrystals (*Ironus dentifurcatus* only)
	B	Body cavity not filled by biocrystals
129	*A*	Cuticle with transverse striations (annulated)
	B	Cuticle articulated
	C	Cuticle with very fine striations, barely visible
	D	Cuticle smooth
130	*A*	Cuticle with transverse striations (annulated)
	B	Cuticle without transverse striations
131	*A*	Cuticle articulated
	B	Cuticle not articulated
132	*A*	Cuticle with very fine striations, barely visible
	B	Cuticle without fine striations
133	*A*	Cuticle striated or articulated
	B	Cuticle smooth
134	*A*	Cuticle with longitudinal striations
	B	Cuticle without longitudinal striations
135	*A*	Cuticular patterns also present on head
	B	Cuticular patterns absent on head
136	*A*	Cuticular patterns also present on tip of tail
	B	Cuticular patterns absent on tip of tail
137	*A*	Cuticular patterns present on main body (*Gnathostoma turgidum* only)
	B	Cuticular patterns absent from about mid-body
138	*A*	Cuticular patterns only cervically on cuticular inflation
	B	Cuticular patterns not only cervically on cuticular inflation
139	*A*	Annules smooth without scales & spines
	B	Annules retrorse with scales & spines
140	*A*	Cuticle with subcuticular longitudinal incisures (*Aphelenchus avenae* only)
	B	Cuticle without subcuticular longitudinal incisures or annules
141	*A*	Cuticle with subcuticular annulation
	B	Cuticle without subcuticular annulation
142	*A*	Cuticle with punctation
	B	Cuticle without punctation
143	*A*	Cuticle with subcuticular punctation
	B	Cuticle without subcuticular punctation
144	*A*	Cuticle with bosses near vulva (*Haemonchus contortus* only)
	B	Cuticle without bosses near vulva
145	*A*	Cuticle with bosses randomly arranged (not head or tail) (*Loa loa* only)
	B	Cuticle without bosses randomly arranged
146	*A*	Cuticle with bosses only on female tail (*Wuchereria bancrofti* only)
	B	Cuticle without bosses on female tail only
147	*A*	Cuticle with warts & 'Kriechleisten' (*Bunonema franzi* only)
	B	Cuticle without warts & 'Kriechleisten'
148	*A*	Cuticle covered with coat of eubacteria or cyanobacteria

C	S	Character and state description
	B	Cuticle not covered with bacteria coat
149	*A*	Cuticle inflated cervically
	B	Cuticle not inflated cervically
150	*A*	Cuticle with bacillary band (*Trichuris muris* only)
	B	Cuticle without bacillary band
151	*A*	Annules smooth without V-like projections
	B	Annules with V-like projections
152	*A*	Annules smooth without spiral thickenings
	B	Annules with spiral thickenings (*Onchocerca cervicalis* only)
153	*A*	Lateral field with longitudinal incisures (stronger than rest of body)
	B	Lateral field without longitudinal incisures (or same as rest of body)
154	*A*	Lateral field with alae
	B	Lateral field without alae
155	*A*	Lateral field areolated (transverse striations continue across lateral field)
	B	Lateral field not areolated
156	*A*	Lateral field punctated differently to rest of body
	B	Lateral field not punctated (or same as rest of body)
157	*A*	Lateral field granular
	B	Lateral field not granular
158	*A*	Hypodermic glands/pores present
	B	Hypodermic glands/pores absent
159	*A*	Ocelli present
	B	Ocelli absent
160	*A*	Pigment spot present
	B	Pigment spot absent
161	*A*	Somatic setae present
	B	Somatic setae absent
	C	Somatic setae papillose
162	*A*	Glandular setae present (*Laxus* only)
	B	Glandular setae absent
163	*A*	Metanemes present
	B	Metanemes absent
164	*A*	Phasmids present
	B	Phasmids absent
165	*A*	Prophasmids present
	B	Prophasmids absent
166	*A*	Postdereids present
	B	Postdereids absent
167	*A*	Caudal pores present
	B	Caudal pores absent
168	*A*	Terminal caudal papillae present
	B	Terminal caudal papillae absent
169	*A*	Female tail conical
	B	Female tail conico-cylindrical
	C	Female tail conical-filiform
	D	Female tail short, conical, strongly bent
	E	Female tail conico-cylindrical with swollen tip (clavate)
	F	Female tail round-cylindrical, longer than anal diameter
	G	Female tail conical with blunt tip and mucron
	H	Female tail short & sharply pointed (*Contracaecum multipapillatum* only)
	I	Female tail conical/rounded with tuft of projections (*Hysterothylacium fortalezae* only)
	J	Female tail conical with many terminal flaps (*Mermis nigrescens* only)
	K	Female tail round-hemispherical, no longer than anal diameter
	L	Female tail absent due to terminal anus
	M	Female tail very short and round due to subterminal anus
170	*A*	Male tail conical
	B	Male tail conico-cylindrical
	C	Male tail conical-filiform
	D	Male tail short, conical, strongly bent

C	S	Character and state description
	E	Male tail conico-cylindrical with swollen tip (clavate)
	F	Male tail round-cylindrical, longer than anal diameter
	G	Male tail conical with blunt tip and mucron
	H	Male tail short & sharply pointed
	I	Male tail conical/rounded with tuft of projections (*Hysterothylacium fortalezae* only)
	J	Male tail mermithoid (*Mermis nigrescens* only)
	K	Male tail round-hemispherical, no longer than anal diameter
	L	Male tail absent due to terminal cloaca
	M	Male tail conical with pair of small appendages near tip of tail and terminal knob (*Setaria digitata* only)
	N	Male tail short and conical (with large bursa)
	O	Male tail conical with spicate terminus (*Aduncospiculum halicti* [or G] only)
	P	Male tail short & stubby (*Gnathostoma turgidum* only)
171	*A*	Terminal setae present
	B	Terminal setae absent
172	*A*	Sexual dimorphism of tail shape present
	B	Sexual dimorphism of tail shape absent
173	*A*	Caudal setae present with same pattern in males and females
	B	Caudal setae more numerous in male
	C	Caudal setae absent
174	*A*	Caudal setae present
	B	Caudal setae absent
175	*A*	Terminal flaps present (pseudobursa) (*Trichinella spiralis* only)
	B	Terminal flaps absent (pseudobursa)
176	*A*	Caudal glands present
	B	Caudal glands absent
	C	No caudal glands observed but terminal duct opening
177	*A*	1 common caudal gland opening
	B	Separate openings of caudal gland ducts
178	*A*	Spinneret well set-off
	B	Spinneret continuous with tail end
179	*A*	Caudal glands situated completely in the tail (caudally)
	B	Caudal glands reach pre-caudally (at least in male)
180	*A*	Dioecious or amphimictic (two separate sexes, copulation required)
	B	Hermaphroditic (if males rare, copulation optional, sperm & oocytes produced in female)
	C	Parthenogenesis prevailing (if males rare, copulation optional, reproduction without sperm)

Table A4.2: Matrix of morphological characters (on CD)

Table A4.3: List of codes for references used for each species in the collection of morphological characters (on CD)

Table A4.4: References used for the morphological analysis of species (on CD)

Appendix V

Table A5.1: Character change list of all morphological characters mapped on the consensus tree of the analysis of all morphological characters under weighted parsimony (on CD)

Figure A5.1: Strict consensus tree from phylogenetic analysis of all morphological characters under weighted parsimony; numbers at nodes correspond to node numbers in Table A5.1 (on CD)

Table A5.2: Character change list of all morphological characters mapped on the consensus tree of the analysis of the combined characters excluding positions with gaps in the molecular data set and all uninformative characters under weighted parsimony (on CD)

Figure A5.2: Strict consensus tree from phylogenetic analysis of all characters combined excluding positions with gaps in the molecular character set and all uninformative characters under weighted parsimony; numbers at nodes correspond to node numbers in Table A5.2 (on CD)

Appendix VI

Table A6.1: Nexus files of the molecular character matrix (on CD)

Table A6.2: Nexus files of the morphological character matrix (on CD)

Table A6.3: Nexus files of the combined character matrix (on CD)

References

Adams, E. N. (1972). Consensus techniques and the comparison of taxonomic trees. *Systematic Zoology* **21**:390-397

Adamson, M. L. (1987). Phylogenetic analysis of the higher classification of the Nematoda. *Canadian Journal of Zoology* **65**:1478-1482

Ahmad, W.; Jairajpuri, M. S. (1979). On the systematics of Nygolaimina new suborder (Nematoda: Dorylaimida). Second National Congress of Parasitology, B. H. U., India, p. 29 (Abstract)

Albertson, D. G.; Thomson, J. N. (1976). The pharynx of *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society of London Series B – Biological Sciences* **275**:299-325

Albrech, P. (1985). Problems with the interpretation of developmental sequences. *Systematic Zoology* **34**(1):46-58

Aleshin, V. V.; Kedrova, O. S.; Milyutina, I. A.; Vladychenskaya, N. S.; Petrov, N. B. (1998). Relationships among nematodes based on the analysis of 18S rRNA gene sequences: Molecular evidence for monophyly of chromadorian and secernentian nematodes. *Russian Journal of Nematology* **6**(2):175-184

Alvarez, B.; Crisp, M. D.; Driver, F.; Hooper, J. N. A.; van Soest, R. W. M. (2000). Phylogenetic relationships of the family Axinellidae (Porifera: Demospongiae) using morphological and molecular data. *Zoologica Scripta* **29**(2):169-198

Anderson, R. C.; Bain, O. (1976). CIH keys to the nematode parasites of vertebrates. CAB International Publications **3**(3):59-116

Andrássy, I. (1959). Taxonomische Übersicht der Dorylaimiden (Nematoda). I. *Acta Zoologica, Academiae Scientiarum Hungaricae, Budapest* **5**(3-4):191-240

Andrássy, I. (1960). Taxonomische Übersicht der Dorylaimen (Nematoda). II. *Acta Zoologica, Academiae Scientiarum Hungaricae, Budapest* **6**(1-2):1-28

Andrássy, I. (1962). Über den Mundstachel der Tylenchiden (Nematologische Notizen 9). *Acta Zoologica Academidae Scientiarum Hungaricae, Budapest* **8**:241-249

Andrássy, I. (1976). Evolution as a Basis for the Systematization of Nematodes. Pitman, London, 288 pp.

Arnheim, N.; Krystal, M.; Schmickel, R.; Wilson, G.; Ryder, O.; Zimmer, E. (1980). Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and ape. *Proceeding of the National Academy of Science USA* **77**:7323-7327

Arnold, E. N. (1981). Estimating phylogenies at low taxonomic levels. *Zeitschrift für Zoologische Systematik und Evolutionsforschung, Sonderdruck* **19**(1):1-35

Baker, R. H.; DeSalle, R. (1997). Multiple sources of character information and the phylogeny of Hawaiian drosophilids. *Systematic Biology* **46**:654-673

Baker, R. H.; Wilkinson, G. S.; deSalle, R. (2001). Phylogenetic utility of different types of molecular data used to infer evolutionary relationships among stalk-eyed flies (Diopsidae). *Systematic Biology* **50**(1):87-105

Baker, R. H.; Yu, X.; DeSalle, R. (1998). Assessing the relative contribution of molecular and morphological characters in simultaneous analysis trees. *Molecular Phylogenetics and Evolution* **9**(3):427-436

Baldwin, J. G.; Edelman, C. D. (1995). Buccal capsule of *Zeldia punctata* (Nemata: Cephalobidae): An ultrastructural study. *Canadian Journal of Zoology* **73**:648-656 1995

Baldwin, J. G.; Giblin-Davis, R. M.; Eddleman, C. D.; Williams, D. S.; Vida, J. T.; Thomas, W. K. (1997). Buccal capsule of *Aduncospiculum halicti* (Nemata: Diplogasterina): An ultrastructure and molecular phylogenetic study. *Canadian Journal of Zoology* **75**:407-423

Ballard, J. W. O.; Thayer, M. K.; Newton, A. F. Jr.; Grismer, E. R. (1998). Data sets, partitions, and characters: Philosophies and procedures for analysing multiple data sets. *Systematic Biology* **47**(3):367-396

Barrett, M.; Donoghue, M. J.; Sober, E. (1991). Against consensus. *Systematic Zoology* **40**(4):486-493

Barriel, V.; Darlu, P.; Tassy, P. (1993). Mammalian phylogeny and conflicts between morphological and molecular data. *Annales des Sciences Naturelles – Zoologie et Biologie Animale* **14**(4):157-171

Barriel, V.; Tassy, P. (1998). Rooting with multiple outgroups: Consensus versus parsimony. *Cladistics* **14**:193-200

Bastian, H. C. II. (1865). Monograph of the Anguillulidae, or free nematodes, marine, land, and fresh water; with descriptions of 100 new species. *Transactions of the Linnaean Society of London* **25**:73-184

Bates, T. (1763). An essay towards solving a problem in the doctrine of chance. *Philosophical Transactions of the Royal Society of London Series B – Biological Sciences* **53**:370-418 [Reprinted *in* Studies in the History of Statistics and Probability. (1970) Pearson, E. S.; Kendall, M. G. (Eds.), Charles Griffin, London, pp. 131-153.]

Baylis, H. A.; Daubney, R. (1926). A Synopsis of the Families and Genera of Nematoda. Richard Clay & Sons, Ltd., Bungay, Suffolk, 277 pp.

Bermingham, E.; Avise, J. (1986). Molecular zoogeography of freshwater fishes in the southeastern United States. *Genetics* **113**:939-965

Bishop, M. J.; Friday, A. E. (1985). Evolutionary trees from nucleic acid and protein sequences. *Proceedings of the Royal Society of London B* **226**:271-302

Blaxter, M. L.; de Ley, P.; Garey, J. R.; Liu, L. X; Scheldemann, P.; Vierstraete, A.; Vanfleteren, J. R.; Mackey, L. Y.; Dorris, M.; Frisse, L. M.; Vida, J. T.; Thomas, W. K. (1998). A molecular evolutionary framework for the phylum nematoda. *Nature* **392**(6671):71-75

Blaxter, M. L.; Dorris, M.; de Ley, P. (2000). Patterns and processes in the evolution of animal parasitic nematodes. *Nematology* **2**(1):43-55

Borellus, P. (1656). Historiarum, et observationum medico-physicarum, centuriae quattuor. Parisiis, 384 pp.

Bradford, J. C. (2002). Molecular phylogenetics and morphological evolution in Cunonieae (Cunoniaceae). *Annals of the Missouri Botanical Garden* **89**(4):491-503

Bremer, B. (1996). Combined and separate analysis of morphological and molecular data in the plant family Rubiaceae. *Cladistics* **12**:21-40

Bremer, B.; Jansen, R. K.; Oxelman, B.; Backlund, M.; Lantz, H.; Kim, K.-J. (1999). More characters or more taxa for a robust phylogeny - case study from the coffee family (Rubiaceae). *Systematic Biology* **48**(3):413-435

Bremer, K. (1988). The limits of amino acid data in angiosperm phylogenetic reconstruction. *Systematic Biology* **42**(4):795-803 1988

Bremer, K. (1990). Combinable component consensus. *Cladistics* **6**:369-372

Bremer, K. (1994). Branch support and tree stability. *Cladistics* **10**:295-304

Brooks, D. R.; Wiley, E. O. (1985). Theories and methods in different approaches to phylogenetic systematics. *Cladistics* **1**:1-11

Bryant, H. N. (1992). The polarization of character transformations in phylogenetic systematics: Role of axiomatic and auxiliary assumptions. *Systematic Zoology* **40**:433-445

Buckley, T. R.; Cunningham, C. W. (2002). The effects of nucleotide substitution model assumptions on estimates of nonparametric bootstrap support. *Molecular Biology and Evolution* **19**(4):394-405

Bull, J. J.; Huelsenbeck, J. P.; Cunningham, C. W.; Swofford, D. L.; Waddell, P. J. (1993). Partitioning and combining data in phylogenetic analysis. *Systematic Biology* **42**(3):384-397

Bütschli, O. (1873). Breiträge zur Kenntnis der freilebenden Nematoden. *Nova Acta Academiae Caesareae Leopoldino-Carolinae Germanicae Naturae Curiosorum* **36**(5):1-124

Bütschli, O. (1874). Zur Kenntnis der freilebenden Nematoden, insbesondere der des Kieler Hafens. *Abhandlungen der Senkenbergischen Naturforschenden Gesellschaft* **9**:236-292

Bütschli, O. (1876). Untersuchungen über freilebende Nematoden und die Gattung *Chaetonotus*. *Zeitschrift für Wissenschaftliche Zoologie* **26**(4):363-413

Campbell, A. J. D.; Gasser, R. B.; Chilton, N. B. (1995). Differences in ribosomal DNA sequences of *Strongylus* species allows identification of single eggs. *International Journal for Parasitology* **25**(3):359-365

Campbell, J. A.; Frost, D. R. (1993). Anguid lizards of the genus *Abronia*: Revisionary notes, descriptions of four new species, a phylogenetic analysis and key. *Bulletin of the American Museum of Natural History* **216**:1-121

Cantrell, S. A.; Hanlin, R. T. (1997). Phylogenetic relationships in the family Hyaloscyphaceae inferred from sequences of ITS regions, 5.8S ribosomal DNA and morphological characters. *Mycologia* **89**(5):745-755

Carnap, R. (1950). Logical Foundations of Probability. University of Chicago Press, Chicago

Carpenter, J. M. (1988). Choosing among multiple equally parsimonious cladograms. *Cladistics* **4**:291-296

Carpenter, J. M. (1996). Uninformative bootstrapping. *Cladistics* **12**:177-181

Carta, L. K.; Wergin, W. P.; Erbe, E. F.; Murphy, C. A. (2003). A comparison of low-temperature SEM for viewing nematode faces. *Journal of Nematology* **35**(1):78-81

Cavalli-Sforza, L. L.; Edwards, A. W. F. (1967). Phylogenetic analysis: Models and estimation procedures. *Evolution* **32**:550-570 and *American Journal of Human Genetics* **19**:233-257

Chilton, N. B.; Gasser, R. B. (1999). Sequence differences in the internal transcribed spacer rDNA amongst four species of hookworm (Ancylostomatoidea, *Ancylostoma*). *International Journal for Parasitology* **29**(12):1971-1977

Chippindale, P. T.; Wiens, J. J. (1994). Weighting, partitioning, and combining characters in phylogenetic analysis. *Systematic Biology* **43**(2):278-287

Chitwood, B. G. (1933). A revised classification of the Nematoda. *Journal of Parasitology* **20**(2):131

Chitwood, B. G. (1937). A revised classification of the Nematoda. *In:* Papers on Helminthology, 30 Year Jubileum K. J. Skrjabin. Edited by Anon., All-Union Lenin Academy of Agricultural Sciences, Moscow, pp. 67-80

Chitwood, B. G. (1958). The designation of official names for higher taxa of invertebrates. *Bulletin of Zoological Nomenclature* **15**:860-895

Chitwood, B. G.; Chitwood, M. B. (1933). The characters of a protonematode. *Journal of Parasitology* **20**:130

Chitwood, B. G.; Chitwood, M. B. (1950). An Introduction to Nematology. 2nd Edition. Monumental Printing Company, Baltimore, MD, 372 pp.

Chitwood, B. G.; Wehr, E. E. (1934). The value of cephalic structures as characters in nematode classification with special reference to the superfamily Spiruroidea. *Zeitschrift für Parasitenkunde* **7**(3):273-335

Churchill, G. A.; von Haeseler, A.; Navidi, W. C. (1992). Sample size for a phylogenetic inference. *Molecular Biology and Evolution* **9**:753-769

Clark, W. C. (1961). A revised classification of the Enoplida (Nematoda). *New Zealand Journal of Science* **4**:123-150

Clark, W. C. (1962). The systematic position of the Alaimidae and the Diphtherophoroidea (Enoplida, Nematoda). *Nematologica* **7**:119-121

Cobb, N. A. (1894). *Tricoma* and other nematode genera. *Proceedings of the Linnaean Society (N. S. W.)* **8**(2):389-412

Cobb, N. A. (1915). Nematodes and their relationships. U. S. Department of Agriculture Yearbook for 1914, pp. 457-490

Cobb, N. A. (1919). The orders and classes of Nemas. *Contributions to a Science of Nematology* **8**:213-216

Cobb, N. A. (1920). One hundred new nemas. (Type species of 100 new genera). *Contributions to a Science of Nematology* **9**:217-343

Colless, D. H. (1983). Wagner trees in theory and practice. *In:* Numerical Taxonomy. Felsenstein, J. (Ed.), NATO ASI Series, G. Springer Verlag, Berlin, pp. 259-278

Coomans, A. (1977). Evolution as a basis for the systematization of nematodes - a critical review and exposé. *Nematologica* **23**(2):129-136

Coomans, A. (2002). Present status and future of nematode systematics. *Nematology* **4**(5):573-582

Coomans, A.; Loof, P. A. A. (1970). Morphology and taxonomy of Bathyodontina (Dorylaimida). *Nematologica* **16**:180-196

Cracraft, J.; Mindell, D. P. (1989). The early history of modern birds: A comparison of molecular and morphological evidence. *In:* The Hierarchy of Life. Fernholm, B.; Bremer, K.; Jörnvall, H. (Eds.), Elsevier Science, Amsterdam, pp. 389-403

Cummings, M. P.; Otto, S. P.; Wakeley, J. (1995). Sampling properties of DNA sequence data in phylogenetic analysis. *Molecular Biology and Evolution* **12**(5):814-822

Davis, J. I. (1993). Character removal as a means for assessing stability of clades. *Cladistics* **9**:201-210

de Coninck, L. A. P. (1942). De symmetrie-verhoudingen aan het voorende der (vrijlevende) Nematoden. *Natuurwetenschappelijk Tijdschrift* **24**:29-68

de Coninck, L. A. P. (1965). Systématique des Nématodes. *In:* Grassé, Traite des Zoologie **4**(2)

de Coninck, L. A. P.; Schuurmans Stekhoven Jr., J. H. (1933). The freeliving marine nemas of the Belgian coast. II. With general remarks on the structure and the system of nemas. *Mémoires du Musée Royal d'Histoire Naturelle de Belgique [Verhandelingen van het Koninklijk Natuurhistorisch Museum van België]* **58**:1-163

de Ley, P.; Blaxter, M. L. (2002). 1. Systematic position and phylogeny. *In:* The Biology of Nematodes. Lee, D. (Ed.), Harwood Acadmic Publishers, Reading, pp. 1-30

de Ley, P.; de Grisse, A.; van der Velde, M. C.; Coomans, A. (1993). Ultrastructure of the stoma in *Tylopharynx*. *Mededelingen van de Faculteit Landbouwwetenschappen Rijksuniversiteit Gent* **58**:763-777

de Ley, P.; van der Velde, M. C.; Mounport, D.; Baujard, P.; Coomans, A. (1995). Ultrastructure of the stoma in Cephalobidae, Panagrolaimidae and Rhaditidae, with a proposal for a revised stoma terminology in Rhabditida (Nematoda). *Nematologica* **41**:153-182

de Man, J. G. (1886). Anatomische Untersuchungen über freilebende Nordsee-Nematoden. Verlag von Paul Frohberg, Leibzig, 82 pp.

de Man, J. G. (1904). Nématodes libres. *In:* Résultats du Voyage du S.Y. Belgica en 1897-1899 : Rapports Scientifiques Zoologie. pp. 1-51

de Queiroz, A. (1993). For consensus (sometimes). *Systematic Biology* **42**(3):368-372

de Queiroz, K. (1985). The ontogenetic method for determining character polarity and its relevance to phylogenetic systematics. *Systematic Zoology* **34**:280-299

deBry, R. W. (2001). Improving interpretation of the decay index for DNA sequence data. *Systematic Biology* **50**(5):742-752

deBry, R. W. (2003). Identifying conflicting signal in a multigene analysis reveals a highly resolved tree: The phylogeny of Rodentia (Mammalia). *Systematic Biology* **52**(5):604-617

Diesing, K. M. (1851). Systema Helmintum. Volume 2. Vindobonae, 588 pp.

Diesing, K. M. (1861). Revision der Nematoden. *Sitzungsberichte der Akademie der Wissenschaften, Wien, Mathematisch-Naturwissenschaftliche Klasse* **42**(28):595-736

Dixon, M. T.; Hillis, D. M. (1993). Ribosomal RNA secondary structure: Compensatory mutations and implications for phylogenetic analysis. *Molecular Biology and Evolution* **10**:256-267

Dolinski, C. M.; Baldwin, J. G. (2003). Fine structure of the stoma of *Bunonema* sp. and *Teratorhabditis palmarum* (Nematoda) and its phylogenetic significance. *Journal of Nematology* **35**(3):244-251

Dolinski, C. M.; Baldwin, J. G.; Thomas, W. K. (2001). Comparative survey of early embryogenesis of Secernentea (Nematoda), with phylogenetic implications. *Canadian Journal of Zoology* **79**:82-94

Dolphin, K.; Belshaw, R.; Orme, C. D. L.; Quicke, D. L. J. (2000). Noise and incongruence: Interpreting results of the incongruence length difference test. *Molecular Phylogenetics and Evolution* **17**(3):401-406

Donoghue, M. J. (1983). The phylogenetic relationships of *Vibrium*. *In:* Advances in Cladistics. Vol. 2 [Proceedings of the 2nd Meeting of the Willi Hennig Society], Platnick, N. I. F.; Funk, V. A. (Eds.), New York, Columbia University Press, pp. 143-166

Donoghue, M. J.; Cantino, P. D. (1984). The logic and limitations of the outgroup substitution approach to cladistic analysis. *Systematic Botany* **9**(2):192-202

Donoghue, M. J.; Doyle, J. A.; Gauthier, J.; Kluge, A. G.; Rowe, T. (1989). The importance of fossils in phylogenetic reconstruction. *Annual Review of Ecology and Systematics* **20**:431-460

Donoghue, M. J.; Sanderson, M. J. (1992). The suitability of molecular and morphological evidence in reconstructing plant phylogeny. *In:* Molecular Systematics of Plants. Soltis, P. S.; Soltis, D. E.; Doyle, J. J. (Eds.), Chapman & Hall, New York, pp. 340-367

Dorchin, N.; Freidberg, A.; Mokady, O. (2004). Phylogeny of the Baldratiina (Diptera: Cecidomyiidae) inferred from morphological, ecological and molecular data sources, and evolutionary patterns in plant-galler relationships. *Molecular Phylogenetics and Evolution* **30**(3):503-515

Dorris, M.; de Ley, P.; Blaxter, M. L. (1999). Molecular analysis of nematode diversity and the evolution of parasitism. *Parasitology Today* **15**(5):188-193

Dorris, M.; Viney, M. E.; Blaxter, M. L. (1999). Molecular phylogenetic analysis of the genus *Strongyloides* and related nematodes. *International Journal for Parasitology* **32**(12):1507-1517

Dougherty, E. C. (1955). The genera and species of the subfamily Rhabditinae Micoletzky, 1922 (Nematoda): A nomenclatorial analysis - including an addendum on the composition of the family Rhabditidae Örley, 1880. *Journal of Helminthology* **29**(3):105-152

Dowton, M.; Austin, A. D. (2002). Increased congruence does not necessarily indicate increased phylogenetic accuracy – the behaviour of the incongruence length difference test in mixed-model analyses. *Systematic Biology* **51**:19-30

Doyle, J. A.; Endress, P. K. (2000). Morphological phylogenetic analysis of basal angiosperms: Comparison and combination with molecular data. *International Journal of Plant Sciences, Supplement* **161**(6):121-153

Doyle, J. J. (1992). Gene trees and species trees: Molecular systematics as one-character taxonomy. *Systematic Botany* **17**:144-163

Drozdovkiy, E. M. (1980). On the division of the class Nematoda into subclasses on the basis of data from comparative anatomy and embryology. *In:* Tezisnye Doklady Konference Vsesoyuznogo Obshchestva Gel'minthologov. [Thesis Reports of the Conference of the All-Union Society of Helminthologists]. Moskva: Nauka, pp. 58-59

Dujardin, F. (1845). Histoire naturelle des helminthes ou vers intestinaux. Paris, 654 pp.

Eberth, C. J. (1863). Untersuchungen über Nematoden. Leibzig, 77 pp.

Efron, B. (1979). Bootstrapping methods: Another look at the jackknife. *Annual Statistics* **7**:1-26

Efron, B. (1982). The jackknife, the bootstrap, and other resampling plans. CBMS-NSF Reginal Conference Series in Applied Mathematics, Society of Industrial and Applied Mathematics, Philadelphia, Monograph **38**:1-92

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association* **82**:171-185

Efron, B.; Gong, G. (1983). A leisurely look at the bootstrap, the jackknife, and the cross-validation. *American Statistics* **37**:36-48

Eldredge, N.; Novacek, M. J. (1985). Systematics and paleontology. *Paleobiology* **11**:65-74

Ellis, R. E.; Sulston, J. E.; Coulson, A. R. (1986). The rDNA of *C. elegans*: Sequence and structure. *Nucleic Acids Research* **14**:2345-2364

Endo, B, Y.; Nickle, W. R. (1994). Ultrastructure of the buccal cavity region and oesophagus of the insect parasitic nematode, *Heterorhabditis bacteriophora*. *Nematologica* **40**:379-398

Endo, B. Y.; Nickle, W. R. (1995). Ultrastructure of anterior and mid-region of infective juveniles of *Steinernema feltiae*. *Fundamental and Applied Nematology* **18**:271-294

Eyre-Walker, A. (1998). Problems with parsimony in sequences of biased base composition. *Journal of Molecular Evolution* **47**:686-690

Faith, D. P. (1991). Cladistic permutation tests for monophyly and nonmonophyly. *Systematic Zoology* **40**(3):366-375

Faith, D. P.; Cranston, P. S. (1991). Could a cladogram this short have arisen by chance alone? - On permutation tests for cladistic structure. *Cladistics* **7**:1-28

Farris, J. S. (1969). A successive approximations approach to character weighting. *Systematic Zoology* **18**(4):374-385

Farris, J. S. (1970). Methods for computing Wagner trees. *Systematic Zoology* **19**:83-92

Farris, J. S. (1971). The hypothesis of nonspecificity and taxonomic congruence. *Annual Review of Ecology and Systematics* **2**:277-302

Farris, J. S. (1972). Estimating phylogenetic trees from distance matrices. *American Naturalist* **106**:645-668

Farris, J. S. (1982). Outgroups and parsimony. *Systematic Zoology* **31**(3):328-334

Farris, J. S. (1983). The logical basis of phylogenetic analysis. *Advances in Cladistics* **2**:1-36

Farris, J. S. (1989). The retention index and rescaled consistency index. *Cladistics* **5**:417-419

Farris, J. S. (2001). Support weighting. *Cladistics* **17**:389-394

Farris, J. S.; Källersjö, M.; Kluge, A. G.; Bult, C. (1994). Testing significance of incongruence. *Cladistics* **10**:315-319

Farris, J. S.; Kluge, A. G., Mickevich, M. F. (1982). Immunological distances and the phylogenetic relationships of the *Rana boylii* species group. *Systematic Zoology* **31**:479-491

Felsenstein, J. (1968). Statistical inference and the estimation of phylogenies. Ph.D. Dissertation, University of Chicago, Chicago

Felsenstein, J. (1978). Case in which parsimony and compatibility methods will be positively misleading. *Systematic Zoology* **27**:401-410

Felsenstein, J. (1981a). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution* **17**:368-376

Felsenstein, J. (1981b). A likelihood approach to character weighting and what it tells us about parsimony and compatibility. *Biological Journal of the Linnaean Society* **16**(3):183-196

Felsenstein, J. (1983). Statistical inference of phylogenies. *Journal of the Royal Statistical Society* **146**:246-272

Felsenstein, J. (1985a). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39**:783-791

Felsenstein, J. (1985b). Confidence limits on phylogenies with a molecular clock. *Systematic Zoology* **34**:152-161

Felsenstein, J. (1993). PHYLIP (Phylogenetic Inference Package, Version 3.5c, Department of Genetics, University of Washington, Seattle

Felsenstein, J.; Kishino, H. (1993). Is there something wrong with the bootstrap on phylogenies? A reply to Hillis and Bull. *Systematic Biology* **42**(2):193-200

Feng, D.; Doolittle, R. F. (1987). Progressive sequencing alignment as a prerequisite to correct phylogenetic trees,. *Journal of Molecular Evolution* **25**:351-360

Ferris, V. R. (1971). Taxonomy of Dorylaimida. *In:* Plant parasitic nematodes. Vol. 1. Zuckerman, B. M.; Mai, F. W.; Rhode, R. A. (Eds.), Academic Press Inc., New York, London, pp. 163-189

Ferris, V. R. (1994). The future of nematode systematics. *Fundamental and Applied Nematology* **17**(2):97-101

Filipjev, I. N. (1918) Nématodes libres marins des environs de Sébastopol. Partie I. [Free-living marine nematodes in the vicinity of Sevastopol. Part I.] *Trudy Osoboi Zoologicheskoi Laboratorii l Sevastopol'skoi Biologicheskoi Stantsii [Travaux du Laboratoire Zoologique et de la Station Biologique de Sébastopol prés l'Academie des Sciences de Russie. Séries 2]* **4**:1-362 [Engl. Translation by Raveh, M. (1968) Israel Program for Scientific Translations, Jerusalem, 255 pp.]

Filipjev, I. N. (1921) Nématodes libres marins des environs de Sébastopol. Partie II. [Free-living marine nematodes in the vicinity of Sevastopol. Part II.] *Trudy Osoboi Zoologicheskoi Laboratorii l Sevastopol'skoi Biologicheskoi Stantsii [Travaux du Laboratoire Zoologique et de la Station Biologique de Sébastopol prés l'Academie des Sciences de Russie. Séries 2]* **4**:353-614 [Engl. Translation by Raveh, M. (1968) Israel Program for Scientific Translations, Jerusalem, 203 pp.]

Filipjev, I. N. (1929). Les nématodes libres de la baie de la Neva et de l'extrémeté orientale du Golfe de Finlande. Première partie. *Archiv für Hydrobiology* **20**:637-699

Filipjev, I. N. (1934). The classification of the free-living nematodes and their relation to the parasitic nematodes. *Smithsonian Miscellaneous Collection* **89**(6):1-63

Fitch, D. H. A. (1997). Evolution of male tail development in rhabditid nematodes related to *Caenorhabditis elegans*. *Systematic Zoology* **46**:145-179

Fitch, W. N.; Margoliash, E. (1967). Construction of phylogenetic trees. *Science* **155**:279-284

Floyd, R.; Abebe, E.; Papert, A., Blaxter, M. (2002). Molecular barcodes for soil nematode identification. *Molecular Ecology* **11**:839-850

Funk, V. A. (1985). Phylogenetic patterns and hybridization. *Annals of the Missouri Botanical Garden* **72**:681-715

Gadéa, E. (1973). Sobre la filogenia interna de los Nematodos. *Publicaciones del Instituto de Biología Aplicada* **54**:87-92

Garnica, S.; Weiss, M.; Oberwinkler, F. (2003). Morphological and molecular phylogenetic studies in South American *Cortinarius* species. *Mycological Research* **107**(10):1143-1156

Gasser, R. A.; Hoste, H.; Chilton, N. B.; Beveridge, I. (1993). Rapid sequencing of rDNA from single worms and eggs of parasitic helminths. *Nucleic Acids Research* **21**(10):2525-2526

Gatesy, J.; DeSalle, R.; Wheeler, W. (1993). Alignment-ambiguous nucleotide sites and the exclusion of systematic data. *Molecular Phylogenetics and Evolution* **2**:152-157

Gauthier, J.; Kluge, A. G.; Rowe, T. (1988). Amniote phylogeny and the importance of fossils. *Cladistics* **4**:105-209

Gelman, A.; Carlin, J. B.; Stern, H. S.; Rubin, D. B. (1995). Bayesian data analysis. Chapman & Hall, New York, 526 pp.

Gerlach, S. A. (1966). Bemerkungen zur Phylogenie der Nematoden. *Mitteilungen aus der Biologischen Bundesanstalt für Land- und Forstwirtschaft Berlin-Dahlem* **118**:25-39

Gerlach, S. A.; Riemann, F. (1973/1974). The Bremerhaven checklist of aquatic nematodes. A catalogue of Nematoda – Adenophorea excluding the Dorylaimida. *Veröffentlichungen des Institutes für Meeresforschung Bremerhaven, Sonderdruck* **4**:1-404 (Part I, 1973), 405-734 (Part II, 1974)

Geyer, C. J. (1991). Markov chain Monte Carlo maximum likelihood. *In:* Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface. Keramidas, E. M. (Ed.), Interface Foundation, Fairfax Station, Virginia, pp. 156-163

Gibbons, L. M. (1986). SEM Guide to the Morphology of Nematode Parasites of Vertebrates. CAB International Publications, Commomwealth Institute of Parasitology, Farnham, 199 pp.

Gilks, W. S.; Richardson, S.; Spiegelhalter, D. (1996). Markov Chain Monte Carlo in Practice. Chapman & Hall, New York

Goldman, N. (1993a). Statistical tests of models of DNA substitution. *Journal of Molecular Evolution* **36**:182-198

Goldman, N. (1993b). Simple diagnostic tests of models of DNA substitution. *Journal of Molecular Evolution* **37**:650-661

Goldman, N.; Anderson, J. P.; Rodrigo, A. G. (2000). Likelihood-based tests of topologies in phylogenetics. *Systematic Biology* **49**(4):652-670

Goloboff, P. A. (1991). Homoplasy and the choice among cladograms. *Cladistics* **7**:215-232

Goloboff, P. A. (1993). Estimating character weights during tree search. *Cladistics* **9**:83-91

Goodey, J. B. (1963a). Soil and freshwater nematodes. 2nd Edition, London, Methuen & Co. Ltd., 544 pp. (original by T. Goodey, 1951, revised and rewritten by J. B. Goodey)

Goodey, J. B. (1963b). Speculations on the identity of the parts of the Tylenchid spear. *Nematologica* **9**:468-470

Goodey, T. (1951). Soil and freshwater nematodes. A monograph. London, Methuen & Co. Ltd., 390 pp.

Gotoh, O. (1982). An improved algorithm for matching biological sequences. *Journal of Molecular Biology* **162**:705-708

Gotoh, O. (1996). Significant improvement in accuracy of multiple protein sequence alignments by iterative refinement as assessed by reference to structural alignments. *Journal of Molecular Biology* **264**:823-838

Gotoh, O. (1999). Multiple sequence alignment: Algorithms and applications. *Advances in Biophysics* **36**:159-206

Graybeal, A. (1998). Is it better to add taxa or characters to a difficult phylogenetic problem? *Systematic Biology* **47**(1):9-17

Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82**:711-732

Gu, X.; Fu, Y.-X.; Li, W.-H. (1995). Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. *Molecular Biology and Evolution* **12**:546-557

Halanych, K. M. (1991). 5S ribosomal RNA sequences inappropriate for phylogenetic reconstruction. *Molecular Biology and Evolution* **8**(2):249-253

Hall, T. (2001). BioEdit. Version 5.0.6 [software program], available from: www.mbio.ncsu.edu/BioEdit/bioedit.html

Hammond, P. M. (1992). Species inventory. *In:* Global Diversity, Status of the Earth's Living Resources. Chapman & Hall, London, pp. 17-39

Hancock, J. M.; Vogler, A. P. (2000). How slippage-derived sequences are incorporated into rRNA variable-region secondary structure: Implications for phylogeny reconstructions. *Molecular Phylogeny and Evolution* **14**(3):366-374

Harshman, J. (1994). The effect of irrelevant characters on bootstrap values. *Systematic Biology* **43**:419-424

Hasegawa, M.; Kishino, H.; Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22**:160-174

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their application. *Biometrika* **57**:97-109

Hein, J. (1989a). A new method that simultaneously aligns and reconstructs ancestral sequences for any number of homologous sequences, when a phylogeny is given. *Molecular Biology and Evolution* **6**:649-668

Hein, J. (1989b). A tree reconstruction method that is economical in the number of pairwise comparisons used. *Molecular Biology and Evolution* **6**:669-684

Hempel, C. G. (1965). Aspects of Scientific Explanation and other Essays in Philosophy of Science. Free Press, Collier-Macmillian, New York, 505 pp.

Hendy, M. D.; Penny, D. (1982). Branch and bound algorithms to determine minimal evolutionary trees. *Mathematical Biosciences* **59**:277-290

Hendy, M. D.; Penny, D. (1989). A framework for the quantitative study of evolutionary trees. *Systematic Zoology* **38**(4):297-309

Hennig, W. (1966). Phylogenetic Systematics. University of Illinois Press, Urbana, 263 pp.

Hickson, R. E.; Simon, C.; Cooper, A.; Spicer, G. S.; Sullivan, J.; Penny, D. (1996). Conserved sequence motifs, alignment, and secondary structure for the third domain of animal 12S rRNA. *Molecular Biology and Evolution* **13**(1):150-169

Hickson, R. E.; Simon, C.; Perrey, S. W. (2000). The performance of several multiple-sequence alignment programs in relation to secondary-structure features for an rRNA sequence. *Molecular Biology and Evolution* **17**(4):530-539

Hillis, D. M. (1987). Molecular versus morphological approaches to systematics. *Annual Review of Ecology and Systematics* **18**:23-42

Hillis, D. M. (1991). Discriminating between phylogenetic signal and random noise in DNA sequences. *In:* Phylogenetic Analysis of DNA sequences. Miyamoto, M. M.; Cracraft, J. (Eds.), Oxford University Press, Oxford, pp. 278-294

Hillis, D. M. (1998) Taxonomic sampling, phylogenetic accuracy, and investigator bias. *Systematic Biology* **47**(1):3-8

Hillis, D. M.; Allard, M. W.; Miyamoto, M. M. (1993). Analysis of DNA sequence data: Phylogenetic inference. *Methods in Enzymology* **224**:456-487

Hillis, D. M.; Bull, J. J. (1993). An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology* **42**(2):182-192

Hillis, D. M.; Dixon, M. T. (1991). Ribosomal DNA: Molecular evolution and phylogenetic ionference. *Quaterly Review of Biology* **66**(4):411-453

Hillis, D. M.; Huelsenbeck, J. P.; Swofford, D. L. (1994). Hobglobin of phylogenetics? *Nature* **369**:363-364

Hillis, D. M.; Mable, B. K.; Moritz, C. (1996). Application of molecular systematics: The state of the field and a look to the future. *In:* Molecular Systematics. 2nd Edition, Hillis, D. M.; Moritz, C.; Mable, B. K. (Eds.), Sinauer Associates, Sunderland, Massachusetts, pp. 515-543

Hoberg, E. P.; Monsen, K. J.; Kutz, S.; Blouin, M. S. (1999). Structure, biodiversity, and historical biogeography of nematode faunas in Holarctic ruminants: Morphological and molecular diagnoses for *Teladorsagia boreoarcticus* n. sp. (N.: Ostertagiinae), a dimorphic cryptic species in muskoxen (*Ovibos moschatus*). *Journal of Parasitology* **85**(5):910-934

Hooper, D. J. (1973). Nematodes. *In:* Viruses and Invertebrates. Gibbs, A. J. (Ed.), North Holland Publishing Co., Amsterdam, pp. 212-228

Hope, W. D. (1988). Ultrastructure of the feeding apparatus of *Rhabdodermania minima* Chitwood 1936 (Enoplida, Rhabdodermaniidae). *Journal of Nematology* **20**(1):118-140

Huelsenbeck, J. P. (1991). Tree-length skewness: An indicator of phylogenetic information. *Systematic Zoology* **40**:257-270

Huelsenbeck, J. P. (1995). Performance of phylogenetic methods in simulation. *Systematic Biology* **44**(1):17-48

Huelsenbeck, J. P., Hillis, D. M.; Nielsen, R. (1996). A likelihood-ratio test of monophyly. *Systematic Biology* **45**(5):546-558

Huelsenbeck, J. P.; Bull, J. J. (1996). A likelihood ratio test to detect conflicting phylogenetic signal. *Systematic Biology* **45**(1):92-98

Huelsenbeck, J. P.; Larget, B.; Miller, R. E.; Ronquist, F. (2002). Potential applications and pitfalls of Bayesian inference of phylogeny. *Systematic Biology* **51**(5):673-688

Huelsenbeck, J. P.; Ronquist, F. (2001). MrBayes: Bayesian inference of phylogeny. *Bioinformatics* **17**:754-755

Huelsenbeck, J. P.; Ronquist, F.; Nielsen, R.; Bollback, J. P. (2001). Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* **294**:2310-2314

Huelsenbeck, J. P.; Swofford, D. L.; Cunningham, C. W.; Bull, J. J.; Waddell, P. J. (1994). Is character weighting a panaces for the problem of data heterogeneity in phylogenetic anaysis? *Systematic Biology* **43**(2):288-291

Hugot, J.-P.; Baujard, P.; Morand, S. (2201). Biodiversity in helminths and nematodes as a field of study: An overview. *Nematology* **3**(3):199-208

Hunt, D. J. (1993). Aphelenchida, Longidoridae and Trichodoridae. Their Systematic and Bionomics. CAB International Publishing, 352 pp.

Hussey, R. S. (1979). Biochemical systematics of nematodes – a review. *Helminthological Abstracts, Series B* **48**:141-148

Hyman, B. C. (1988). Nematode mitochondrial DNA: Anomalies and applications. *Journal of Nematology* **20**:523-531

Hyman, B. C. (1996). Molecular systematics and population biology of phytonematodes: Some unifying principles. *Fundamental and Applied Nematology* **19**(4):309-313

Hyman, L. H. (1951). The Invertebrates. Vol. III: Acanthocephala, Aschelminthes and Entoprocta - the Pseudocoelomate Bilateria. McGraw-Hill, New York, 572 pp.

Inglis, W. G. (1974). The marine Enoplida (Nematoda): A comparative study of the head. *Bulletin of the British Museum (Natural History) (Zoology)* **11**:265-376

Inglis, W. G. (1983). An outline classification of the phylum nematoda. *Australian Journal of Zoology* **31**:243-255

Jairajpuri, M. S. (1964) Studies on Campydoridae and Leptonchidae (Nematoda: Dorylaimoidea) with description of *Basirotyleptus basiri* n. gen., n. sp., from India. *Proceedings of the Helminthological Society of Washington* **31**:59-64

Jairajpuri, M. S. (1969) Studies on Mononchida of India. I. The genera *Hadronchus*, *Iotonchus* and *Miconchus*, and a revised classification of Mononchida, new order. *Nematologica* **15**:557-581

Jairajpuri, M. S. (1983) Observations on *Campydora* (Nematoda: Dorylaimida). *Nematologia Mediterranea* **11**:33-42

Jairajpuri, M. S.; Ahmad, S.; Bajaj, H. K. (1976). On the morphology and systematic position of *Campydora* Cobb, 1920 (Nematoda: Dorylaimida). *First National Convention of Indian Helminthology*, pp. 25-26 (Abstract)

Jairajpuri, M. S.; Ahmad, W. (1992). Dorylaimida: Free-living, Predaceous and Plant-Parasitic Nematodes. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi, 458 pp.

Jardine, N. (1969). A logical basis for biological classification. *Systematic Zoology* **18**:37-52

Jayasree, K.; Warwick, R. M. (1977). Free-living marine nematodes of a polluted sandy beach in the firth of Clyde, Scotland - description of seven new species. *Journal of Natural History* **11**:289-302

Kaiser, H. (1977). Untersuchungen zur Morphologie, Biometrie, Biologie und Systematik von Mermithiden. Ein Beitrag zum Problem der Trennung morphologisch schwer unterscheidbarer Arten. *Zoologische Jahrbücher (Systematik)* **194**:20-71

Källersjö, M.; Albert, V. A.; Farris, J. S. (1999). Homoplasy increases phylogenetic structure. *Cladistics* **15**:91-94

Källersjö, M.; Farris, J. S.; Kluge, A. G.; Bult, C. (1992). Skewness and permutation. *Cladistics* **8**:275-287

Kampfer, S.; Sturmbauer, C.; Ott, J. (1998). Phylogenetic analysis of rDNA sequences from adenophorean nematodes and implications for the Adenophorea-Secernentea controversy. *Invertebrate Biology* **117**:29-36

Kass, R. E.; Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association* **90**(430):773-795

Kidd, K. K.; Astolfi, P.; Cavalli-Sforza, L. L. (1974). Error in the reconstruction of evolutionary trees. *In:* Genetic Distance. Crow, J. F.; Denniston, C. (Eds.), Plenum, New York, pp. 121-136

Kidd, K. K.; Sgaramella-Zonta, L. A. (1971). Phylogenetic analysis: Concepts and methods. *American Journal of Human Genetics* **23**:235-252

Kim, J. (1998). Large-scale phylogenies and measuring the performance of phylogenetic estimation. *Systematic Biology* **47**(1):43-60

Kishino, H.; Hasegawa, M. (1989). Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *Journal of Molecular Evolution* **29**:170-179

Kitching, I, J.; Forey, P. L.; Humphries, C. J.; Williams, D. M. (1998). Cladistics, the Theory and Practice of Parsimony Analysis. 2nd Edition, The Systematics Association Publication No. 11, Oxford University Press, Oxford, 228 pp.

Kjer, K. M. (1995). Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: An example of alignment and data presentation from the frogs. *Molecular Phylogenetics and Evolution* **4**(3):314-330

Kluge, A. G. (1984). The relevance of parsimony to phylogenetic inference. *In:* Cladistics: Perspectives on the Reconstruction of Evolutionary History. Duncan, T.; Stuessy, T. (Eds.), Columbia University Press, New York, pp. 24-38

Kluge, A. G. (1985). Ontogeny and phylogenetic systematics. *Cladistics* **1**:13-27

Kluge, A. G. (1989). A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Systematic Zoology* **38**(1):7-25

Kluge, A. G.; Farris, J. S. (1969). Quantitative phyletics and the evolution of anurans. *Systematic Zoology* **18**(1):1-32

Kluge, A. G.; Wolf, A. G. (1993). Cladistics: What's in a word? *Cladistics* **9**:183-199

Kuhner, M. K.; Felsenstein, J. (1994). A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. *Molecular Biology and Evolution* **11**(3):459-468

Kumar, S.; Tamura, K.; Nei, M. (1993). MEGA: Molecular Evolutionary Genetics Analysis. Version 1.0, Pennsylvania State University, University Park, Pennsylvania

Lake, J. A. (1994). Reconstructing evolutionary trees from DNA and protein sequences: Paralinear distances. *Proceedings of the National Academy of Sciences* **91**:1455-1459

Lamboy, W. F. (1994). The accuracy of maximum parsimony method for phylogeny reconstruction with morphological characters. *Systematic Botany* **19**(4):489-505

Lamboy, W. F. (1996). Morphological characters, polytomies, and homoplasy indeces: A response to Wiens and Hillis. *Systematic Botany* **21**(2):243-253

Lambshead, P. L. D. (1993). Recent developments in marine benthic biodiversity research. Océanis **19**(6):5-24

Lanyon, S. M. (1985). Detecting internal inconsistencies in distance data. *Systematic Zoology* **34**:397-403

Larget, B.; Simon, D. L. (1999). Markov Chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. *Molecular Biology and Evolution* **16**(6):750-759

Lassmann, T.; Sonnhammer, E. L. L. (2002). Quality assessment of multiple alignment programs. *FEBS Letters* **529**:126-130

Lee, C.; Grasso, C.; Sharlow, M. F. (2002). Multiple sequence alignment using partial order graphs. *Bioinformatics* **18**(3):452-464

Lee, M. S. Y. (2001). Unalignable sequences and molecular evolution. *Trends in Ecology and Evolution* **16**(12):681-685

Leidig, F (1854). *Zoologisches Archiv für Anatomie und Physiologie*, pp. 284-295

Li, S. (1996). Phylogenetic tree construction using Markov chain Monte Carlo. Ph.D. Dissertation, Ohio State University, Columbus

Lieven, A. F. v. (2002) The sister group of the Diplogasterina (Nematoda). *Russian Journal of Nematology* **10**(2):127-137

Linnaeus, C. (1767) Systema naturae per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentitii, synonymis, locis. Editio duodecima, reformata, Holmiae, **1**(2):533-1327

Linstow, O. von (1905). Neue Helminthen. *Archive für Naturgeschichte* **71**(1):267-276

Littlewood, D. T. J.; Smith, A. B. (1995). A combined morphological and molecular phylogeny for sea urchins (Echinoidea: Echinodermata). *Philosophical Transactions of the Royal Society of London Series B – Biological Sciences* **347**:213-234

Litvaitis, M. K.; Bates, J. W.; Hope, W. D.; Moens, T. (2000). Inferring a classification of the Adenophorea (Nematoda) from nucleotide sequences of the D3 expansion segment (26/28S rDNA). *Canadian Journal of Zoology* **78**:911-922

Lockhard, P. J.; Steel, M. A., Hendy, M. D.; Penny, D. (1994). Recovering evolutionary trees under a more realistic model of sequence evolution. *Molecular Biology and Evolution* **11**(4):605-612

Lorenzen, S. (1977). Revision der Xyalidae (freilebende Nematoden) auf der Grundlage einer kritischen Analyse von 56 Arten aus Nord- und Ostsee. *Veröffentlichungen des Institutes für Meeresforschung, Bremerhaven* **16**:197-261

Lorenzen, S. (1978). Discovery of stretch receptor organs in nematodes - structure, arrangement and functional analysis. *Zoologica Scripta* **7**:175-178

Lorenzen, S. (1981). Entwurf eines phylogenetischen Systems der freilebenden Nematoden. *Veröffentlichungen des Institutes für Meeresforschung Bremerhaven, Sonderdruck* **7**:1-472

Lorenzen, S. (1983). Phylogenetic Systematics: Problems, Achievements and its Application to the Nematoda. *In:* Concepts of Nematode Systematics. Stone, A. R.; Platt, H. M.; Khalil, L. F. (Eds.), Academic Press, London, pp. 11-23

Lorenzen, S. (1994). The Phylogenetic Systematics of Freeliving Nematodes. Ray Society, London, 383 pp.

Lundberg, J. G. (1972). Wagner networks and ancestors. *Systematic Zoology* **21**:398-413

Lyons-Weiler, J.; Hoelzer, G. A.; Tausch, R. J. (1998). Optimal outgroup analysis. *Biological Journal of the Linnaean Society* **64**:493-511

Maddison, W. P. (1989). Reconstructing character evolution on polytomous cladograms. *Cladistics* **5**:365-377

Maddison, W. P. (1990). A method for testing the correlated evolution of two binary characters: Are gains or loses concentrated on certain branches of a phylogenetic tree? *Evolution* **44**:539-557

Maddison, W. P. (1993). Missing data versus missing characters in phylogenetic analysis. *Systematic Biology* **42**:576-581

Maddison, W. P.; Donoghue, M. J.; Maddison, D. R. (1984). Outgroup analysis and parsimony. *Systematic Zoology* **33**(1):83-103

Maddison, W. P.; Maddison, D. R. (1992). MacClade: Analysis of Phylogeny and Character Evolution. Version 3.0, Sinauer Associates, Sunderland, Massachusetts, [software program]

Maggenti, A. R. (1963). Comparative morphology on nemic phylogeny. *In:* The lower Metazoan. Dougherty; Brown; Hanson; Hartmann; (Eds.), University of California Press, Berkeley, pp. 273-282

Maggenti, A. R. (1970). System analysis and nematode phylogeny. *Journal of Nematology* **2**:7-15

Maggenti, A. R. (1982). Nemata. *In:* Synopsis and Classification of Living Organisms. Vol. 1, Parker, S. P. (Ed.), McGraw-Hill Book Co., New York, pp. 879-929

Maggenti, A. R. (1983). Nematode higher classification as influenced by species and families concepts. *In:* Concepts of Nematode Systematics. Stone, A. R.; Platt, H. M.; Khalil, L. F. (Eds.), Academic Press, London [388 pp.]

Maggenti, A. R.; Luc, M.; Raski, D. J.; Fortuner, R.; Geraert, E. (1987). A reappraisal of Tylenchina (Nemata). 2. Classification of the suborder Tylenchina (Nemata: Diplogasteria). *Revue de Nématologie* **10**(2):135-142

Malakhov, V. V. (1974). The life cycle of the free-living marine nematode *Pontonema vulgare* in the White Sea. *Vyestnik Moskovskogo Universiteta* **1**:10-13

Malakhov, V. V. (1977). Postembryonic development of the free-living marine nematode *Pontonema vulgare* Bastian. *Biologia Belogo Marya* **4**:148-153

Malakhov, V. V. (1994). Nematodes - Structure, Development, Classification, and Phylogeny. Hope, D. (Ed.), Smithsonian Institution Press, Washington and London, 286 pp. [translated by G. V. Bentz]

Malakhov, V. V. (1998). Embryological and histological peculiarities of the order Enoplida, a primitive group of nematodes. *Russian Journal of Nematology* **6**(1):41-46

Malakhov, V. V.; Akimushkiva, M. I. (1976). The embryonic development of the free-living marine nematode *Enoplus brevis*. *Zoologicheskii Zhurnal* **55**(12):1788-1798

Malakhov, V. V.; Cherdantsev, V. G. (1975). The embryonic development of a free-living nematode *Pontonema vulgare*. *Zoologicheskii Zhurnal* **54**(2):165-174

Malakhov, V. V.; Cherdantsev, V. G. (1977). Early stages of embryonic development of the free-living marine nematode *Anoplostoma vivipara*. *Biologiya Beogo Morya* **4**:154-161

Margush, T.; McMorris, F. R. (1981). Consensus n-trees. *Bulletin of Mathematical Biology* **43**:239-244

Marshall, C. (1992). Substitution bias, weighted parsimony, and aminote phylogeny as inferred from 18S rRNA sequences. *Molecular Biology and Evolution* **9**:370-373

Mau, B. (1996). Bayesian phylogenetic inference via Markov chain Monte Carlo methods. Ph.D. Dissertation, University of Wisconsin, Madison

May, R. M. (1988). How many species are there on Earth? *Science* **241**:1441-1449

Mayden, R., L.; Wiley, E. O. (1992). The fundamentals of phylogenetic systematics. *In:* Systematics, Historical Ecology, and North American Fishes. Mayden, R, L. (Ed.), Stanford University Press, Stanford, pp. 114-185

Mayr, E. (1969). Principles of Systematic Zoology. McGraw-Hill, New York, 428 pp.

McCarthy, C. (1997). Chromas. Grffith University, Brisbane [software program]

McDowell, T.; Bremer, B. (1998). Phylogeny, diversity, and distribution in *Exostema* (Rubiaceae): Implications of morphological and molecular analyses. *Plant Systematics and Evolution* **212**(3-4):215-246

Meier, R.; Wiegmann, B. M. (2002). A phylogenetic analysis of Coelopidae (Diptera) based on morphological and DNA sequence data. *Molecular Phylogenetics and Evolution* **25**(3):393-407

Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, H.; Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics* **21**:1087-1091

Mickevich, M. F. (1978). Taxonomic congruence. *Systematic Biology* **27**:143-158

Mickevich, M. F.; Farris, J. S. (1981). The implications of congruence in Menidia. *Systematic Zoology* **30**(3):351-370

Mickevich, M. F.; Weller, S. J. (1990). Evolutionary character analysis: Tracing character change on a cladogram. *Cladistics* **6**(2):137-170

Micoletzky, H. (1922). Die freilebenden Erd-Nematoden. *Archiv für Naturgeschichte* **87A**(8/9):1-650

Micoletzky, H. (1925). Die freilebenden Süßwasser- und Moornematoden Dänemarks. nebst Anhang über Amöbosporidien und andere Parasiten bei freilebenden Nematoden. *Kongelige Danske Videnskabernes Selskabs Skrifter, Naturvidenskabernes og Mathematikk Afd. 8* **10**(2):55-310

Milinkovitch, M. C.; LeDuc, R. G.; Adachi, J.; Farnir, F.; Georges, M.; Hasegawa, M. (1996). Effects of character weighting and species sampling on phylogeny

reconstruction: A case study based on DNA sequence data in cetaceans. *Genetics* **144**:1817-1833

Milinkovitch, M. C.; Lyons-Weiler, J. (1998). Finding optimal ingroup topologies and convexities when the choice outgroups is not obvious. *Molecular Phylogeny and Evolution* **9**(3):348-357

Mindell, D. P.; Honeycutt, R. L. (1990). Ribosomal RNA in vertebrates: Evolution and phylogenetic implications. *Annual Review of Ecology and Systematics* **21**:541-566

Miyamoto, M. M. (1983). Biochemical variation in *Eleutherodactylus bransfordii*: Geographic patterns and cryptic species. *Systematic Zoology* **321**:43-51

Miyamoto, M. M. (1985). Consensus cladograms and general classification. *Cladistics* **1**:186-189

Miyamoto, M. M.; Boyle, S. M. (1989). The potential importance of mitochondrial DNA sequence data to eutherian mammal phylogeny. *In:* The Hierarchy of Life. Fernholm, B.; Bremer, K.; *Jörnvall*, H. (Eds.), Elsevier Science, Amsterdam, pp. 437-450

Miyamoto, M. M.; Fitch, W. M. (1995). Testing species phylogenies and phylogenetic methods with congruence. *Systematic Biology* **44**(1):64-76

Morgenstern, B. (1999). DIALIGN 2: Improvement of the segment-to-segment approach to multiple sequence alignment. *Bioinformatics* **15**(3):211-218

Morrison, D. A.; Ellis, J. T. (1997). Effects of nucleotide sequence alignment on phylogeny estimation: A case study of 18S rDNAs of Apicomplexa. *Molecular Biology and Evolution* **14**(4):428-441

Mortensen, M. L.; Rostrup, S.; Ravn, K. (1908). Oversigt over Landbrugsplanternes Syngdomme i 1907. *Tidsskr. Bandbr. Planteavl.* **15**:145-158

Müller, O. F. (1786). Animalcula infusoria fluviatilla et marina, etc. Havniae, 367 pp.

Mullin, P, G..; Szalanski, A. L.; Harris, T. S.; Powers, T. O. (2001). Status of *Campydora* (Nematoda: Campydorina). *Journal of Nematology* **33**(4):268 (Abstract)

Mullis, K. B.; Faloona, F. A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology* **155**:335-350

Nadler, S. A. (1990). Molecular approaches to studying helminth population genetics and phylogeny. *International Journal for Parasitology* **20**(1):11-29

Nadler, S. A. (1992). Phylogeny of some Ascaridoid nematodes, inferred from comparison of 18S and 26S rRNA sequences. *Molecular Biology and Evolution* **9**:932-944

Nadler, S. A.; Hoberg, E. P.; Hudspeth, D. S. S.; Rickard, L. G. (2000). Relationships of *Nematodirus* species and *Nematodirus battus* isolates (Nematoda: Trichostrongyloidea) based on nuclear ribosomal DNA sequences. *Journal of Parasitology* **86**(3):58-61

Nadler, S. A.; Huspeth, D. S. S. (1998). Ribosomal DNA and phylogeny of the Ascaridoidea (Nemata: Secernentea): Implications for morphological evolution and classification. *Molecular Phylogenetics and Evolution* **10**(2):221-236

Needham, F. (1745). An account of some new microscopical discoveries. London, 126 pp.

Needleman, S. B.; Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequences of two proteins. *Journal of Molecular Biology* **48**:443-453

Neff, N. (1986). A rational basis for a prior character weighting. *Systematic Zoology* **35**:110-123

Nelson, G. J. (1979). Cladistic analysis and synthesis: Principles and definitions, with a historical note on Adanson's Famille des Plantes (1763-1764). *Systematic Zoology* **28**:1-21

Nelson, G. J. (1978). Ontogeny, phylogeny, paleontology, and the biogenic low. *Systematic Zoology* **27**:324-345

Nelson, G. J. (1985). Outgroups and ontogeny. *Cladistics* **1**:29-45

Nicholas, W. L.; Stewart, A. C. (1997). Ultrastructure of *Gonionchus australis* (Xyalidae, Nematoda). *Journal of Nematology* **29**(2):133-143

Nickle, W. R.; Hooper, D. J. (1991). The Aphelenchina: Bud, leaf and insect nematodes. *In:* Manual of Agricultural Helminthology. Nickle, W. R. (Ed.), Marcel Dekker, New York, pp. 465-507

Nixon, K. C.; Carpenter, J. M. (1993). On outgoups. *Cladistics* **8**:413-426

Nixon, K. C.; Carpenter, J. M. (1996). On simultaneous analysis. *Cladistics* **12**:221-241

Nixon, K. C.; Wheeler, Q. D. (1992). Extinction and the origin of species. *In:* Extinction and Phylogeny. Wheeler, Q. D.; Novacek, M. (Eds.), Columbia University Press, New York, pp. 119-143

Noller, H. F. (1984). Structure of ribosomal RNA. *Annual Review of Biochemistry* **53**:119-162

Notredame, C.; Higgins, D. G.; Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* **302**(1):205-217

O'Brien, E. A.; Notredame, C.; Higgins, D. G. (1998). Optimisation of ribosomal RNA profile alignments. *Bioinformatics* **14**(4):332-341

Oken, L. (1815). Lehrbuch der Naturgeschichte. Leibzig, **3**(1):1-842

Okimoto, R.; Chamberlin, H. M.; Macfarlane, J. L.; Wolstonholme, D. R. (1991). Repeated sequence sets in mitochondrial DNA molecules of root knot nematode (*Meloidogyne*): Nucleotide sequences, genome location and potential host-race identification. *Nucleic Acids Research* **19**(7):1619-1626

Okimoto, R.; Macfarlane, J. L.; Clary, D. O.; Wolstenholme, D. R. (1992). The mitochondrial genome of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*. *Genetics* **130**:471-498

Örley, L. (1880). Az anguillulidák magánrajza. (Monographie der Anguilluliden). *Természetrajzi Füzetek* **4**(1-2):16-150

Örley, L. (1880). Die Rhabditiden und ihre medizinische Bedeutung. Friedländer, R., Berlin, 84 pp.

Osche, G. (1952). Systematik und Phylogenie der Gattung *Rhabditis* (Nematoda). *Zoologische Jahrbücher (Systematik)* **81**(3):190-280

Page, R. D. M. (1996). TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Application in the Biosciences* **12**:357-358

Page, R. D. M. (2000). Comparative analysis of secondary structure of insect mitochondrial small subunit ribosomal RNA using maximum weighted matching. *Nucleic Acids Research* **28**(20):3839-3845

Paramonov, A. A. (1968). Plant-parasitic nematodes. Vol. I. Origin of Nematodes. Ecological and Morphological Characteristics of Plant Nematodes. Jerusalem: Israel Program for Scientific Translation, 390 pp.

Patterson, C. (1982). Morphological characters and homology. *In:* Problems in Phylogenetic Reconstruction. Joysey, K. A. F.; Friday, A. E. (Eds.), Academic Press, London, pp. 21-74

Patterson, C. (1983). How does phylogeny differ from ontogeny? *In:* Development and Evolution. Goodwin, B. C.; Holder, N.; Wylie, C. C. (Eds.), Cambridge University Press, Cambridge, pp. 1-31

Patterson, C. (1987). Molecules and Morphology in Evolution: Conflict of Compromise? Cambridge University Press, Cambridge, 229 pp.

Patterson, C.; Williams, D. M.; Humphries, C. J. (1993). Congruence between molecular and morphological phylogenies. *Annual Review of Ecology and Systematics* **24**:153-188

Penny, D. (1982). Towards a basis for classification: The incompleteness of distance measures, incompatibility analysis and phenetic classification. *Journal of theoretical Biology* **96**:129-142

Penny, D.; Foulds, L. R.; Hendy, M. D. (1982). Testing the theory of evolution by comparing evolutionary trees constructed from five different protein sequences. *Nature* **297**:197-200

Penny, D.; Hendy, M. D. (1986). Estimating the reliability of evolutionary trees. *Molecular Biology and Evolution* **3**:403-417

Perez-Losada, M.; Hoeg, J. T.; Kolbasov, G. A.; Crandall, K. A. (2002). Reanalysis of the relationships among the Cirripedia and the Ascothoracida and the phylogenetic position of the Facetotecta (Maxillopoda : Thecostraca) using 18S rDNA sequences. *Journal of Crustacean Biology* **22**(3):661-669

Pesole, G.; Sbisa, E.; Mignotte, F.; Saccone, C. (1991). The branching order of mammals: Phylogenetic trees inferred from nuclear and mitochondrial molecular data. *Journal of Molecular Evolution* **33**:537-542

Peterson, K. J.; Eernisse, D. J. (2001). Animal phylogeny and the ancestry of bilaterians: Inference from morphology and 18S rDNA gene sequences. *Evolution and Development* **3**(3):170-205

Phillips, A.; Janies, D.; Wheeler, W. (2000). Multiple sequence alignment in phylogenetic analysis. *Molecular Phylogenetics and Evolution* **16**(3):317-330

Platnik, N. I.; Griswold, C. E.; Coddington, J. A. (1991). On missing entries in cladistic analysis. *Cladistics* **7**:337-343

Platt, H. M. (1983). The free-living marine nematode genus *Sabatieria* (Nematoda: Comesomatidae). *Bulletin of the British Museum (Natural History) (Zoology)* **45**(5):239-252

Platt, H. M. (1985). The free-living marine nematode genus *Sabatieria* (Nematoda: Comesomatidae). Taxonomic revision and pictorial keys. *Zoological Journal of the Linnean Society* **83**:27-78

Platt, H. M.; Warwick, R. M. (1983). Freeliving Marine Nematodes: Part I British Enoplids. Pictoral key to world genera and notes for the identification of British species. Synopses of the British Fauna (New Series) No. 28, Cambridge University Press, 307 pp.

Platt, H. M.; Warwick, R. M. (1988). Freeliving Marine Nematodes: Part II British Chromadorids. Pictoral key to world genera and notes for the identification of British species. Synopses of the British Fauna (New Series) No. 38, Brill, E. J.; Backhuys, Dr. W. (Eds.), 502 pp.

Pleijel, F. (1995). On character coding for phylogeny reconstruction. *Cladistics* **11**:309-315

Poe, S.; Swofford, D. L. (1999). Taxon sampling revisited. *Nature* **398**:299-300

Poinar, G. O. Jr. (1985). Fossil evidence of nematode (Tylenchida) parasitism in Staphylinidae (Coleoptera). *Nematologica* **31**:353-357

Poinar, G. O. Jr. (2003). Trends in the evolution of insect parasitism by nematodes as inferred from fossil evidence. *Journal of Nematology* **35**(2):129-132

Posada, D.; Crandall, K. A. (1998). MODELTEST: Testing the model of DNA substitution. *Bioinformatics Applications Notes* **14**(9):817-818

Powers, T. O.; Harris, T. S.; Hyman, B. C. (1993). Mitochondrial DNA sequence divergence among *Meloidogyne incognita*, *Romanomermis culicivorax*, *Ascaris suum*, and *Caenorhabditis elegans*. *Journal of Nematology* **25**(4):564-572

Powers, T. O.; Todd, T. C.; Burnell, A. M.; Murray, P. C. B.; Fleming, C. C.; Szalanski, A. L.; Adams, B. A.; Harris, T. S. (1997). The rDNA internal transcibed spacer region as a taxonomic marker for nematodes. *Journal of Nematology* **29**(4):441-450

Radford, E. A.; Watson, M. F.; Preston, J. (2001). Phylogenetic relationships of species of *Aciphylla* (Apiaceae, subfamily Apioideae) and related genera using molecular, morphological, and combined data sets. *New Zealand Journal of Botany* **39**(2):183-208

Rannala, B.; Yang, Z. (1996). Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. *Journal of Molecular Evolution* **43**:304-311

Reeves, J. H. (1992). Heterogeneity in the substitution process of amino acid sites of proteins coded for by mitochondrial DNA. *Journal of Molecular Evolution* **35**:17-31

Richter, S. (1971). Zum Feinbau von Mermithiden (Nematoda), I. Der Bohrapparat der vorparasitischen Larven von *Hydromermis contorta* (Linstow, 1889) Hagmeier, 1912. *Zeitschrift für Parasitenkunde* **36**:32-50

Riemann, F. (1966). Die interstitielle Fauna im Elbe-Aestuar, Verbreitung und Systematik. *Archiv für Hydrobiologie, Sonderdruck* **8**(1-2):1-279

Rodrigo, A. G.; Kelly-Borges, M.; Berquist, P. R.; Berquist, P. L. (1993). A randomisation test of the null hypothesis that two cladograms are sample estimates of a parametric phylogenetic tree. *New Zealand Journal of Botany* **31**:257-268

Rudolphi, C. A. (1808). Entozoorum sive vermium intestinalium historia naturalis. Vol. 1. Amstelaedami, 527 pp.

Rusin, L. Y.; Aleshin, V. V.; Tchesunov, A. V.; Atrashkevich, G. I. (2003). The 18S ribosomal RNA gene of *Soboliphyme baturini* Petrow, 1930 (Nematoda: Dioctophymida) and its implications for phylogenetic relationships within Dorylaimia. *Nematology* **5**:615-628

Rzhetsky, A.; Nei, M. (1992). A simple method for estimating and testing minimum-evolution trees. *Molecular Biology and Evolution* **9**:945-967

Rzhetsky, A.; Nei, M. (1995). Tests of applicability of several substitution models for DNA sequence data. *Molecular Biology and Evolution* **12**(1):131-151

Sachs, H. (1950). Die Nematodenfauna der Rinderextremente. Eine ökologisch-systematische Studie. *Zoologische Jahrbücher (Systematik)* **79**:209-272

Saitou, N.; Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**:406-425 1987

Sanderson, M. J. (1989). Confidence limits on phylogenies: The bootstrap revisited. *Cladistics* **5**:113-129

Schneider, A. (1866). Monographie der Nematoden. Druck und Verlag von Georg Reimer, Berlin, 357 pp.

Schuh, R. T.; Polhemus, J. T. (1981). Analysis of taxonomic congruence among morphological, ecological, and biogeographical data sets for the Leptopodomorpha (Hemiptera). *Systematic Zoology* **29**:1-26

Schuurmans Stekhoven Jr., J. H. (1935). Nematoda: Systematischer Teil: Nematoda errantia. *In:* Die Tierwelt der Nord- und Ostsee. Gimpe & Wagler (Ed.s), Leipzig, **5b**:1-173

Shaffer, H. B.; Clark, J. M.; Kraus, F. (1991). When molecules and morphology clash: A phylogenetic analysis of the North American ambystomatid salamanders (Caudata: Ambystomatidae). *Systematic Biology* **40**:284-303

Sharkey, M. (1994). Discriminate compatibility measures and the reduction routine. *Systematic Biology* **43**:526-542

Shimodaira, H.; Hasegawa, M. (1999). Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution* **16**(8):1114-1116

Shoshani, J.; Golenberg, E. M.; Yang, H. (1998). Elephantidae phylogeny: Morphological versus molecular results. *Acta Theriologica* **5**:89-122

Siddiqi, M. R. (1961). On *Xiphinema opisthohysterum* n. sp. and *X. pratense* Loos, 1949, two dorylaimid nematodes attacking fruit trees in India. *Zeitschrift für Parasitenkunde* **20**:457-465

Siddiqi, M. R. (1973). Systematics of the genus *Trichodorus* Cobb, 1913 (Nematoda: Dorylaimida), with descriptions of three new species. *Nematologica* **19**:259-278

Siddiqi, M. R. (1980). The origin and phylogeny of the nematode orders Tylenchida Thorne, 1949 and Aphelenchida n. ord. *Helminthological Abstracts Series B* **49**:143-170

Siddiqi, M. R. (1983). Phylogenetic relationships of the soil nematode orders Dorylaimida, Mononchida, Triplonchida and Alaimida, with a revised classification of the subclass Enoplia. *Pakistan Journal of Nematology* **1**(2):79-110

Siddiqi, M. R. (2000). Tylenchida, Parasites of Plants and Insects. 2nd Edition, CAB International Publications, Wallingford, Oxon, 833 pp.

Sidow, A. T.; Nguyen. T.; Speed, T. P. (1992). Estimating the fraction of invariable codons with a capture-recapture method. *Journal of Molecular Evolution* **35**:253-260

Simon, D.; Larget, B. (1998). Bayesian Analysis in Molecular Biology and Evolution (BAMBE). Version 1.0, Department of Mathematics and Computer Science, Duquesne University Pittsburgh, Pennsylvania

Simpson, G. G. (1945). The principles of classification and a classification of mammals. *Bulletin of the American Museum of Natural History* **85**:1-350

Smith, A. B. (1989). RNA sequence data in phylogenetic reconstruction: Testing the limits of its resolution. *Cladistics* **5**:321-344

Smith, A. B. (1994). Rooting molecular trees: Problems and strategies. *Biological Journal of the Linnaean Society* **51**:279-292

Smith, A. B.; Lafay, B.; Christen, R. (1992). A comparison of morphological and molecular rates of evolution: 28S ribosomal RNA in echinoids. *Philosophical Transactions of the Royal Society of London Series B – Biological Sciences* **338**:365-382

Sober, E. (1983). Parsimony in systematics: Philosophical issues. *Annual Review of Ecology and Systematics* **14**:335-357

Sober, E. (1988). Reconstructing the Past: Parsimony, Evolution, and Inference. MIT Press, Cambridge, Massachusetts

Sokal, R. R.; Rohlf, F. J. (1981). Taxonomic congruence in the Leptopodomorpha re-examined. *Systematic Zoology* **30**:309-325

Sokal, R. R.; Shao, K. (1985). Characters stability in 39 data sets. *Systematic Zoology* **34**:83-89

Sota, T.; Vogler, A. P. (2001). Incongruence of mitochondrial and nuclear gene trees in the carabid beetles *Ohomopterus*. *Systematic Biology* **50**:39-59

Spiridonov, S. E.; Yushin, V. V. (2000). Ultrastructure of the stoma in *Heth mauriesi* Adamson, 1982 (Rigonematida: Hethidae). *Nematology* **2**(4):417-424

Stanton, J. M.; McNicol, C. D.; Steele, V. (1998). Non-manual lysis of second stage *Meloidogyne* juveniles for identification of pure and mixed samples based on polymerase chain reaction. *Australian Plant Pathology* **27**:112-115

Steel, M. A. (1989). Distribution of bicoloured evolutionary trees. Ph.D. Dissertation, Massey University, Palmerston North, New Zealand

Steel, M. A.; Hendy, M. D.; Penny, D. (1993a). Parsimony can be consistent! *Systematic Biology* **42**(4):581-587

Steel, M. A.; Székely, L.; Erdös, P. L.; Waddell, P. J. (1993b). A complete family of phylogenetic invariants for any number of taxa under Kimura's 3ST model. *New Zealand Journal of Botany* **31**:289-296

Steel, M. A.; Lockhard, P. J.; Penny, D. (1993c). Confidence in evolutionary trees from biological sequence data. *Nature* **364**:440-442

Steel, M. A.; Szikely, L. A.; Hendy, M. D. (1994). Reconstructing trees when sequence sites evolve at variable rates. *Journal of Computational Biology* **1**:153-163

Steiner, G. (1933). The nematode *Cylindrogaster longistoma* (Stefanski) Goodey, and its relationship. *Journal of Parasitology* **20**:66-68

Stewart, A. C.; Nicholas, W. L. (1994). Fine structure of the head and cervical region of *Ceramonema carinatus* (Chromadorida: Ceramonematidae). *Journal of Nematology* **26**(2):188-211

Stuessy, T. F. (1990). Plant Taxonomy. The Systematic Evaluation of Comparative Data. New York, Columbia University Press, 514 pp.

Sudhaus, W. (1976). Vergleichende Untersuchungen zur Phylogenie, Systematik, Ökologie, Biologie und Ethologie der Rhabditidae (Nematoda). *Zoologica* **43**(125):1-229

Sudhaus, W.; Fitch, D. (2001). Comparative studies on the phylogeny and systematics of the Rhabditidae (Nematoda). *Journal of Nematology* **33**(1):1-70

Swofford, D. L. (1991). When are phylogeny estimates from molecular and morphological data incongruent? *In:* Phylogenetic Analysis of DNA sequences. Miyamoto, M. M.; Cracraft, J. (Eds.), Oxford University Press, Oxford, pp. 295-333

Swofford, D. L. (1993). PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1, formerly distributed by Illinois Natural History Survey, Champaign, Illinois [software program]

Swofford, D. L. (2002). PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods). Version 4, Sinauer Associates, Sunderland, Massachusetts [software program]

Swofford, D. L.; Olsen, G. J. (1990). Phylogeny reconstruction. *In:* Molecular Systematics. 2nd Edition, Hillis, D. M.; Moritz, C.; Mable, B. K. (Eds.), Sinauer Associates, Sunderland, Massachusetts, pp. 411-501

Swofford, D. L.; Olsen, G. J.; Waddell, P. J.; Hillis, D. M. (1996). Phylogenetic Inference. *In:* Molecular Systematics. 2nd Edition, Hillis, D. M.; Moritz, C.; Mable, B. K. (Eds.), Sinauer Associates, Sunderland, Massachusetts, pp. 407-514

Tamura K.; Nei M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* **10** (3):512-526

Tarjan, A. C.; Hopper, B. E. (1974). Nomenclatorial compilation of plant and soil nematodes. The Society of Nematologists Inc., E. O. Painter Printing Co., DeLeon Springs, Florida, 419 pp.

Tarrio, R.; Rodrígues-Trelles, F.; Ayala, F. J. (2000). Tree rooting with outgroups when they differ in their nucleotide composition from the ingroup: The *Drosophila saltans* and *willistoni* groups, a case study. *Molecular Phylogenetics and Evolution* **16**(3):344-349

Tchesunov (Chesunov), A. V. (1990a). Cephalic cuticular structures in some free-living nematodes from the superfamily Monhysteroidea (Chromadoria, Monhysterida), Communication 1. *Zoologichesky Zhurnal* **69**(11):18-24

Tchesunov (Chesunov), A. V. (1990b). Cephalic cuticular structures in some free-living nematodes from the superfamily Monhysteroidea (Chromadoria, Monhysterida), Communication 2. *Zoologichesky Zhurnal* **69**(12):13-23

Tehler, A. (1995). Morphological data, molecular data, and total evidence in phylogenetic analysis. *Canadian Journal of Botany – Revue Canadienne de Botanique, Supplement* **73**:667-676

Templeton, A. R. (1983). Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and apes. *Evolution* **37**(2):221-244

The *C. elegans* Consortium (1998). Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Nature* **282**:2012-2018

Thomas, W. K.; Vida, J. T.; Frisse, L. M.; Mundo, M.; Baldwin, J. G. (1997). DNA sequences from formalin-fixed nematodes: Integrating molecular and morphological approaches to taxonomy. *Journal of Nematology* **29**(3):250-254

Thomas, W. K.; Wilson, A. C. (1991). Mode and tempo of molecular evolution in the nematode *Caenorhabditis*: Cytochrome oxidase II and Camodulin sequence. *Genetics* **128**:269-279

Thompson, I. R.; Ladiges, P. Y.; Ross, J. H. (2001). Phylogenetic studies of the tribe Brongniartieae (Fabaceae) using nuclear DNA (ITS-1) and morphological data. *Systematic Botany* **26**(3):557-570

Thompson, J. D.; Gibson, T. J.; Plewniak, F.; Jeanmougin, F.; Higging, D. G. (1997). The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25**(24):4876-4882

Thompson, J. D.; Higgins, D. G.; Gibson, T. J. (1994). CLUSTAL_W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap-penalties and weight matrix choice. *Nucleic Acids Research* **22**:4673-4680

Thorne, G. (1939). A monograph of the nematodes of the superfamily Dorylaimoidea. *Capita Zoologica* **8**(5):1-261

Timm, R. W. (1969). The genus *Isolaimium* Cobb, 1920 (Order Isolaimida, Isolaimiidae New Family). *Journal of Nematology* **1**(2):97-106

Titus, T. A.; Frost, D. R. (1996). Molecular homology assessment and phylogeny in the lizard family Opluridae (Squamata:Iguania). *Molecular Phylogenetics and Evolution* **6**(1):49-62

Urbancik, W.; Bauer-Nebelsick, M.; Ott, J. A. (1996a). The ultrastructure of the cuticle of nematoda. 1. The body cuticle within the Stilbonematinae (Adenophorea, Desmodoridae). *Zoomorphology* **116**(2):51-64

Urbancik, W.; Bauer-Nebelsick, M.; Ott, J. A. (1996b). The ultrastructure of the cuticle of nematoda. 2. The cephalic cuticle of Stilbonematinae (Adenophorea, Desmodoridae). *Zoomorphology* **116**(2):65-75

van der Velde, M. C.; Coomans, A. (1991). The ultrastructure of the buccal cavity of the monhysterid nematodes *Geomonhystera disjuncta* and *Diplolaimella dievengatensis*. *Revue de Nématologie* **14**:133-146

van der Velde, M. C.; de Ley, P.; Monport, D.; Baujard, P.; Coomans, A. (1994). The ultrastructure of the buccal cavity and the cuticle of three Cephalobidae (Nematoda: Rhabditida). *Nematologica* **40**:541-563

von Lieven, A. F.; Sudhaus, W. (2000). Comparative and functional morphology of the buccal cavity of Diplogastrina (Nematoda) and a first outline of the phylogeny of this taxon. *Journal of Zoological Systematics and Evolutionary Research* **38**(1):37-63

Voronov, D. A. (1999). The embryonic development of *Pontonema vulgare* (Enoplida: Oncholaimidae) with a discussion of nematode phylogeny. *Russian Journal of Nematology* **7**(2):105-114

Voronov, D. A.; Panchin, Y. V.; Spiridonov, S. E. (1998). Nematode phylogeny and embryology. *Nature* **395**:28

Waddell, P. J. (1995). Statistical methods of phylogenetic analysis, including Hadamard conjugations, LogDet transforms, and maximum likelihood. Ph.D. Dissertation, Massey University

Waddell, P. J.; Kishino, H.; Ota, R. (2000). Rapid evaluation of the phylogenetic congruence of sequence data using likelihood ratio tests. *Molecular Biology and Evolution* **17**(12):1988-1992

Waddell, P. J.; Penny, D. (1996). Evolutionary trees of apes and humans from DNA sequences. *In:* Handbook of Human Symbolic Evolution. Lock, A. L.; Peters, C. R. (Eds.), Clarendon Press New York, Oxford University Press, Oxford

Ward, S.; Thomson, N.; White, J. G.; Brenner, S. (1975). Electron microscopical reconstruction of the anterior sensory anatomy of the nematode *Caenorhabditis elegans*. *Journal of Comparative Neurology* **160**:313-338

Warwick, R. M.; Platt, H. M.; Somerfield, P. J. (1998). Freeliving Marine Nematodes: Part III Monhysterids. Pictoral key to world genera and notes for the identification of British species. Synopses of the British Fauna (New Series) No. 53, Field Studies Council, 296 pp.

Watrous, L. E.; Wheeler, Q. D.; 91981). The outgroup comparison method of character analysis. *Systematic Zoology* **30**(1):1-11

Watts, K. J.; Courtney, C. H.; Reddy, G. R. (1999). Development of a PCR- and probe-based test for the sensitive and specific detection of the Dog Heartworm, *Dirofiliaria immitis*, in its mosquito intermediate host. *Molecular and Cellular Probes* **14**:425-430

Wheeler, W, C.; Gatesy, J.; DeSalle, R. (1995). Elision: A method for accommodating multiple molecular sequence alignments with alignment-ambiguous sites. *Molecular Phylogenetic and Evolution* **4**:1-9

Wheeler, W. C. (1990). Nucleic acid sequence phylogeny and random outgroups. *Cladistics* **6**:363-367

Wheeler, W. C.; Gladstein, D. L. (1994). MALIGN: A multiple nucleic and sequence alignment program. *Journal of Heredity* **85**:417

Wheeler, W. C.; Honeycutt, R. L. (1988). Paired sequence difference in ribosomal RNAs: Evolutionary and phylogenetic considerations. *Molecular Biology and Evolution* **5**:90-96

Whiting, M. F. (1998). Phylogenetic position of the Strepsiptera: Review of molecular and morphological evidence. *International Journal of Insect Morphology and Embryology* **27**(1):53-60

Wiens, J. J. (2003). Missing data, incomplete taxa, and phylogenetic accuracy. *Systematic Biology* **52**(4):528-538

Wiens, J. J.; Chippindale, P. T. (1994). Combining and weighting characters and the prior agreement approach revisited. *Systematic Biology* **43**(4):564-566

Wiens, J. J.; Hillis, D. M. (1996). Accuracy of parsimony analysis using morphological data: A reappraisal. *Systematic Botany* **21**(2):237-243

Wiens, J. J.; Reeder, T. W. (1995). Combining data sets with different numbers of taxa for phylogenetic analysis. *Systematic Biology* **44**(4):548-558

Wieser, W. (1954). Free-living marine nematodes II. Chromadoroidea. Chile Reports. *Lunds Universitets Arsskr. N. F. Avd. 2* **50**(16):1-148

Wilbur, J.; Lipman, D. (1983). The context dependent comparison of biological sequences. *SIAM Journal of Applied Mathematics* **44**:557-567

Wiley, E. O. (1981). Phylogenetics: The Theory and Practice of Phylogenetic Systematics. New York, John Wiley and Sons, 439 pp.

Wilkinson, M. (1992). Ordered versus unordered characters. *Cladistics* **8**:375-385

Wilkinson, M. (1995). Comparison of two methods of character construction. *Cladistics* **11**:297-308

Wilkinson, M.; Benton, M. J. (1996). Sphenodonid phylogeny and the problems of multiple trees. *Philosophical Transactions of the Royal Society of London Series B – Biological Sciences* **351**:1-16

Williams, P. L.; Fitch, W. M. (1989).Finding the minimal change in a given tree. *In:* The Hierarchy of Life. Fernholm, B.; Bremer, K.; Jörnvall, H. (Eds.), Elsevier Science, Amsterdam, pp. 453-470

Winterton, S. L.; Yang, L. L.; Wiegmann, B. M.; Yeates, D. K. (2001). Phylogenetic revision of Agapophytinae subf. n. (Diptera: Therevidae) based on molecular and morphological evidence. *Systematic Entomology* **26**(2):173-211

Won, J. H.; Rho, B. J.; Song, J. I. (2001). A phylogenetic study of the Anthozoa (Phylum Cnidaria) based on morphological and molecular characters. *Coral Reefs* **20**(1):39-50

Wright, K. A.; Thompson, J. N. (1981). The buccal capsule of *Caenorhabditis elegans* (Nematoda: Rhabditoidea): An ultrastructural study. *Canadian Journal of Zoology* **59**:1952-1961

Wuyts, J.; Van de Peer, Y.; Winkelmans, T.; De Wachter, R. (2002). The European database on small subunit ribosomal RNA. *Nucleic Acids Research* **30**(1):183-185

Yang, Z. (1993). Maximum likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Molecular Biology and Evolution* **10**:1396-1401

Yang, Z. (1994a). Maximum Likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *Journal of Molecular Evolution* **39**:306-314

Yang, Z. (1994b). Statistical properties of the maximum-likelihood method of phylogenetic estimation and comparison with distance matrix methods. *Systematic Biology* **43**(3):329-342

Yoder, A. D.; Irwin, J. A.; Payseur, B. A. (2001). Failure of the ILD to determine data compatibility for slow loris phylogeny. *Systematic Biology* **50**:408-424

Zarlenga, D. S.; Hoberg, E. P.; Stringfellow, F.; Lichtenfels, J. R. (1998). Comparison of two polymorphic species of *Ostertagia* and phylogenetic relationships within the Ostertagiinae (Nematoda: Trichostrongylidea) inferred from ribosomal DNA repeat and mitochondrial DNA sequences. *Journal of Parasitology* **84**(4):806-812

Zhang, Y. C.; Baldwin, J. G. (1999). Ultrastructure of the esophagus of *Diplenteron* sp. (Diplogasterida) to test hypotheses of homology with Rhabditida and Tylenchida. *Journal of Nematology* **31**:1-19

Zhang, Y. C.; Baldwin, J. G. (2000). Ultrastructure of the post-corpus of *Zeldia punctata* (Cephalobina) for analysis of the evolutionary framework of nematodes related to *Caenorhabditis elegans* (Rhabditina). *Proceedings of the Royal Society of London Series B* **267**(1449):1229-1238

Zhang, Y. C.; Baldwin, J. G. (2001). Ultrastructure of the postcorpus of the esophagus of *Teratocephalus lirellus* (Teratocephalidea) and its use for interpreting character evolution in Secernentea (Nematoda). *Canadian Journal of Zoology* **79**:16-25

Zharkikh, A.; Li, W.-H. (1992a). Statistical properties of bootstrap estimation of phylogenetic variability from nucleotide sequences. I. Four taxa with a molecular clock. *Molecular Biology and Evolution* **9**:1119-1147

Zharkikh, A.; Li, W.-H. (1992b). Statistical properties of bootstrap estimation of phylogenetic variability from nucleotide sequences. II. Four taxa without a molecular clock. *Journal of Molecular Evolution* **35**:356-366

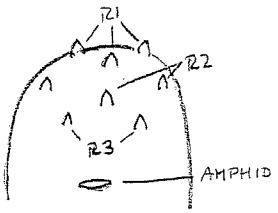
Zharkikh, A.; Li, W.-H. (1993). Inconsistency of the maximum parsimony method: The case of five taxa with a molecular clock. *Systematic Biology* **42**:113-366

Zharkikh, A.; Li, W.-H. (1995). Estimation of confidence in phylogeny: The complete-and-partial bootstrap technique. *Molecular Phylogenetics and Evolution* **4**(1):44-63a

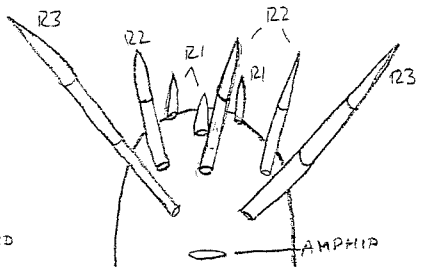
Plates

The following nine pages show figures of many of the characters analysed in this thesis. They are not drawn to scale and and often do not represent a single species. They are only schematic drawings providing the reader with further information on the individual characters.

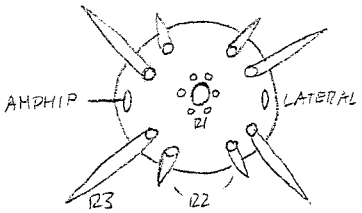
PLATE I.



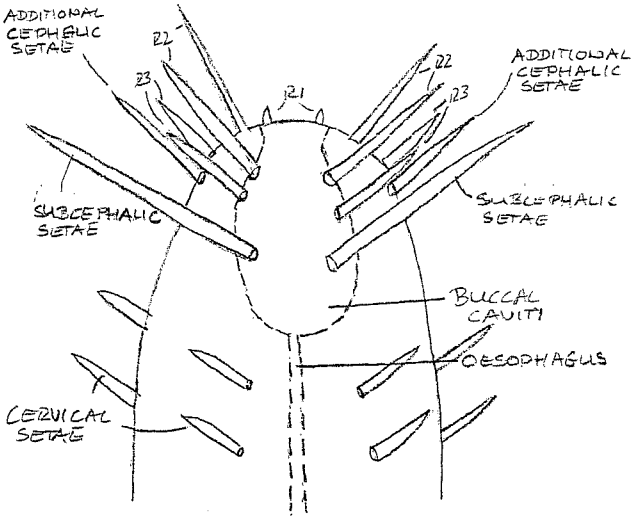
1. ALL SENSILLA PAPILLIFORM (LATERAL VIEW)



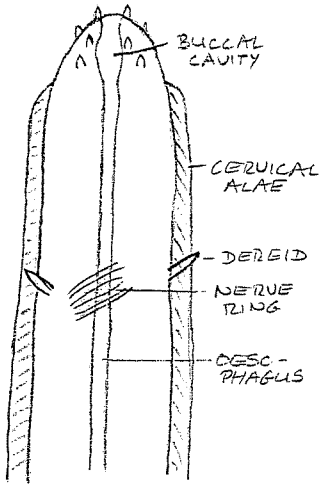
2. ALL SENSILLA SETIFORM, R2 WITH 2 SEGMENTS, R3 WITH 3 SEGMENTS (LATERAL VIEW)



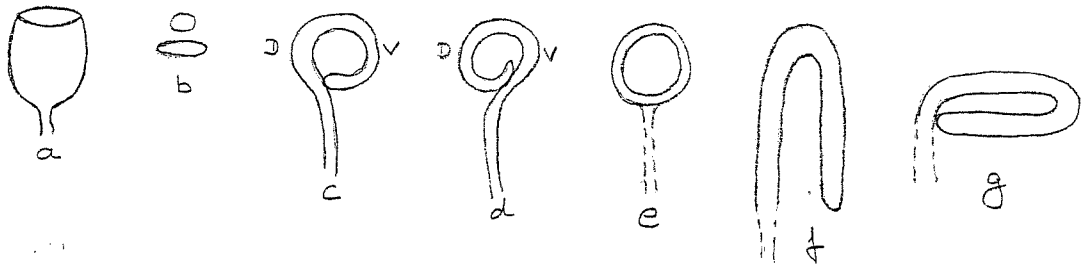
3. EN FACE VIEW, R1 PAPILLIFORM, R2+3 SETIFORM AND AT SAME LEVEL



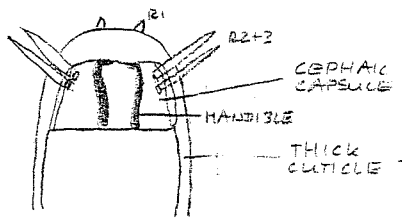
4. ADDITIONAL ANTERIOR SETAE



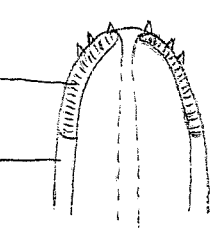
5. POSITION OF CERVICAL ALAE AND DEREIDS, CEPHALIC SENSILLA ALL PAPILLIFORM



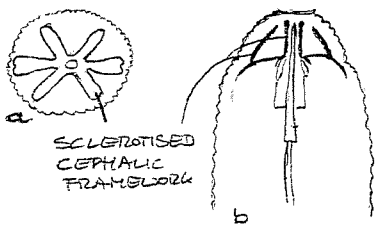
6. AMPHIDS: a - POCKET-SHAPED, b - PORE-SHAPED, c - VENTRALLY-SPIRAL, d - DORSALLY-SPIRAL (V-VENTRAL SIDE, D-DORSAL SIDE), e - ROUND, f - ELONGATED LOOP, g - SLIT-LIKE



7. CEPHALIC CAPSULE IN ENOPLUS

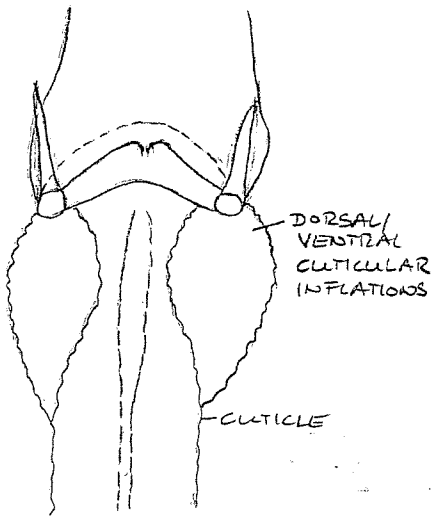


8. CEPHALIC CAPSULE IN DESMODORIDAE

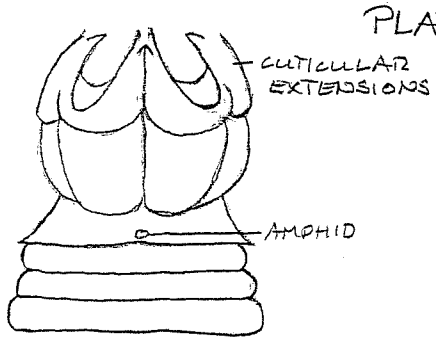


9. HEAD STRUCTURE OF TYLENCHID

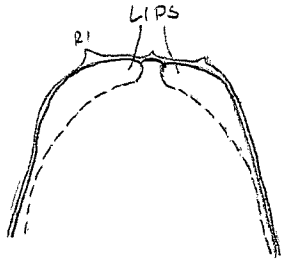
PLATE II.



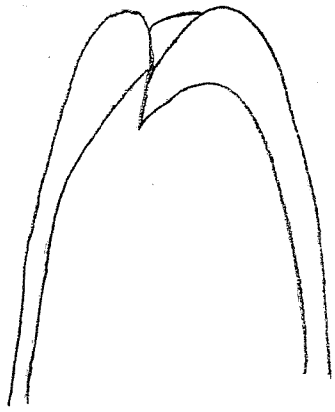
10. HEAD STRUCTURE OF SOME PLECTIDAE (LATERAL VIEW)



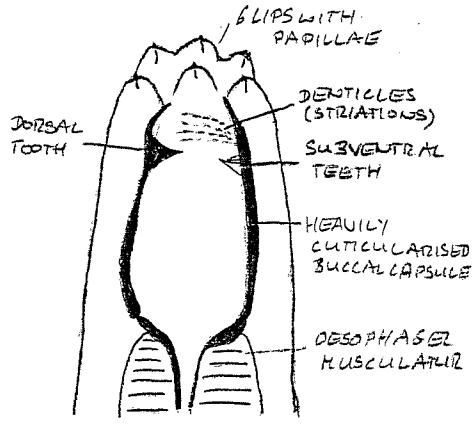
11. HEAD STRUCTURE OF TERATOCEPHALUS LIDELLUS (LATERAL VIEW)



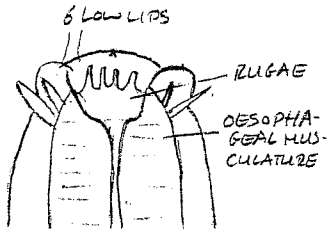
12. 3 SMALL LIPS (THIRD BEHIND, NOT VISIBLE)



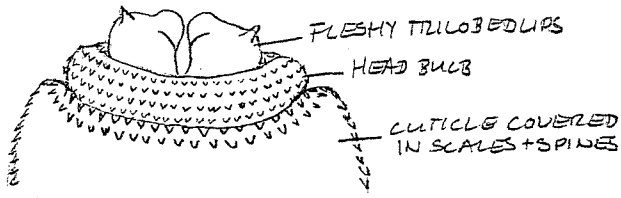
13. ONE DEEPLY-BILOBED LIP, SECOND ONE JUST VISIBLE BEHIND, AS IN BATHYLAIMUS



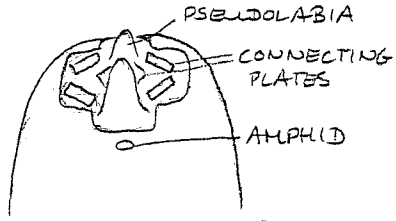
14. HEAD STRUCTURE OF MONONCHIDA, HERE WITH 6 WELL DEVELOPED LIPS



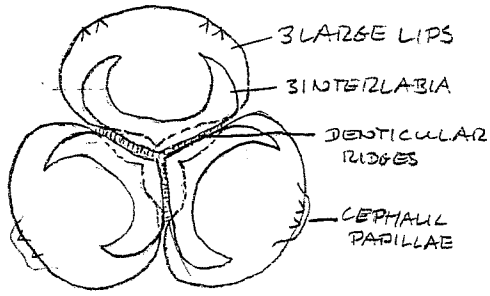
15. HEAD STRUCTURE OF CHROMADORIDAE, TEETH OMITTED



16. HEAD STRUCTURE OF GNATHOSTOMA

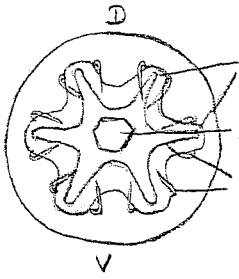


17. HEAD STRUCTURE OF ASCAROPHIS (LATERAL VIEW)

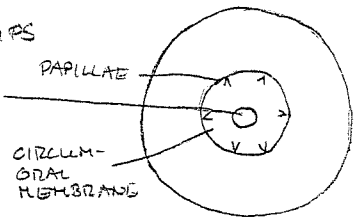


18. LIP STRUCTURE OF ASCARIDS (EN FACE VIEW)

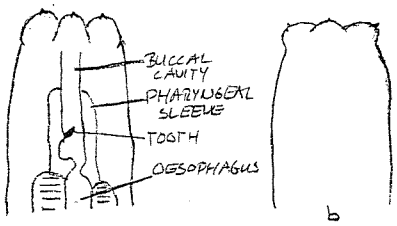
PLATE III.



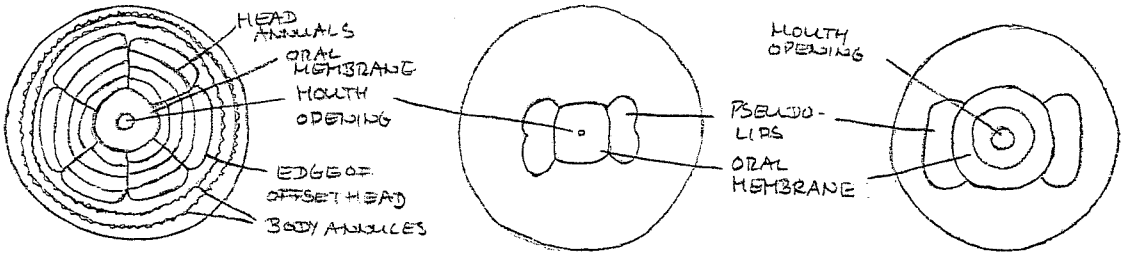
19. HEAD STRUCTURE OF PANAGROBELUS (EN FACE VIEW)



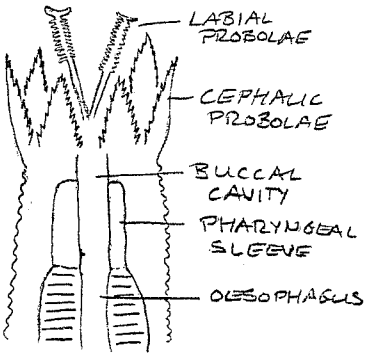
20. CIRCUMORAL MEMBRANE IN PLACE OF LIPS (EN FACE VIEW)



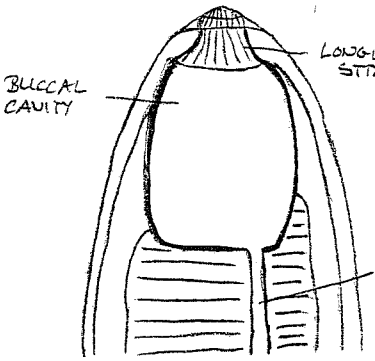
21. 6 OPEN (a) AND 6 CLOSED (b) LIPS IN RHABDITIDAE AND BUCCAL CAVITY STRUCTURE

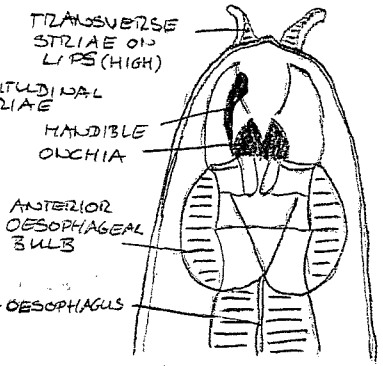


22. 3 TYPES OF HEAD STRUCTURES FOUND IN THE TYLENCHINA, LIPS ARE ALWAYS MERGED TO FORM ORAL MEMBRANE



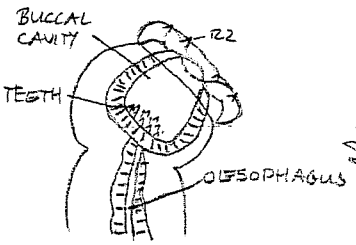
23. HEAD STRUCTURE OF CEPHALOBIDEA



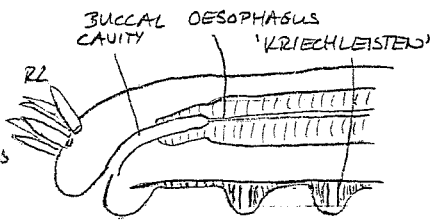


24. HEAD STRUCTURE OF SPHAEROLAIMUS (OMITTING SENSILLA)

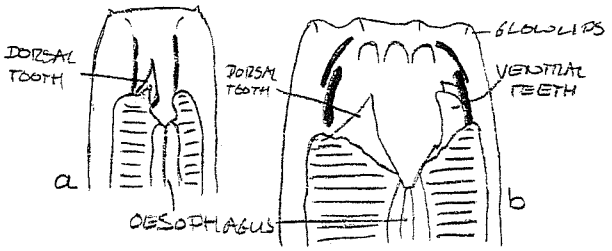
25. HEAD STRUCTURE OF ENOPLOIDES (OMITTING SENSILLA)



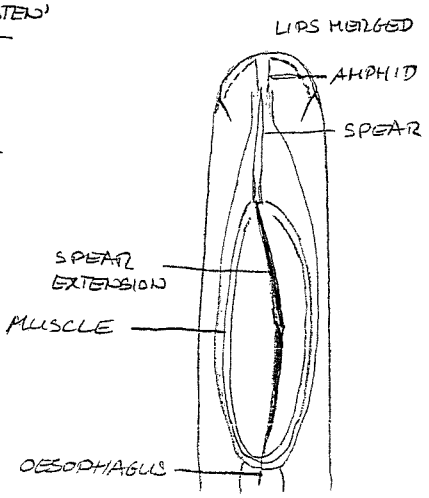
26. HEAD STRUCTURE OF SYNGAMUS



27. HEAD STRUCTURE OF BUNONEMA

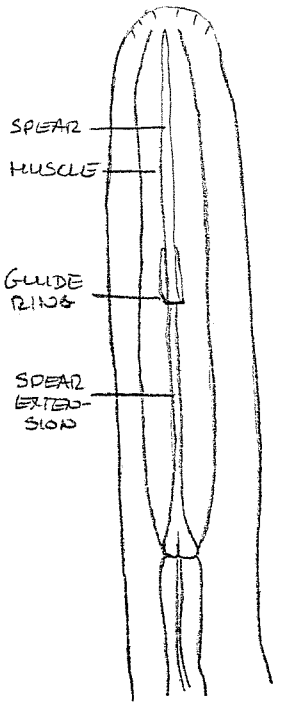


28. HEAD STRUCTURE OF ADUNCOSPICULUM
a - EUNOSTOMATOUS ♀, b - STENOSTOMATOUS ♀

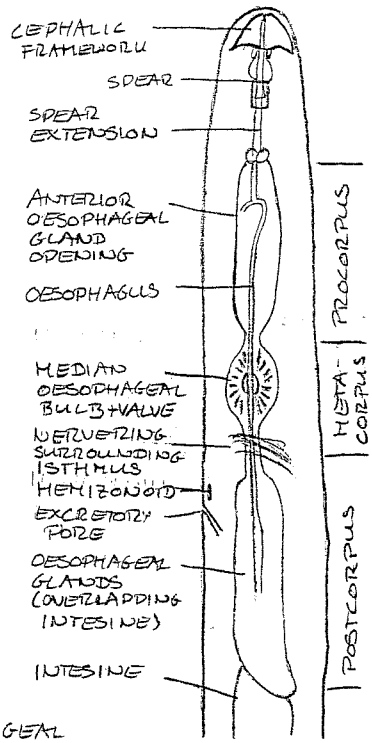


29. ANTERIOR OESOPHAGEAL STRUCTURE OF DIPHTHEROPHOROIDEA

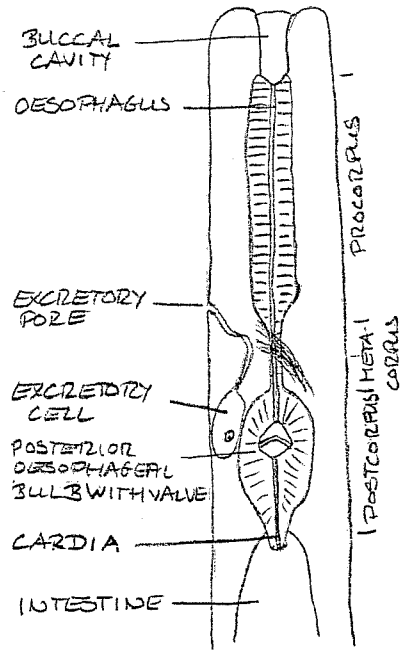
PLATE IV.



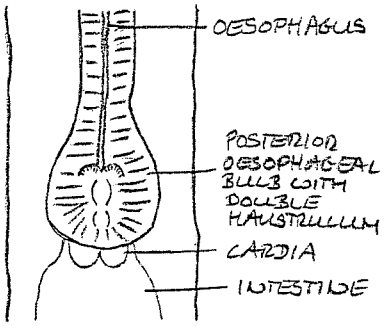
30. ANTERIOR OESOPHAGEAL STRUCTURE OF XIPHINEMA



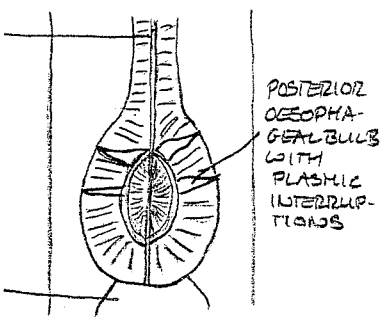
31. OESOPHAGUS STRUCTURE OF TYLENCHIDS



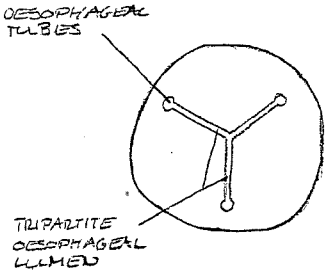
32. OESOPHAGUS OF PLECTIDAE



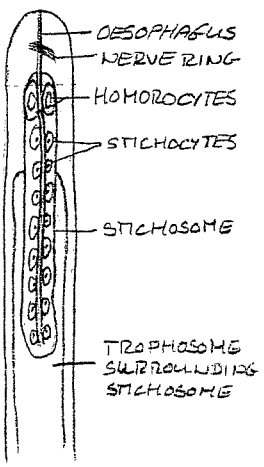
33. POSTERIOR OESOPHAGUS OF RHABDITIDAE



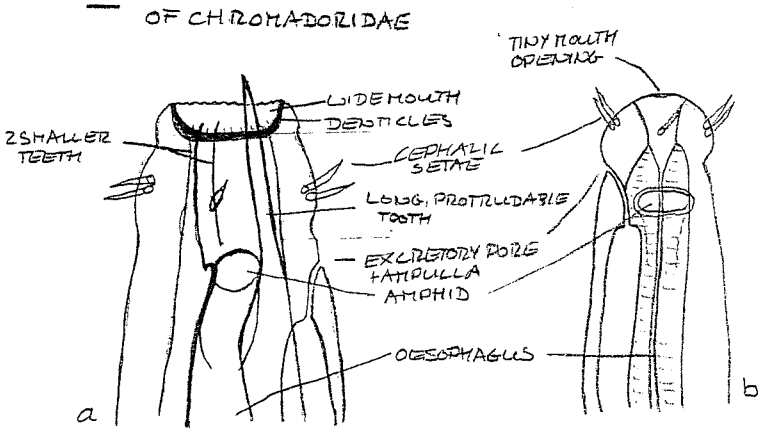
34. POSTERIOR OESOPHAGUS OF CHROMADORIDAE



35. OESOPHAGUS IN TRANSVERSE SECTION

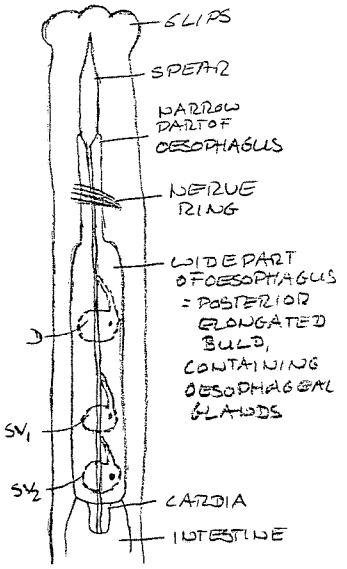


36. STICHOSOME AND TROPHOSOME IN MERMITHIDA

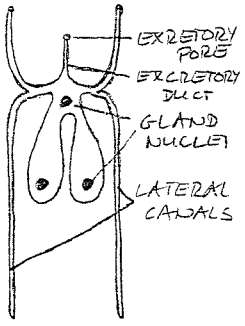


37. HEAD STRUCTURE OF CALYPTRONEMA MAXWEBERI
a - ♀, b - ♂

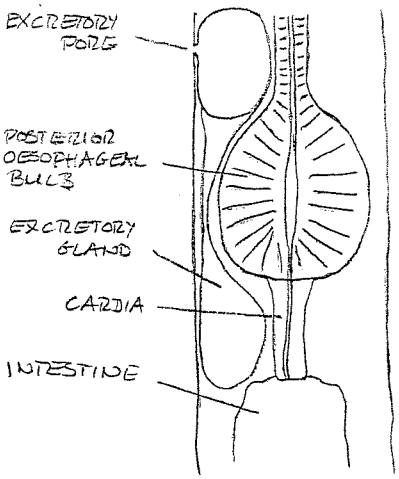
PLATE V.



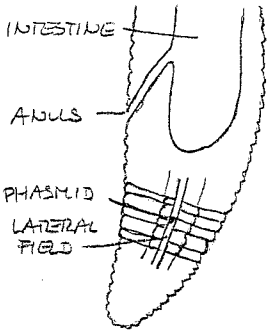
38. OESOPHAGUS OF DORYLAIMIDS (D - DORSAL GLAND, SV - SUBVENTRAL GLANDS)



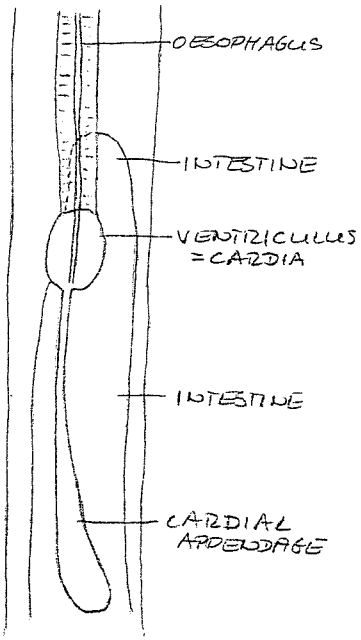
39. MOST ELABORATE EXCRETORY SYSTEM AS FOUND IN RHABDITIDAE



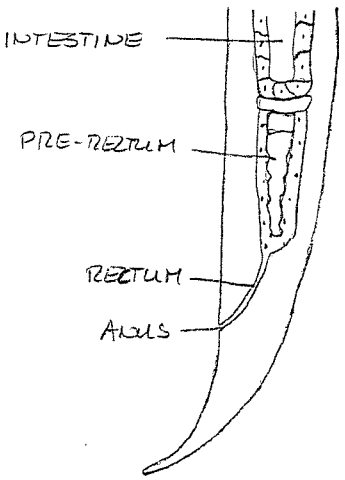
40. POSTERIOR OESOPHAGEAL STRUCTURE OF RHOMBOIDAE



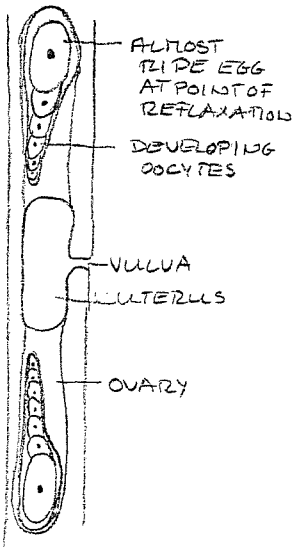
41. POSTERIOR END OF TYLENCHID



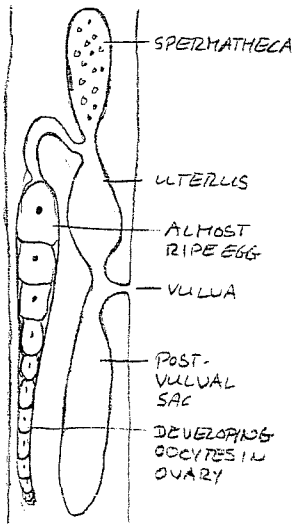
42. CARDIAL REGION IN RAPHIDASCARIDAE (SPIRURINA)



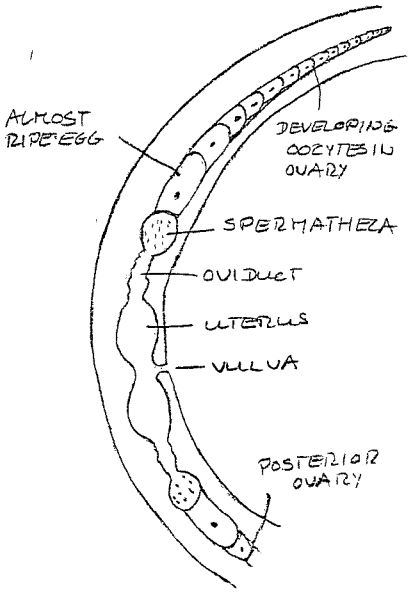
43. POSTERIOR END OF DORYLAIMIDA



44. 2 ANTIDROMOUSLY RELEXO, AMPHIDELPHIC OVARIES

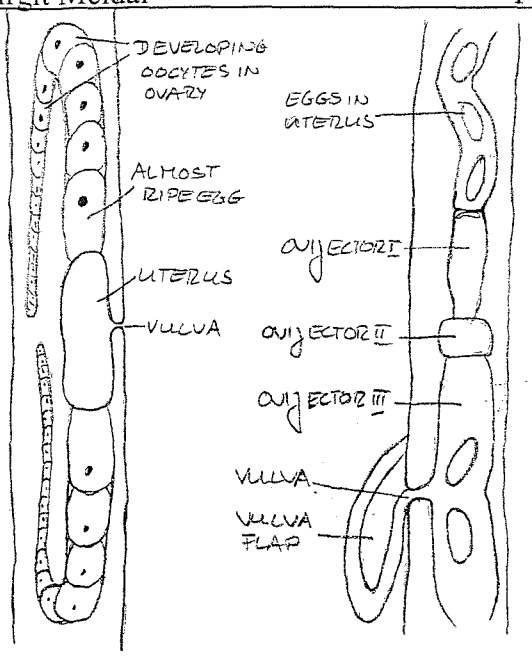


45. OVARY OF CEPHALOBIDAE



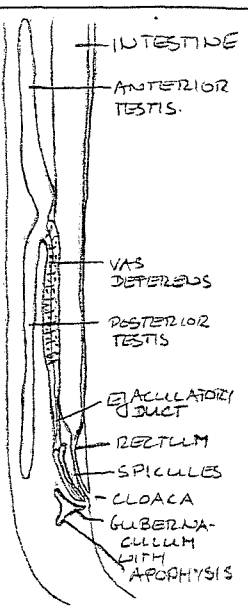
46. 2 OUTSTRETCHED, DIDELPHIC OVARIES AS IN TYLENCHIDS

PLATE VI.

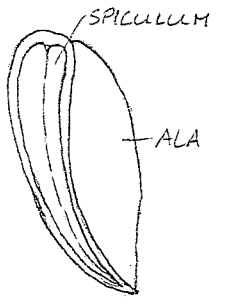


47. 2 AMPHIDELPHIC, HOMODROMOUSLY REFLEXED OVARIES AS IN RHABDITIDE

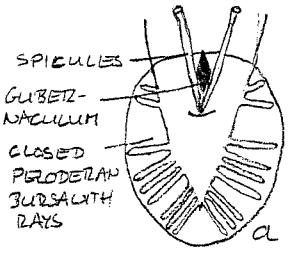
48. VULVA REGION OF OSTERTAGIA (AMPHI-, DI-DELPHIC)

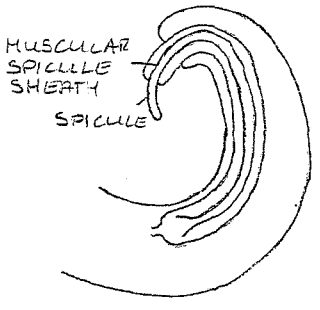


49. 2 AMPHIORCHIC TESTES

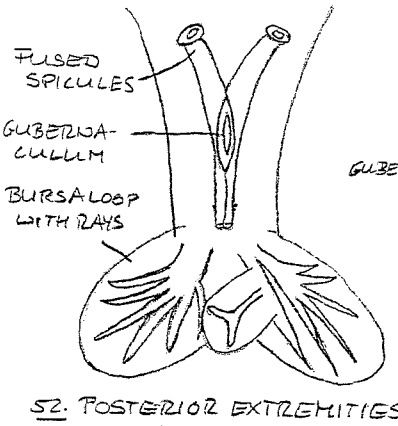


50. SPICULUM WITH VENTRAL ALA

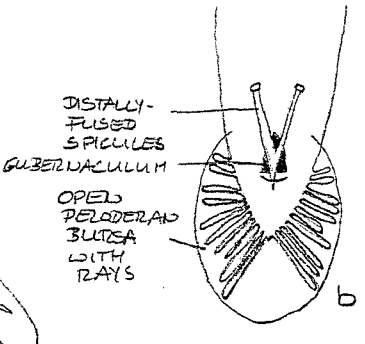




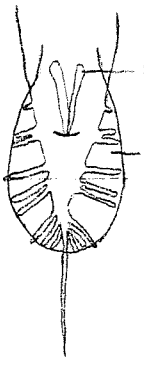
51. SINGLE SPICULE COVERED BY SHEATH IN TRICHURIS



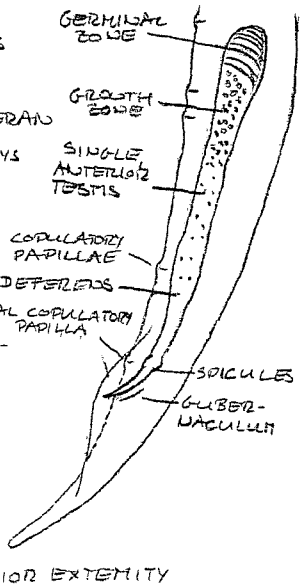
52. POSTERIOR EXTREMITIES OF HAEMONCHUS



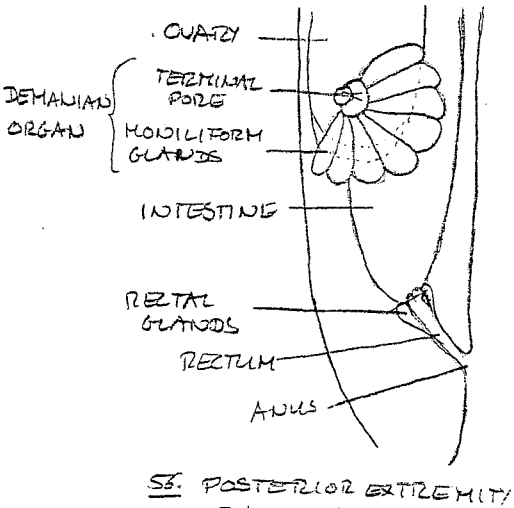
53. PELODERAN POSTERIOR EXTREMITY OF RHABDITIDAE (a AND b)



54. LEPTODERAN POSTERIOR EXTREMITY OF RHABDITIDAE

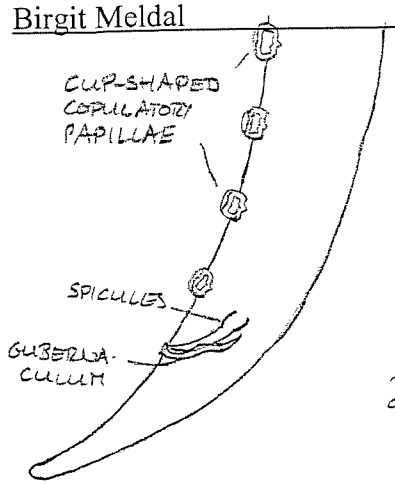


55. POSTERIOR EXTEMITY OF CAMPYDORA

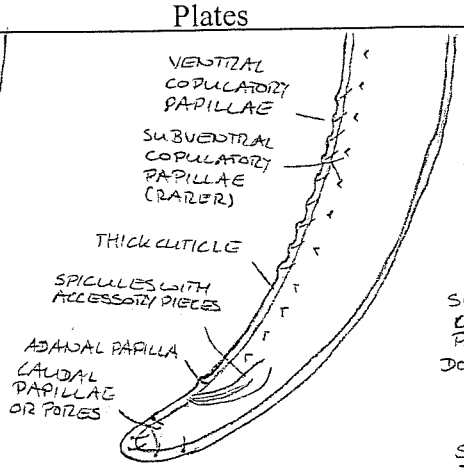


56. POSTERIOR EXTREMITY OF ADONCHOLAIMUS

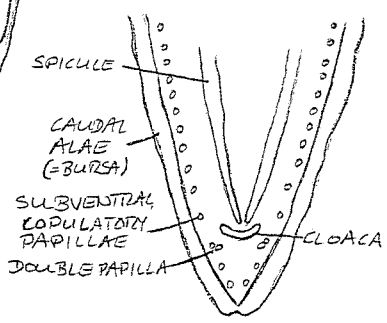
PLATE VII.



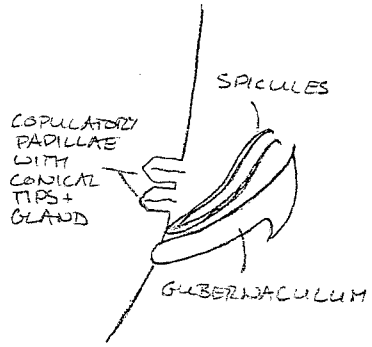
57. POSTERIOR EXTREMITY OF CHROMADORIDAE



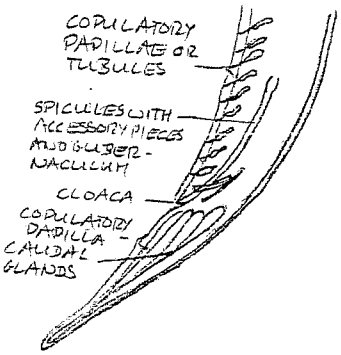
58. POSTERIOR EXTREMITY OF DORYLAIMIDA



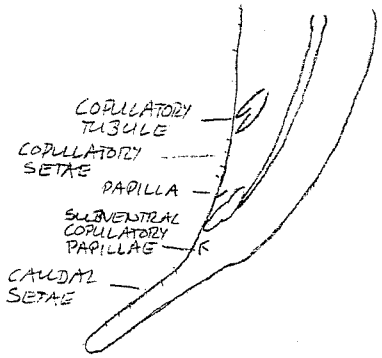
59. POSTERIOR EXTREMITY OF ASCARIDS



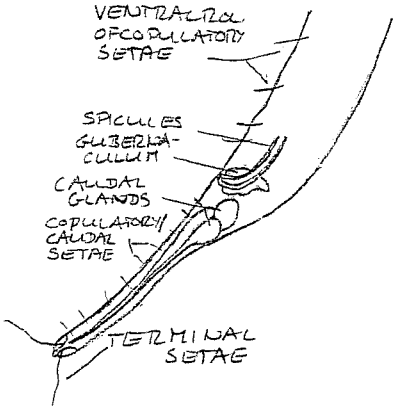
60. CLOACAL REGION OF *MICRODORA BIPAPILLATA*



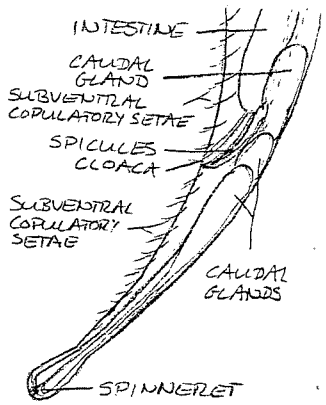
61. POSTERIOR EXTREMITY OF *MONONCHUS*



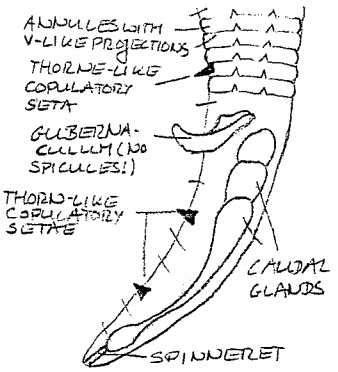
62. POSTERIOR EXTREMITY OF *ENDOPLOIDES*



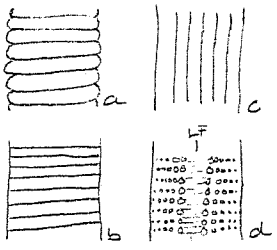
63. POSTERIOR EXTREMITY OF *DAPTONEMA*



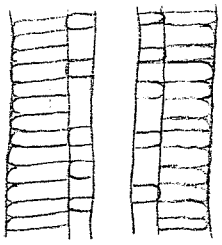
64. POSTERIOR EXTREMITY OF *VISCOSIA*



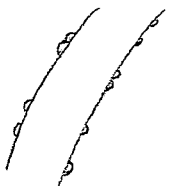
65. POSTERIOR EXTREMITY AND CUTICULAR PATTERN OF *MONOPOSTHIA*



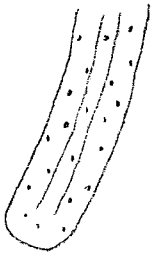
66. CUTICULAR PATTERNS
a - ARTICULATED
b - ANNULATED (STRIATED)
c - LONGITUDINAL STRIAE
d - PUNCTATED, DIFFERENTIATED IN LATERAL FIELD (LF) (eg. CHROMADORIDAE)



67. LATERAL FIELD WITH LONGITUDINAL RIDGES FORMING ALAE AND PARTIALLY AREOLATED, AS IN TYLENCHIDS

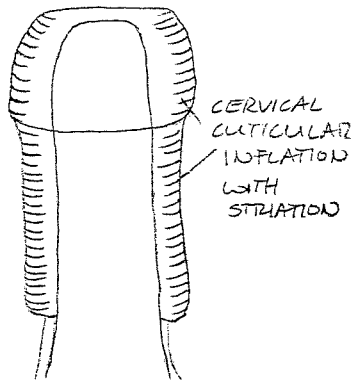


68. CUTICULAR BOSSES OF *LOA*

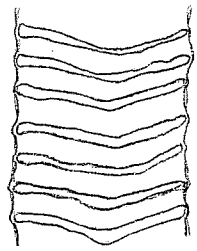


69. CUTICULAR BOSSES ON TAIL OF *LUCHERIA*
(NOTE ALSO SIMPLE LATERAL FIELD)

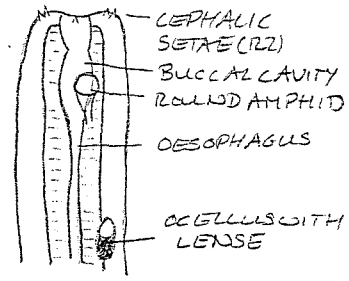
PLATE VIII.



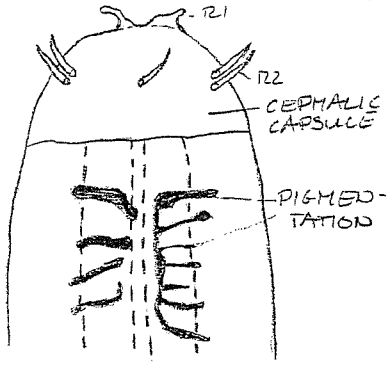
70. ANTERIOR EXTREMITY OF NEMATODIRUS



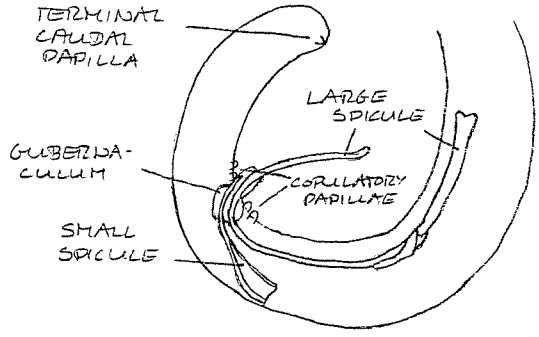
71. EXTERNAL CUTICULAR PATTERN OF ONCHOCERCA



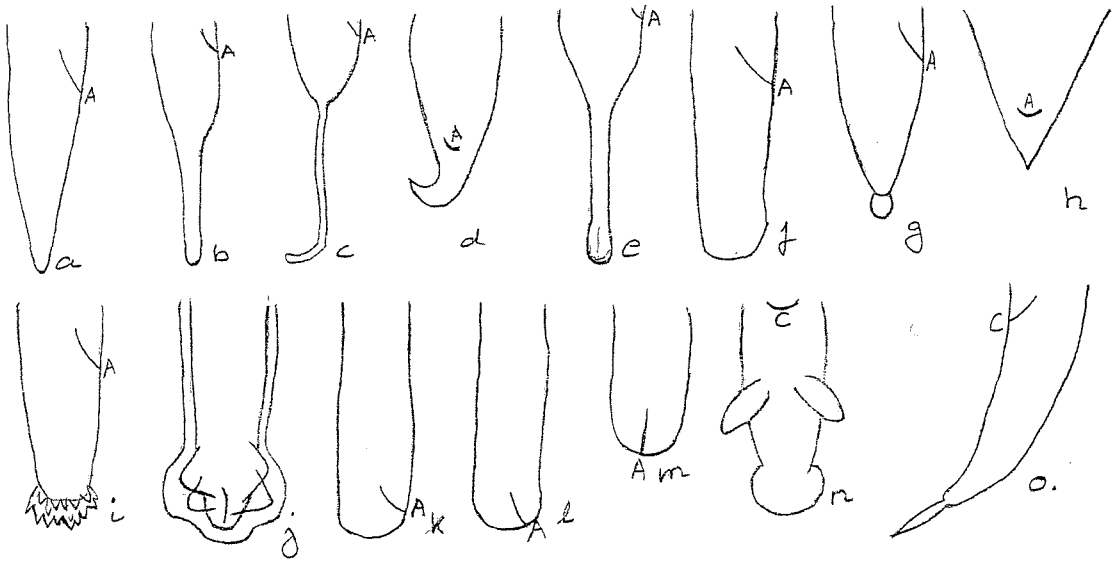
72. ANTERIOR EXTREMITY OF DIPLOLAIMELLOIDES



73. HEAD OF ENOPLUS WITH PIGMENTATION

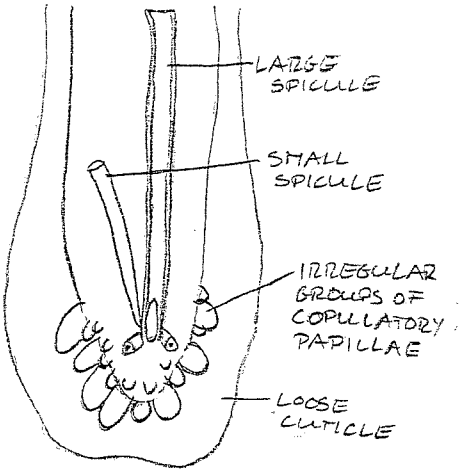


74. POSTERIOR EXTREMITY OF BRUGIA

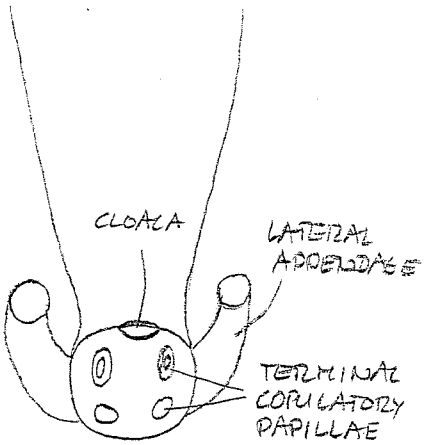


75. TAIL SHAPES: a - CONICAL, b - CONICO-CYLINDRICAL, c - CONICAL-FILIFORM, d - SHORT+BENT, e - CLAVATE, f - ROUND-CYLINDRICAL, g - CONICAL+MUCRON, h - SHORT+SHARPLY POINTED, i - CONICAL+TUFTS OF PROJECTIONS (HYSTEROTHYLACIUM), j - BERMIS (FLAPS), k - HEMISPHERICAL, l - SUBTERMINAL ANUS, m - NO TAIL (TERMINAL ANUS), n - SETARIA ♂: CONICAL APPENDAGES + TERMINAL KNOB, o. - ADUNCOSPICULUM ♂: CONICAL+SPICATE TERMINUS
(A - ANUS (OR CLOACA), C - CLOACA, ALL LATERAL VIEW APART FROM h, j, n, THAT ARE VENTRAL VIEWS)

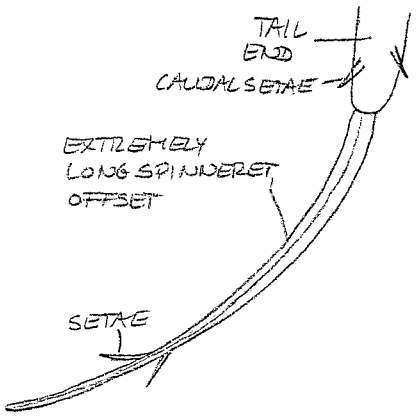
PLATE IX.



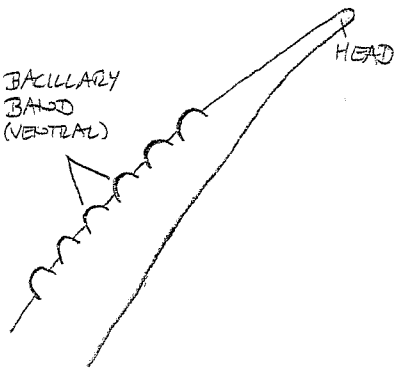
76. POSTERIOR EXTREMITY OF GNATHOSTOMA (VENTRAL VIEW)



77. POSTERIOR EXTREMITY OF TRICHINELLA (PSEUDOBURSA)



78. TAIL END OF SPILOPHORELLA PARADOXA



79. ANTERIOR EXTREMITY OF TRICHURIS