Molecular systematics of Bryozoa.

Tsyganov-Bodounov, Anton

MOLECULAR SYSTEMATICS OF BRYOZOA

Anton Tsyganov-Bodounov

Submitted to Swansea University
in fulfilment of the requirements for the
Degree of Doctor of Philosophy

Swansea University
Prifysgol Abertawe

2008

The current systematic status of Bryozoa and phylogenetic relationships between its orders (Cheilostomata, Ctenostomata, Cyclostomata) and within their families are uncertain. Their present classification is based on the zooid frontal wall and fossil record data, however there is an inconsistency with molecular 16S rDNA gene data Dick *et al.* (2000) where ctenostomes, cyclostomes and cheilostomes were shown to be paraphyletic. Larval morphology has also been emphasised as an area lacking sufficient information.

In the present study molecular sequence data for the 18S rDNA gene have been collected for over 30 species of Bryozoa, based on material collected in South Wales. Bryozoa specific oligonucleotide primers for 18S rDNA were developed, tested and optimised.

Based on the collected 18S rDNA sequences and the secondary structure alignment of the sequences a phylogenetic analysis was performed using Bayesian methods. A mixed evolutionary model was used for different regions of the alignment of 18S rDNA, including an rRNA-specific model.

The resulting trees suggest a monophyletic Cyclostomata. The position of Cheilostomata and Ctenostomata are uncertain and vary depending on whether a sequence of *Alcyonidium gelatinosum* is or is not included in the analysis. Without *A.gelatinosum*, Ctenostomata are a monophyletic clade within paraphyletic Cheilostomata. Addition of *A.gelatinosum* makes Ctenostomata paraphyletic incorporating monophyletic Cheilostomata. Based on these findings, suggestions for further research are given.

In addition, a secondary structure model for *Bugula turbinata* is presented. This is the first bryozoan 18S rRNA structure model and should be of utility in future systematics studies.

A method of larval analysis and visualisation was evaluated using confocal laser microscopy. This method facilitates observation of the external morphology of larvae including a partial 3D reconstruction so that their morphotype based on the Zimmer and Woolacott (1977) system can be identified. This method is superior to previously used epi-fluorescent microscopy approaches due to its much higher resolution and the lower number of artefacts encountered.

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed .. .. (candidate)

Date ..... 8/Apr/2008 .....

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).

Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed ... . (candidate)

Date ..... 8 Apr /2008 .....

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans after expiry of a bar on access approved by the Swansea University.

Signed .... (candidate)

Date ..... 8/Apr/2008 .....

I would like to thank my supervisor Prof. David O F Skibinski for his support, guidance and understanding of all my needs during the project as well as allowing me to develop it in the way and direction which was my own. The same gratitude goes to my supervisor Dr. Peter J Hayward for showing me what Bryozoa are, developing my interest in the group, assisting with field trips and sample collections as well as guidance for all morphological and taxonomic work. I would like to thank Dr. Joanne Porter for assisting me with sampling trips, sharing her experience and providing me with some data for the study, I am also thankful to her for showing support and enthusiasm for my work. I would also like to thank Dr. Andrea Waeschenbach who shared with me important ideas and recommendations during my work. I thank Jean Leamon for teaching me lab techniques and assisting me with field trips on a regular basis. Keith Neilor and Dr. John Lancaster allowed me to join them on several boat trips and showed full support for my needs. Dr. Max Telford kindly provided his software for the data analysis. I would like to acknowledge the Genetics Society for awarding me with a grant which enabled me to take part in Workshop in Molecular Evolution in Woods Hole, MA (USA). I would like thank all colleagues and friends at Swansea University for their support and care during these four years. Finally, I am immensely grateful to all my family, especially to Sarah and Armelle Brunet de Sairigné for their constant support, understanding and patience.

This study was funded by NERC PhD student grant (NER/S/J/2002/12005).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 GENERAL INTRODUCTION

The phylum Bryozoa Ehrenberg, 1831 is represented by sessile colonial aquatic animals, which can be commonly found on the seashore during low tides, encrusting rocks and algae. This group is the largest phylum of the lophophorate group of invertebrates and is commonly represented around the world throughout shelf epifauna. The estimate of the number of extant species is from 3000 (Ryland 1970) to around 5600 (Todd 2000), but possibly larger than that and essentially is unknown (Hayward and Ryland 1998).

Bryozoa are coelomate modular colonial sessile animals. All Bryozoa possess a distinctive organ, the lophophore, a feeding organ which is also found in Phoronida and Brachiopoda, and hence the above cluster group is called Lophophorates, however the composition and uniformity of this group are disputed in literature, and this group was shown to be not monophyletic (Passamaneck and Halanych 2006).

1.1 Bryozoan colony

The bryozoan colony consists of modular blocks – zooids – which grow asexually from the sexually produced and dispersed larva which settles and gives rise to a new colony. Colony main elements are autozooids, but also so-called heterozooids are present in some groups (discussed below). These include avicularia, vibracula, kenozooids, gonozooids and nanozooids. The zooids in the colony are surrounded by walls, sometimes calcified or gelatinous, which are referred to as cystid. The part of the zooids which corresponds to the lophophore, gut and musculature is referred to as polypide (to distinguish it from the cystid). The walls of the cystid include the frontal membrane or a specialised hydrostatic sac are responsible for the protrusion of the lophophore. The individual zooids in the colony are linked via a network called the funiculus through the pores connecting individual zooids. This network is responsible for the transport of metabolic products through the entire colony.

Bryozoan colonies form incrustations on the substrata or grow in series or chains. Colonies of Bryozoa exhibit a great variety of shapes and ways of formation. However, most commonly these are incrustation of the substrate or lobed

or bushy colonies, which adhere to the substratum via rhizoid-like attachments or through direct cementation (Ryland 1970, Hayward and Ryland 1983).

1.2 Reproduction of Bryozoa

Most Gymnolaemata and Stenolaemata are colonial hermaphrodites with gonochroic zooids (Reed 1991). The type of hermaphroditism present is zooidal hermaphroditism and zooidal gonochorism. Some sexual polymorphism is observed with its clear differences in stenolaemates where female maternal zooids are present, a gonozooid. Embryo brooding is very common, especially for Gymnolaemata, however, some species produce many small eggs, which are released directly into the sea. Spermatozoa are released into the surrounding water and then cross fertilisation takes place (Ryland 1970). For those species which release their eggs into the sea the fertilisation occurs during or just immediately before the egg release.

Reproduction seasons of Bryozoa have been scarcely studied and usually information about the reproduction of any particular species has to be collected form an array of literature rather than from any particular study dedicated to the reproduction periods (Reed 1991). In this work (Chapter 2) the reproduction period of those species which were collected is further discussed. Reproduction period recording is dependent on the presence of eggs/embryos in the colonies and is often correlated to the geographical distribution of the species (Ryland 1970, Reed 1991).

The majority of bryozoans brood their embryos and release completely developed and mature larvae. Gymnolaemates and in particular ctenostomes brood their embryos in an introvert which sometimes (often in Ctenostomata) results in the degeneration of the polypide. Brooding sac is also observed in some species and an ovicell is also common (Cheilostomata).

In Stenolaemata brood chamber specialisation is characteristic of the class where a female maternal zooid is modified as a gonozooid and numerous embryos produced in the brood chamber.

Released larvae are short-lived and their release is linked to light stimulus in most shallow water species. In most cases it is followed by positive phototaxis just after the release of the larvae and negative phototaxis immediately prior to the larval settlement, in addition some negative geotactic responses are reported (Ryland

1977). However some stenolaemates are reported to have their larval release after the sunset (Reed 1991).

In general, larvae which are released from the colonies are fully developed so that their settlement and metamorphosis can begin very shortly after their release (within minutes) (Reed 1991).

For Gymnolaemata larvae a detailed system of larval morphology was proposed by Zimmer and Woollacott (1977), this system is described in more detail in Chapter 6 where larval morphology is also reviewed.

In general however, morphology of the larvae is fully adapted for the locomotion and sensory organs, which aid it in the substrate searching and settlement process and is not linked to the morphology of an adult form (Reed 1991). Larvae can be separated into planktotrophic and lecithotrophic types, the latter being prevalent in most Gymnolaemata (Zimmer and Woolacott 1977). Many organs are common between all larvae due to their similar functions.

Much less is known about the larval morphology of the Stenolaemata and the account of larval behaviour is limited to six species (Nielsen 1970). The larvae appear to be lecithotrophic without many secondary organs observed.

Following settlement, larvae reorganise themselves into preancestrula (or primary disk in stenolaemates, Nielsen 1970) after which a process of histogenesis follows and the first zooid of the colony appears, capable of feeding. The tissues specific to the larva itself undergo histolysis.

1.3 Classification of Bryozoa

There persists an apparent confusion between Bryozoa *sensu stricto* and that used by Nitsche in 1869, *i.e.* Bryozoa with the subdivision into Ectoprocta (*sensu* Bryozoa) and Entoprocta. However, the grouping of Bryozoa with Entoprocta is no longer accepted (Ryland 1970) and the most suitable and correct name for the phylum was suggested to be Bryozoa (Mayr 1968). Further molecular evidence in support of morphological data has shown that Entoprocta are unrelated to Bryozoa or even other lophophorates (Mackey *et al.* 1996). However, some authors (Nielsen 2001) are still substituting Ectoprocta with Bryozoa, and this still leads to some confusion (*e.g.* Giribet 2000).

The phylum Bryozoa traditionally has been subdivided into three classes: Gymnolaemata, Stenolaemata and Phylactolaemata, with the latter having a

distinctively different morphology: horseshoe-shaped lophophore and epistome present among other characters and its representatives being exclusively freshwater species. The affinity of Phylactolaemata with Bryozoa is uncertain and is disputed based on ontogenetic development (Jebram 1973) as well as their distinctive morphological characters and palaeontological record (Mundy *et al.* 1981). The problem is aggravated by the apparent scarcity of palaeontological data (Taylor and Larwood 1990). Also, recent studies based on molecular data although inconclusive about the origin of this class all separate Bryozoa *sensu stricto* and Phylactolaemata and place the latter class closer to Entoprocta (Giribet 2000, Mackey 1996, Glenner 2004).

Bryozoans are mostly marine representatives, all of class Stenolaemata and the majority of class Gymnolaemata. The latter class is the largest and most abundant group. Phylactolaemata are entirely represented by freshwater species having strong differences from the rest of Bryozoa are not reviewed here.

The class Gymnolaemata consists of around 650 genera[1] and over 3000 species (Ryland 1970) and subdivided into two orders: Ctenostomata and Cheilostomata. The former order characterised by the representatives which are not calcified, have chitinous exoskeleton and colonies which form either gelatinous sheets or branching networks of zooids. Zooids of Bryozoa of this order are cylindrical and without avicularia, the orifice being closed by a sphincter muscle. The order is further subdivided into two suborders, Stolonifera and Carnosa. The former order consists of eight families and includes such commonly found and abundant representatives as *Bowerbankia* (Figure 1A). Suborder Carnosa has ten families (nine represented in British fauna) and has such common representatives as *Flustrellidra hispida* and several *Alcyonidium* species.

[1] Currently a working Treatise (D.P.Gordon, personal communication) lists 1047 genera only for order Cheilostomata. This is reflective of the constantly undergoing changes in the nomenclature of many Bryozoa.

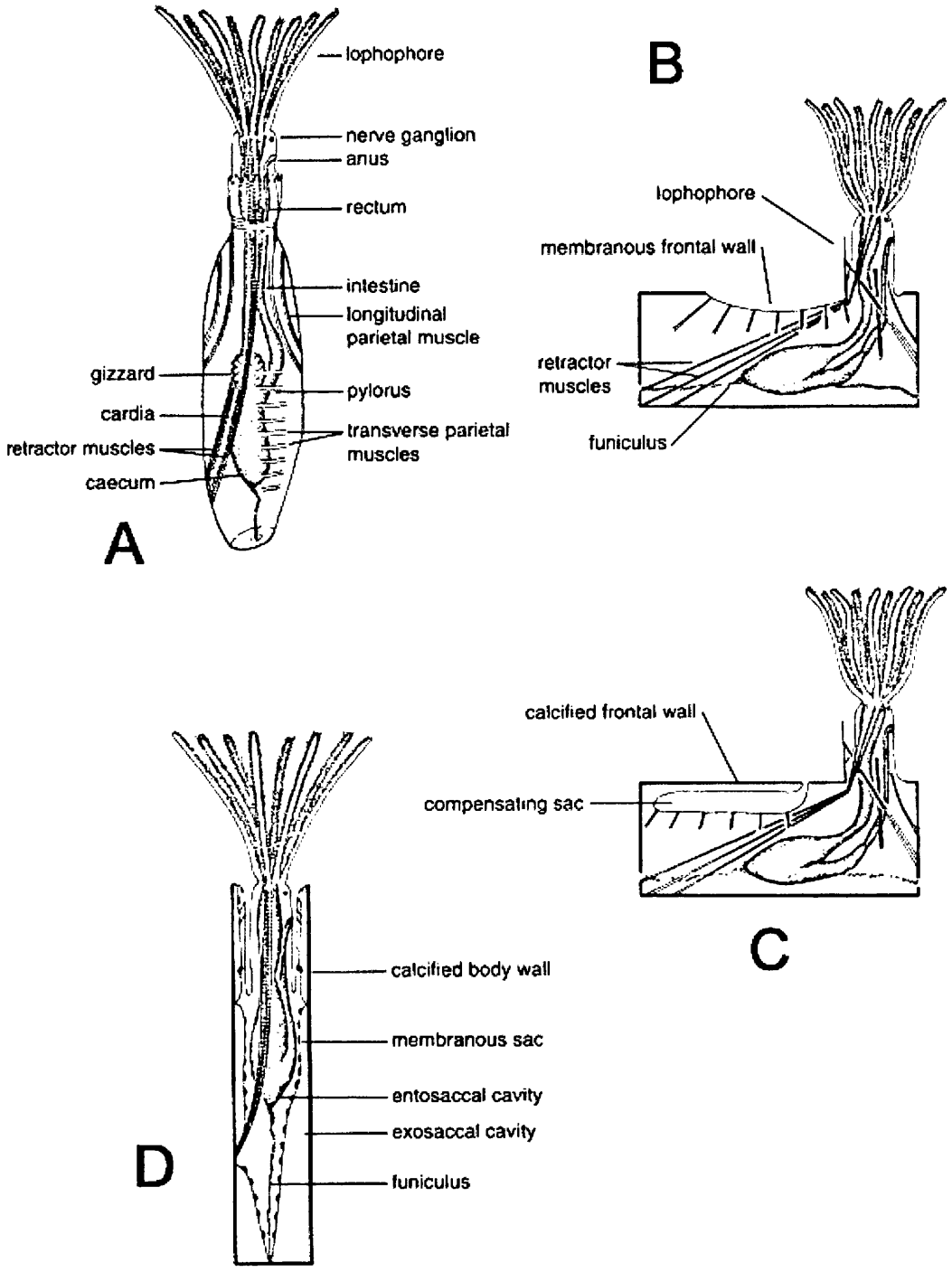


Figure 1 Different types of autozooids found in three orders of Bryozoa. A) Ctenostomate type of autozooid from Stolonifera. B) Anascan type of cheilostomate autozooid. C) Ascophoran type of cheilostomate autozooid. D) Cyclostomate type of autozooid (see text for details). Image taken from Hayward and Ryland (1998) Fig. 1.

Cheilostomata could be distinguished by box like zooids, which are always enclosed by walls calcified to varying degree (Figure 1B,C); this group shows the largest polymorphism among living Bryozoa. Another distinctive characteristic of this group is the presence of an operculum (a calcified hinged flap) although it is

secondary missing in some genera such as *Bugula*. Polymorphism of this group is largely attributed to the variation of the calcified frontal wall and zooid protrusion mechanism. Heterozooids differentiated by the zooid polymorphs avicularia and vibracula are distinctive of this group and develop from the modified shape of the autozooid due to the homologous change in the enlarged operculum. Kenozooids lack an orifice and operculum thus distinguishing them from the other two types of heterozooids.

The order Cheilostomata is subdivided into five suborders Inovicellina, Scrupariina, Malacostegina, Flustrina and Ascophorina. Ascophorina being further subdivided into infraorders Acanthostegomorpha, Hippothoomorpha and Umbonulomorpha and Lepraliomorpha (Hayward and Ryland 1999). The order had 40 families in Britain as of 1999, however as of 2007 the working Treatise on Cheilostomata lists over 140 families and 21 *incertae sedis* (D.P. Gordon and P.J. Hayward, personal communication).

Formerly Cheilostomata were subdivided into two suborders: Anasca and Ascophora. The division was based on the hydrostatic mechanism of eversion of the lophophore. Anasca have a soft frontal membrane in the cystid, which is responsible for the lophophore eversion (Figure 1B). The depression of the frontal membrane raises hydrostatic pressure of the coelom and everts the lophophore. In Ascophora on the other hand the frontal membrane is internalised beneath a solid wall and the process of lophophore eversion is controlled by the ascus (a sac), which fills with seawater as the lophophore everts (Figure 1C). However, it was recognised that the above division of Cheilostomata is not sufficient and further subgroups can be identified. These subdivisions can be based on the more detailed study of the frontal walls and the way in which the membrane is protected and were designated as the following suborders – Inovicellata, Scrupariina, Malacostega, Flustrina and Ascophora (P.G. Gordon, working Treatise personal communication). Despite the fact that the subdivision into two suborders (Anasca and Ascophora) is no longer recognised, it is still widely used in the literature and awareness of them is important.

The order Ctenostomata is smaller compared to Cheilostomata, represented by zooids with membranous or gelatinous walls, which are never calcified. The order is further subdivided into two suborders Carnosa and Stolonifera. The former is represented by such abundant and common species as *Alcyonidium* and

Flustrellidra whereas the suborder Stolonifera has among its representatives another common genus – *Bowerbankia*. There are 17 families and around 40 genera in Carnosa (Ryland 1970, Hayward 1985).

The suborder Stolonifera have heterozooids knows as kenozooids – these are stolon-forming zooids which lack many organs of autozooids. Suborder Carnosa on the other hand lacks stolon-forming kenozooids.

The class Stenolaemata is represented by five orders: Cystoporata, Trepostomata, Cryptostomata, Fenestrata and Cyclostomata the first four of which are fossil and the only extant order is Cyclostomata. The colonies of this group are characterised by tubular elongated autozooids, which are calcified (Figure 1D). The terminal membrane of the cyclostomes is functionally identical to that of the anascan frontal membrane of Cheilostomata. Some Cyclostomata (crisiids) form large bushy forms which are attached to the substrate via kenozooids simplified in their function which act as rhizoids. Apart from kenozooids, gonozooids and nanozooids, the latter being described only for a few genera, heterozooids are uncommon in cyclostomes. Cyclostomata are further subdivided into five[2] suborders Tubuliporina, Articulata, Cancellata, Cerioporina and Rectangulata (Taylor 2000). The order is represented by 9 families[3] (Hayward & Ryland 1985) and 250 genera (Ryland 1970, Hayward 1985) in British waters.

1.4 Evolution and palaeontology of Bryozoa

Bryozoa are believed to have originated in the lower Ordovician (approximately 480 mya) with the majority of taxa belonging to Stenolaemata[4] (Taylor and Larwood 1990, McKinney and Jackson 1989), see Figure 2.

[2] Ключе (1962) identifies another suborder: Isoporina, however due to missing diagnosis of this order in his original manuscript this order is disputed (Taylor 2000).
[3] The exact number of families is uncertain as Ключе (1962) for instance lists 11 families for the Cyclostomata found in the Russian northern seas.
[4] Boring Ctenostomata are believed to have originated at the same time as the rest of the Stenolaemata (personal communication with P.Taylor, reported in A. Waeschenbach 2003).

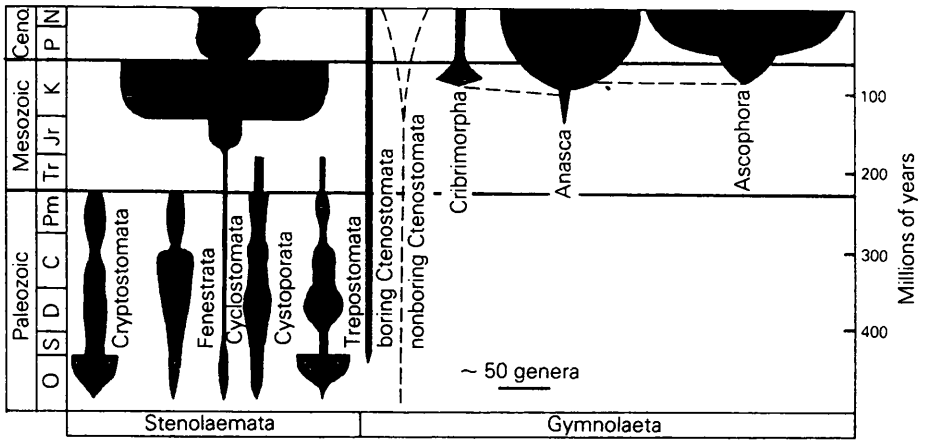


Figure 2 Geological histories of major bryozoan taxa. Diagram taken from McKinney and Jackson (1989), Fig 1.14. Additional notes about origin time of boring Ctenostomata are in the text.

What followed after that is what is called the Ordovician stenolaemate radiation (Taylor and Larwood 1990). The Ordovician radiation was the expansion of now mostly extinct stenolaemates of five orders (Cryptostomata, Fenestrata, Cystoporata, Trepostomata and Cyclostomata). Following the extinction of the first four orders in the Permian or Triassic periods[5] (approx. 200 mya), the surviving cyclostomes underwent another radiation in the mid-Mesozoic era (*i.e.* Jurassic and Cretaceous periods – around 140 mya). Since the late Cretaceous, Cheilostomata have become the dominant. Their massive radiation occurred in the mid-Cretaceous period and is referred to as the late Mesozoic cheilostome radiation (Taylor and Larwood 1990). The demise of the Cyclostomata group happened during and after the Cretaceous-Tertiary mass extinction event, the so-called K-T event, which took place 65 mya (MacLeod *et al.* 1997; Alvarez *et al.* 1980). Following the mid-Mesozoic cyclostome radiation the number of cyclostome genera recorded had reached 170, and then following the K–T event the number of genera declined to approximately 75 towards the late Palaeocene (McKinney and Taylor 2001), losing 79 cyclostome genera over the K-T boundary. The same fate was followed by cheilostomes with 81 genera in total being lost during the K-T event (McKinney and Taylor 2001). However the limited number of sites which have palaeontological records of Bryozoa on both sides of the K-T boundary has an impact on the

[5] Cryptostomata are now believed to have gone extinct at the same time as Cystoporata and Trepostomata, *i.e.* upper Triassic (personal communication with P.D. Taylor, reported in A. Waeschenbach 2003).

estimation of the full impact of this event on the taxonomic diversity of the group (McLeod *et al.* 1997).

1.5 Phylogenetic studies

The phylogenetic position of Bryozoa in relation to other phyla is not certain and relationships within a larger group, the Lophophorata have been a long debated issue (Halanych *et al.* 1995; Mackey *et al.* 1996; Zrzavy *et al.* 1998; Adoutte *et al.* 1999; Adoutte *et al.* 2000; Giribet *et al.* 2000; Hayward and Ryland 2000; Nielsen 2000; Nielsen 2001; Giribet 2002; Anderson *et al.* 2004). Traditionally, Bryozoa, Phoronida and Brachiopoda have been united into Lophophorata, based on their possession of a lophophore (Hyman 1959). This classification is based essentially on one character, the lophophore. Some authors (Nielsen 2001) argued for the unification of the Entoprocta and Ectoprocta in a superphylum under one name – Bryozoa.

In a phylogenetic study of Lophophorata and other Metazoa using 18S rRNA gene (Mackey *et al.* 1996), Entoprocta and Ectoprocta were separated into two clades, and were shown not to be sister taxa as was believed by Nielsen (2001). Further, in a large study of triploblastic taxa based again on the 18S rRNA data, Gymnolaemata and Stenolaemata were separated both from Phylactolaemata and Entoprocta, and the location of lophophorates was shown to lie between Protostomia and Deuterostomia, further introducing uncertainty (Giribet *et al.* 2000). In a more recent study (Glenner *et al.* 2004) using 18S rRNA data and Bayesian methodology the relationships within the Lophotrochozoa were evaluated. The Lophotrochozoa was recently created based on 18S rRNA data analysis, and encompasses Lophophorata and molluscs and annelids (Halanych 1995). The study of Glenner *et al.* (2004) showed a clear separation of Lophotrochozoa as a distinct group; in addition Gymnolaemata and Stenolaemata were in a separate clade from Phylactolaemata and Entoprocta.

The studies related to the phylogenetic relationship of Bryozoa are unfortunately limited and controversial in their findings. The relationships and the complexity of the orders within Bryozoa can be clearly seen from the constantly changing systematics – such as removal of the original Anasca-Ascophora grouping in Cheilostomata as well as changing number of families and genera in the group (see above).

Several studies were performed recently, some using molecular phylogenetic analysis to examine relationships within Bryozoa. The main characteristic used for the taxonomy of Bryozoa is the structure of their cystid or in other words their skeleton. This is true for both fossil and extant species.

Todd (2000) specifically noted that despite the fact that morphological characters are readily available because of the highly skeletised nature of Bryozoa, the systematics of Bryozoa is poorly understood, and that, at the time (2000) there were only two reports which used computerised (cladistic) studies of Bryozoa. However, these studies were criticised (Todd 2000) as lacking data matrices and thus hard to evaluate. The findings of study based on the combined information from fossil and extant species found Ctenostomata to be paraphyletic, with Stenolaemata and Cheilostomata nesting within the Ctenostomata (Todd 2000). At the same time as Todd's (2000) findings, an examination of the phylogeny of Cheilostomata was undertaken using information derived from frontal wall structure (Gordon 2000). As a result nine possible models for the evolution of ascophorans were presented.

Both Todd (2000) and Gordon (2000) recognised the necessity and urgency of conducting molecular based studies dedicated to Bryozoa. The situation with the interrelationship within Bryozoa was further complicated by Dick *et al.* (2000), which was the first molecular study of Bryozoa, using the data from mitochondrial 16S rRNA. In that study Ctenostomata and Cheilostomata showed paraphyly whereas Cyclostomata showed polyphyly. The validity of the findings of that study are discussed further in this work, in particular DNA sequence alignment methodology. The authors (Dick *et al.* 2000) themselves cautioned about the use of 16S rRNA gene as the suitability of this gene in phylogenetic studies is limited by its ability to resolve divergences only as far back as mid-Cretaceous, which can be insufficient given the palaeontological record of Bryozoa.

One more molecular study of Bryozoa which recently appeared is that of Hao *et al.* (2005). This study re-evaluated Cheilostomata phylogenetic relationships based on the 16S rRNA gene. However, the methods used by these authors are

questionable and the validity of their findings is further discussed in related chapters[6] of this work.

1.6 Ribosomal RNA and nuclear 18S gene

Ribosomal RNAs (rRNA) are among the building blocks of the ribosomes, which are responsible for the protein synthesis in cells. Each eukaryotic ribosome consists of two subunits: small subunit (SSU) 40S[7] and large subunit (LSU) 60S (Figure 3). 18S rRNA is located in the SSU of the ribosome.

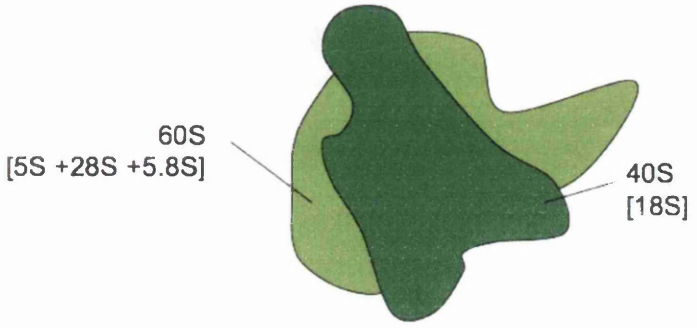


Figure 3 Ribosome subunits and their corresponding rRNAs. 60S is a LSU formed by 5S 28S and 5.8S rRNAs and several proteins, 40S is a SSU formed by 18S rRNA and several proteins.

rRNAs are synthesised as a large precursor unit in eukaryotes. Each unit contains 18S, 5.8S and 28S rRNAs as well as two internal spacers (ITS-1, ITS-2) and one external transcribed spacer (ETS), which are spliced out during ribosome synthesis (Figure 4). These units are referred to as rDNA operon, and in eukaryotes they are repeated as multiple tandems throughout the genome.

[6] Each chapter in this study deals with slightly different topics where discussion of relevant studies is undertaken. For instance in case of Hao *et al.* (2005) sequence alignment issues are discussed in Chapter 4, whereas method of DNA extraction is discussed in Chapter 2.
[7] S in the name of the subunit stands for Svedberg – a non-SI unit of particle sedimentation.

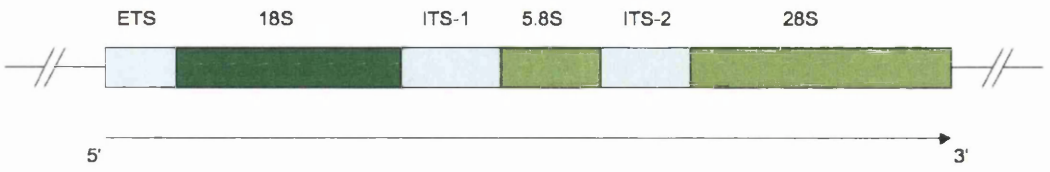


Figure 4 An rDNA operon of eukaryotes, containing 18S, 5.8S and 28S rRNAs and ETS and two ITS (see text). 18S rRNA (shown in dark green) once processed comprises part of 40S SSU of the ribosomes (together with proteins).

Each rRNA is folded into secondary and tertiary structure prior to its integration into the ribosomes. These rRNAs are highly conserved in all organisms as both SSU and LSU have regions of high conservation, which are responsible for the function of the ribosome. Each rRNA gene also has highly variable regions (usually corresponding to the loops of the secondary structure) and highly conserved regions, which are often represented by the stems of the secondary structure and more complex tertiary units of organisation.

The length of the 18S rRNA is in general considered to be 1800 bp, however some studies have shown great extensions in the hypervariable regions and 18S was described as long as 2469 bp for some aphids (Kwon *et al.* 1991) and 2864 bp for parasitic insects (Gillespie *et al* 2005).

One of the main advantages of molecular methods for phylogenetic research is the extensive data sets of independent characters, theoretically limited by the number of nucleotides in the gene which is used in the study. Also of advantage is the information character uniformity among all living organisms (*i.e.* genetic code) and the fact that genetic code variation is always inheritable (Hillis 1987). The advantages of using ribosomal RNA (rRNA) genes and in particular 18S rRNA are many and were noted and used very early for reconstructions of animal phylogeny. For instance, 18S rRNA sequences were used for the first ever molecular phylogenetic study (Field *et al.* 1988) dedicated to the origin of the Metazoa.

Because of the varying substitution rates 18S rRNA sequences can be used to resolve deep phylogenies as far back as Precambrian (Hillis 1991) as well as more closely related organisms (Olsen and Woese 1993). Thus rRNA stores information relevant to both recently evolved taxa as well as those that have evolved a long time ago.

However, the different rates of substitution (as high as 10-fold) of 18S rRNA due to the presence of highly variable regions (see Figure 5) presents an additional problem associated with possible dilution of the amount of phylogenetic information. The substitutional saturation of highly evolving regions can cause loss of resolution and contribute noise, thus biasing the results of the tree reconstruction when parsimony and distance based methods are used (Abouheif and Meyer 1998).

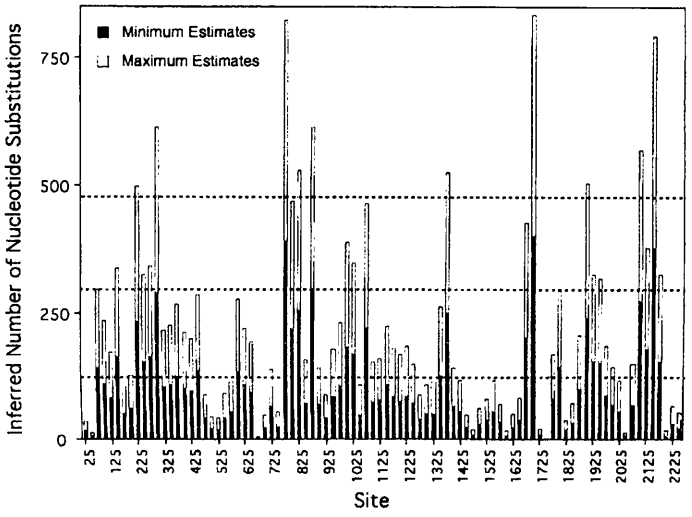


Figure 5 Estimated number of nucleotide substitutions per site of the 18S rRNA, calculated for each 25 bp. Maximum (black) and minimum (white) estimates are shown. Image taken and modified from Abouheif and Meyer (1998), Fig 1.

This issue is discussed in great detail in the Chapter 5, which deals with the alignment and secondary structure of the bryozoan 18S.

rDNA[8] is present in the genome in multiple copies and these copies were shown to be evolving in such a way that homogenisation of information occurs between the copies and hence called "concerted evolution" (Hillis 1991). This gives a special advantage to the phylogenetic studies of closer related species as no intraspecific variation is present and thus smaller sample sizes can be used.

GenBank has accumulated a vast number of 18S sequences for many organisms and has several 18S rRNA sequences of Bryozoa. Unfortunately, the validity of some of the bryozoan sequences deposited in GenBank is questionable (see Chapter 3 for detailed discussion) and therefore not all sequences can be used for phylogenetic reconstruction.

[8] rDNA is a term which refers collectively to the entire set of rRNA genes and their spacers (Hillis 1991), also see Figure 4.

Secondary structure has a direct impact on the rates of substitution in the different regions of rRNA and thus requires special treatment of rRNA sequences and their alignments. The stems of rRNA evolve much slower compared to the loops and bulges which are evolving more freely. This issue has a big impact on the alignment of the 18S sequences and thus is discussed in a separate chapter.

Highly conserved regions of rRNA are suggested (Hillis 1991) to be of great aid in designing so-called universal primers, *i.e.* oligonucleotides which could be suitable for amplification of the 18S gene from a diverse group of organisms. In the bryozoan context it indeed would be an ideal situation if one set of primers could be used for all bryozoan species. Unfortunately, universal primers do not appear to work well with all bryozoan species and thus a larger set of primers is required. The issue of primers development is discuss at length in Chapter 3.

1.7 Aims

This study has several objectives. Firstly, development of working sets of oligonucleotide primers for the 18S rRNA gene, which could be used to collect sequence information from as wide a number of species of Bryozoa as possible. These will include representatives from the orders Cheilostomata, Ctenostomata and Cyclostomata.

The second objective is based on the acquired 18S rRNA sequences and possibly some 18S sequences from the NCBI GenBank database to build a working phylogeny of Bryozoa. Whilst performing the analysis of the sequences obtained here an evaluation of the secondary rRNA structure will be performed and incorporated into the multiple alignment of the sequences.

In addition to molecular phylogenetic work, an evaluation of a microscopy method based on the confocal laser microscopy will be performed in the hope that this method could be used in the future for assigning larval types based on the system proposed by Zimmer and Woollacott (1977).

2 MORPHOLOGICAL METHODS

2.1 Sample collection general observations

In this chapter a general overview of the sampling procedure is given including location of sampling sites, why they were chosen and how the sampling was done. During several seasons lasting from October 2003 until March 2006, sample collection was performed on a regular basis and whenever possible, depending only on the weather conditions and tides. The sites were visited at low tide only, data about which was taken in advance from the Admiralty Tidal Tables published annually for the UK by The UK Hydrographic Office. These tables offer low/high tide readings for any given date for Milford Haven (major sea traffic point). However, to acquire a more precise time reading for the desired location an electronic version of the tables[9] was used, which allowed specification of a precise location. On average two sufficiently low water tides occur every month and it was planned to visit at least one site at each low tide and sometimes more if timing between tides allowed. Low water spring tides (LWST) allowed exceptional access to the infralittoral zone of some sites such as Watwick Bay. Site-specific collection procedures are described below but in general colonies were picked up together with the rocks on which they were found, or detached from the substrata if possible.

Because of the relatively large distance between sites (see map on Figure 6 below) it was not possible to sample every location every time within a short low tide time and therefore different sites were visited on a simple rota basis. However, in some cases a specific target sampling was performed when it was known that certain species could be found at a specific location. For instance, *Crisia* species were mainly found in the Watwick Bay, and *Alcyonidium* species were abundant at the Pembroke Ferry site.

All sampling sites were reached by car and samples once collected returned to the laboratory for further analysis. Samples were always transported in a temperature insulated container to minimise temperature shock to the colonies.

[9] The electronic version of the tidal tables provided by the UK Admiralty and can be accessed for free on http://easytide.ukho.gov.uk/ website.

2.2 Sampling locations and physical environment

Sampling sites (Figure 6) were selected throughout South West Wales from Swansea Bay towards Dale and Skomer Island based on the previous records in the literature, mainly from the Synopses of the British Fauna (Hayward and Ryland 1979, 1985; Ryland and Hayward 1977; Hayward 1985). In addition a survey review of local sampling sites related to Bryozoa was previously presented by Porter (1999). Also personal communication and consultations were done with Dr PJ Hayward about many species locations. Three locations (Watwick Bay, Lydstep, Pembroke Ferry) were selected as the main sites as they were known to have numerous representatives of Bryozoa and visited on a regular basis. In addition, some other sites were visited on a one-off basis (see below for detailed description) and on four occasions trawling was undertaken with the R/V Noctiluca, which belongs to Swansea University. Boat trips were done specifically in order to find reproducing *Flustra foliacea* from the locations previously known to have this species (PJ Hayward, personal communication).

The general conditions of the coastline around Milford Haven could be described as rocky or stony with many cliffs and some eroded reefs (Nelson-Smith 1965). The region is bathed by the Atlantic ocean water coming from the Arctic and southern region and supports a very varied fauna for the British Isles (Nelson-Smith 1965). Spring tides in Milford Haven have a mean range of 6 metres (Nelson-Smith 1965) and can reach up to 8 metres during LWST. The time of the lowest tide during the day is around 13:00 hrs GMT which allowed a very consistent access to the area for the sampling because of the daylight. During the equinoctial low water spring tides exceptional access to the infralittoral zone was possible at most sites. The salinity of the coastal waters is between 34 ‰ and 34.6 ‰ (Nelson-Smith 1965). Water temperature varied from approximately 8°C in winter months to around 13°C in summer with slightly higher temperatures in the littoral zone in summer.

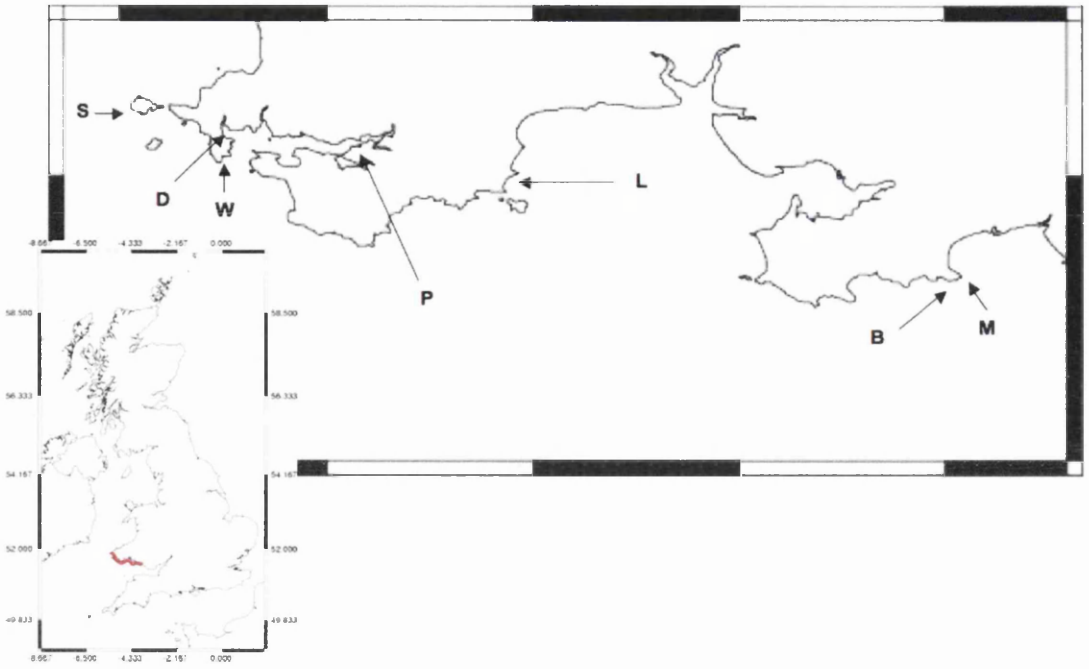


Figure 6 Map showing location of the sampling sites. S – Skomer Island, D – Dale Harbour, W – Watwick Bay, P – Pembroke Ferry, L – Lydstep Bay, B – Bracelet Bay, M – Mumbles Pier. For more detailed location positions please see maps below. Map generated using Coastal Extractor[10].

2.3 Sampling site descriptions

2.3.1 Lydstep Bay (Carmarthen Bay)

This site (Figure 7) is situated near Tenby town and forms part of the Lydstep Leisure Centre. It is a large sandy beach with many tourist activities including water sports and is affected by sea traffic of the nearby Ferry line during the summer season. The southeastern part of the beach is rocky with large boulders overhanging from the shore and extends towards Giltar Point (the furthest point of the SE shore, see Figure 7) in the direction of which most of the sample collection was done. During LWST it was possible to reach Giltar Point where under the large boulders, exposed during these tides, many Crisiidae colonies could be found. This site had an abundant bryozoan fauna with 35 species found belonging to all three bryozoan orders.

[10] Coastal Extractor is a free online Java tool (National Geographical Centre) which allows generation of world coastal maps: http://rimmer.ngdc.noaa.gov/coast/

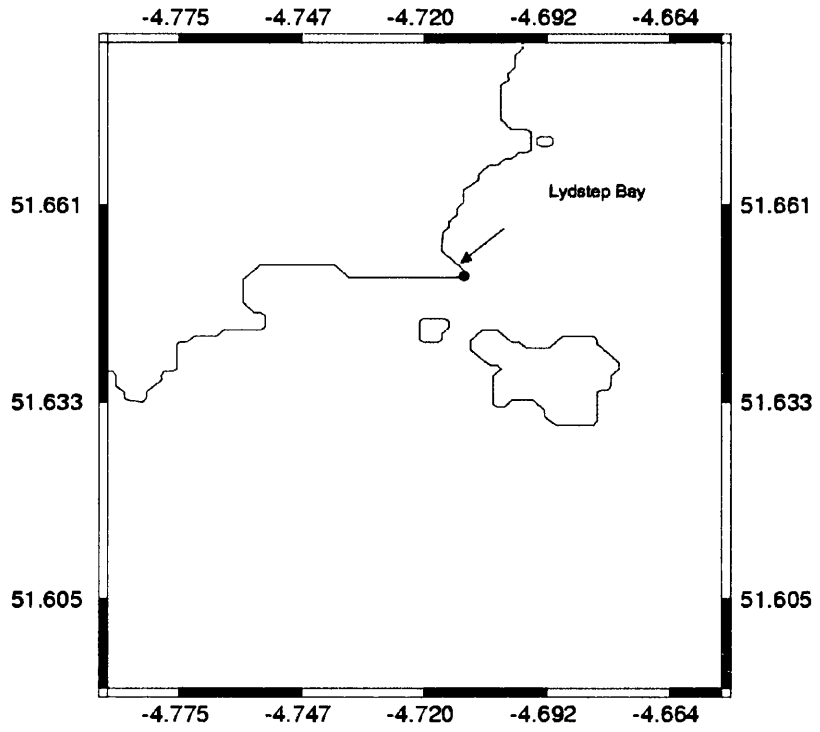


Figure 7 Lydstep Bay sampling point location. Precise position of sampling is indicated by the arrow. Red point indicates Giltar Point. Map generated using Coastal Extractor.

2.3.2 Dale Harbour

This location (Figure 8) was in the village of Dale in a small rocky shore with many macroalgae present (such as *Fucus serratus* and *Fucus vesiculosus*). This site was well exposed during low tides and mainly *Alcyonidium* spp. and *Flustrellidra hispida* were collected from here.

2.3.3 Watwick Bay (Milford Haven)

Further out from Dale Harbour towards St. Ann's Head a relatively small lagoon is situated – Watwick Bay (Figure 8). This is a sandy shore with rocky sides and very little tourist activity. The southeastern side of the shore was used for sampling, with lowest tides giving access to the West Blockhouse Point. As with Lydstep Bay this was one of the most visited sites with 37 bryozoan species recorded from three orders. When the LWST zone was exposed at the furthest SE point of this shore many Crisiidae species could be collected at this site under very large overhanging rocks and boulders. This site is relatively dangerous during higher tides because of many algae covering underwater rocks, and the need to walk on them far out toward

the West Blockhouse point, therefore it is best visited with a companion and in summer, light footwear is preferable.

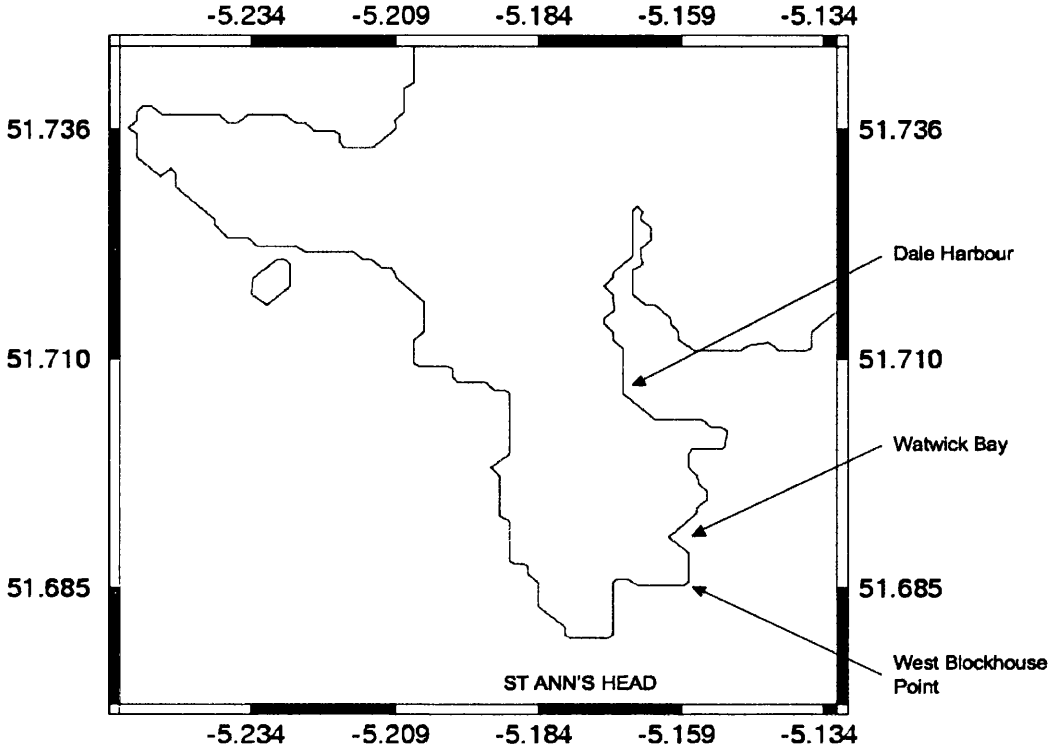


Figure 8 Dale Harbour and Watwick Bay sampling sites location. Precise positions of sampling are indicated by the arrows. Map generated using Coastal Extractor.

2.3.4 Pembroke Ferry (Cleddau bridge, Milford Haven)

This site is located in the estuary of the Daucleddau river (Figure 9). The exact location is underneath Cleddau bridge at the site of the old ferry connection which no longer exists. This site has a small reef directly under the bridge, which can be reached through the rocky shore during low tidess. This reef has many small stones encrusted by algae, sponges and many *Alcyonidium* species. This site was visited mainly to collect *Alcyonidium* species, also *Scruparia chelata* was found here in reproducing stage, growing on *Alcyonidium hirsutum* colony.

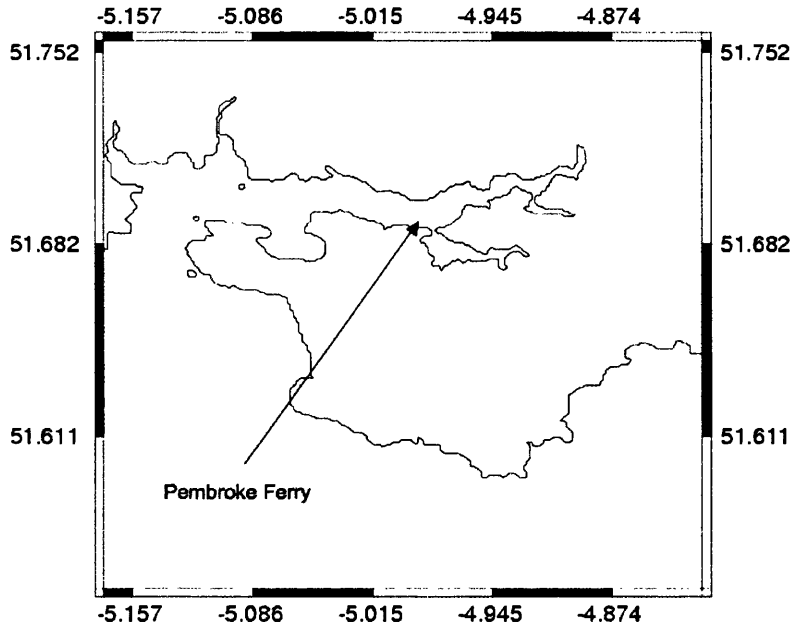


Figure 9 Pembroke Ferry (Cleddau Bridge) sampling site location. Precise position of sampling is indicated by the arrow. Map generated using Coastal Extractor.

2.3.5 Mumbles pier (Swansea Bay)

This site is located near Swansea University, under the Mumbles pier and also under the lifeboat station (marked as "M" on the main map, Figure 6). It is fully exposed and easily reachable during low tides. The rocky shore is mixed with some sandy patches. The actual sampling was done along an old pipe (diameter approx 70 cm) running alongside the shore. This pipe acts as a reef and substratum for many algae, especially *Fucus serratus.* Hydroids and bryozoans are attached to the algae and the pipe directly. In particular, this site was visited to collect several species of Ctenostomata such as *Bowerbankia*, *Alcyonidium*, *Crisia*, *Walkeria uva* and *Flustrellidra hispida.* In addition to the pipe some sampling was done alongside the pier pillars which are covered by many *Mytilus edulis* banks and many sponges. The pillars had previously had sightings of the ctenostome *Anguinella palmata* (PJ Hayward, personal communication). Unfortunately, no specimens of this species were found.

2.3.6 Other sampling sites

Several other sampling sites were visited on a one off basis, namely Bracelet Bay (Figure 6, marked as "B"). This location was visited in particular to collect *Crisia klugei.* This is a rocky shore with many littoral rockpools just to the west of

Mumbles Head. Also, in an attempt to find reproducing *Flustra foliacea,* several boat trips were done to the locations off Swansea Bay alongshore towards Oxwich Point, in particular "White Oyster Ledge" (Figure 10) was sampled (by an otter trawl). In addition, several species of Bryozoa were collected by Dr J. Porter from Skomer Island (Figure 6, marked as "S" on the map) by scuba diving.

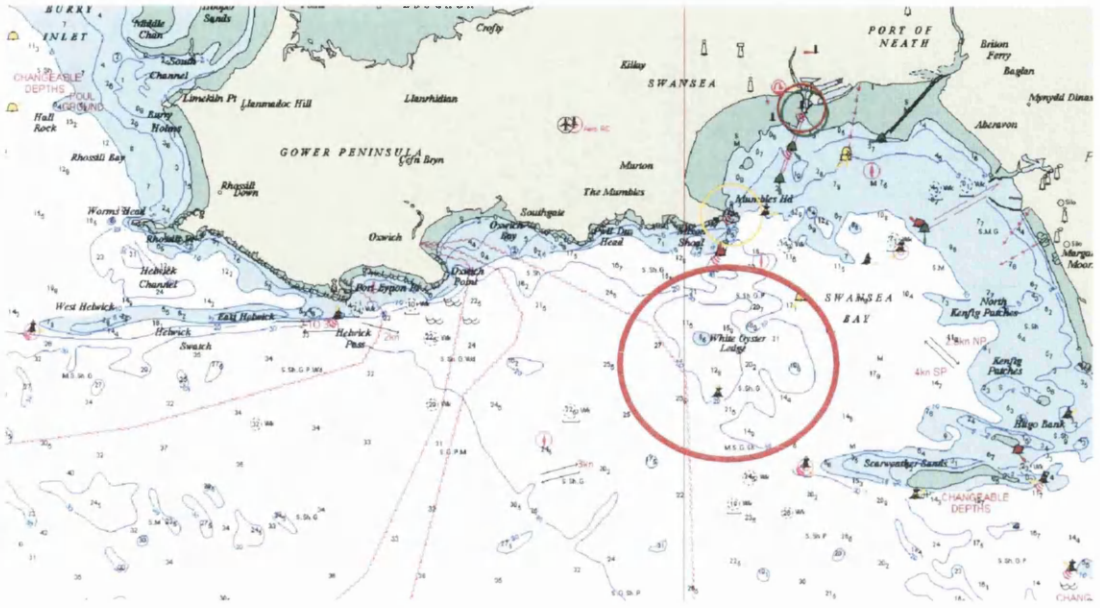


Figure 10 Location of the boat sampling. White Oyster Ledge is shown by the red circle. See text for details. Map generated using Admiralty chart.

2.4 Sample handling and identification

Once samples were collected and transferred to the laboratory they were analysed within the shortest possible period. All live specimens were stored in the temperature controlled (CT) room adjusted to the current sea temperature, based on the data from the National Data Buoy Centre[11]. The room had a permanent supply of sea water, part of the general laboratory supply from Swansea Bay for the School of Biological Sciences. Several tanks were built for this purpose with constant air supply and the possibility to change water regularly. In addition, a dark tank (light tight) was built to store those colonies containing mature larvae for the purpose of live larval released to be used in the confocal microscopy experiments (see below). This tank was stored in the same CT room with separate water and air supply.

[11] The network of weather marine buoys is maintained by the US National Oceanic and Atmospheric Administration. This network encompasses weather buoys from all over the world including those maintained by the UK Met Office. The readings for this study were taken from the buoy located near Milford Haven.

Live specimens were identified using appropriate Synopses of the British Fauna and sometimes, in ambiguous cases, consultations were done with Prof J Ryland and Dr PJ Hayward. Identification was done using the Olympus Stereo microscope (SZ60) and cold fibre optics lighting to minimise the temperature shock to living colonies. Once identified, samples were logged into a database, which recorded date of sample collection, location, colony description and the reproductive stage of the colony and any special notes on the substrate they were found on. Subsequently DNA was extracted from the colony using the technique described in Chapter 3.

The above mentioned database contains nearly 300 records of all sighting events. A summary table of the database showing each species recorded based on the location and month is presented here (Table 1). In addition a summary of the reproductive cycle, as recorded, is given in Table 2. This latter table only lists those species that had embryos at any developmental stage during the period from October 2003 until March 2006. Whilst the Synopses of the British Fauna have extensive information about reproductive cycles and species distributions, it is hoped that this table can add information about the breeding cycle for those species for which information is imprecisely known or missing.

In total 42 species were recorded as reproducing and DNA was extracted from these species. However, 18S rRNA sequences were not obtained for all of these species mainly because of insufficient DNA extracted (sometimes only one embryo was available). There were also problems with oligonucleotide primers. For instance *Alcyonidium diaphanum, Bugula neritina, Celleporella hyalina, Chorizopora brongniartii, Omalosecosa ramulosa* only gave a few embryos which could be extracted and given the fact that for many species primers had to be optimised to work with polymerase chain reaction (see Chapter 3 for a detailed discussion) these species were either not sequenced at all or only partial sequences were obtained. One species *Omalosecosa ramulosa* was only sighted and collected from the Skomer Island (by Dr J Porter) and no further sightings of this species was made during the regular sampling trips to the sites described above. Another example is *Celleporella hyalina.* Because of the difficulty of obtaining the sequences and lack of DNA material due to few samplings it was only possible to obtain a partial 18S sequence of this species. Finally, DNA extraction was not done

successfully for some species even though they were reproducing due to small size of the embryos, which were lost during the extraction process.

Table 1 Sampling sites and reproduction information of Bryozoa collected in South Wales. The species names are in alphabetical order. The last column (Repr.) indicates if the species sample was in a reproducing state (*i.e.* embryos/larvae were observed).

Order	Family	Species	Location	Collection Month	Repr.
Cheilostomata	Aeteidae	*Aetea anguina*	Skomer Isl.	August	
			Watwick Bay	August	
	Calloporidae	*Amphiblestrum auritum*	Lydstep Bay	October	
	Bugulidae	*Bicellariella ciliata*	Lydstep Bay	June	yes
	Bugulidae	*Bugula fulva*	Lydstep Bay	June	yes
			Watwick Bay	June	yes
	Bugulidae	*Bugula neritina*	Lydstep Bay	August	
			Mumbles Pier	June	yes
	Bugulidae	*Bugula plumosa*	Lydstep Bay	June	yes
			Mumbles Pier	June	yes
	Bugulidae	*Bugula turbinata*	Lydstep Bay	June	yes
	Calloporidae	*Callopora dumerilii*	Lydstep Bay	May	yes
			Watwick Bay	September	yes
	Calloporidae	*Callopora lineata*	Skomer Isl.	August	
			Lydstep Bay	May	yes
				May	yes
				August	yes
			Watwick Bay	May	yes
	Calloporidae	*Callopora rylandi*	Lydstep Bay	January	yes
				May	yes
				June	yes
			Watwick Bay	June	yes
				September	yes
				May	yes
				August	yes
	Calloporidae	*Cauloramphus spinifera*	Watwick Bay	March	yes
	Cellariidae	*Cellaria fistulosa*	Skomer Isl.	August	
	Celleporidae	*Cellepora pumicosa*	Lydstep Bay	August	
	Hippothoidae	*Celleporella hyalina*	Lydstep Bay	March	yes
				May	yes
			Mumbles Pier	August	
			Pembroke Ferry	September	yes
			Watwick Bay	January	yes
	Celleporidae	*Celleporina hassallii*	Lydstep Bay	June	yes
			Watwick Bay	June	yes

Order	Family	Species	Location	Collection Month	Repr.
				September	yes
				May	yes
	Chorizoporidae	*Chorizopora brongniartii*	Watwick Bay	October	yes
	Electridae	*Conopeum reticulum*	Lydstep Bay	March	
	Cribrilinidae	*Cribrilina cryptooecium*	Skomer Isl.	August	
			Lydstep Bay	March	yes
			Watwick Bay	November	yes
				February	yes
				October	yes
				January	yes
	Hippoporinidae	*Cryptosula pallasiana*	Lydstep Bay	October	
			Pembroke Ferry	September	
	Electridae	*Electra pilosa*	Lydstep Bay	October	
	Escharellidae	*Escharella immersa*	Skomer Isl.	August	
			Lydstep Bay	January	yes
				March	yes
				April	yes
				May	yes
				March	yes
			Pembroke Ferry	September	yes
			Watwick Bay	February	yes
				September	yes
				March	yes
				May	yes
				August	yes
	Escharellidae	*Escharella variolosa*	Watwick Bay	June	
	Exochellidae	*Escharoides coccinea*	Lydstep Bay	January	yes
				March	yes
				April	yes
				May	yes
				August	yes
			Watwick Bay	November	yes
				February	yes
				March	yes
				September	yes
				May	yes
				August	yes
	Microporellidae	*Fenestrulina malusii*	Skomer Isl.	August	
	Flustridae	*Flustra foliacea*	White Oyster Ledge	October	

Order	Family	Species	Location	Collection Month	Repr.
			Mumbles Pier	October	yes
	Hippothoidae	*Haplopoma graniferum*	Skomer Isl.	August	yes
			Lydstep Bay	September	yes
				March	yes
	Membraniporidae	*Membranipora membranacea*	Watwick Bay	March	yes
	Microporellidae	*Microporella ciliata*	Lydstep Bay	October	yes
				March	yes
				May	yes
				August	yes
			Watwick Bay	March	
	Celleporidae	*Omalosecosa ramulosa*	Skomer Isl.	August	yes
				September	yes
	Escharinidae	*Phaeostachys spinifera*	Pembroke Ferry	September	
			Watwick Bay	February	yes
				March	yes
				June	yes
				May	yes
	Bitectiporidae	*Schizomavella linearis*	Lydstep Bay	October	yes
				May	yes
				August	yes
			Watwick Bay	November	yes
				September	yes
				February	yes
	Schizoporellidae	*S.linearis var. hastata*	Lydstep Bay	October	yes
			Watwick Bay	September	yes
	Scruparíidae	*Scruparia ambigua*	Lydstep Bay	March	
	Scruparíidae	*Scruparia chelata*	Skomer Isl.	August	
			Boat Collection	August	
			Lydstep Bay	May	
			Mumbles Pier	September	
			Pembroke Ferry	June	yes
			Watwick Bay	May	
	Scrupocellariidae	*Scrupocellaria reptans*	Lydstep Bay	June	yes
			Watwick Bay	August	
	Umbonulidae	*Umbonula littoralis*	Lydstep Bay	January	yes
			Watwick Bay	November	yes
				March	yes
				September	yes

Order	Family	Species	Location	Collection Month	Repr.
				August	yes
				January	yes
Ctenostomata	Alcyonidiidae	*Alcyonidium diaphanum*	Pembroke Ferry	September	yes
			Watwick Bay	September	
	Alcyonidiidae	*Alcyonidium gelatinosum*	Skomer Isl.	August	
			Dale Harbour	January	yes
			Lydstep Bay	May	
			Pembroke Ferry	September	yes
			Watwick Bay	September	yes
				March	yes
	Alcyonidiidae	*Alcyonidium hirsutum*	Dale Harbour	January	yes
			Lydstep Bay	May	
			Mumbles Pier	February	
			Pembroke Ferry	September	yes
			Watwick Bay	March	yes
				January	yes
	Alcyonidiidae	*Alcyonidium mytili*	Watwick Bay	September	
	Alcyonidiidae	*Alcyonidium polyoum*	Mumbles Pier	June	yes
			Pembroke Ferry	September	yes
				September	yes
			Watwick Bay	February	
	Vesiculariidae	*Bowerbankia citrina*	Mumbles Pier	June	yes
	Vesiculariidae	*Bowerbankia gracilis*	Lydstep Bay	May	
			Mumbles Pier	August	yes
			Pembroke Ferry	September	
	Vesiculariidae	*Bowerbankia imbricata*	Skomer Isl.	August	
			Mumbles Pier	June	yes
	Flustrellidridae	*Flustrellidra hispida*	Bracelet Bay	March	
			Dale Harbour	February	yes
			Lydstep Bay	May	yes
			Mumbles Pier	February	yes
				June	yes
				November	yes
			Watwick Bay	March	yes
				May	yes
	Walkeriidae	*Walkeria uva*	Bracelet Bay	March	yes
			Mumbles Pier	September	
			Watwick Bay	September	
Cyclostomata	Crisiidae	*Crisia aculeata*	Mumbles Pier	November	

Order	Family	Species	Location	Collection Month	Repr.
			Watwick Bay	August	yes
	Crisiidae	*Crisia cornuta*	Watwick Bay	March	
	Crisiidae	*Crisia denticulata*	Skomer Isl.	August	yes
			Bracelet Bay	March	yes
			Lydstep Bay	September	yes
			Watwick Bay	September	yes
				March	yes
				August	yes
				January	yes
	Crisiidae	*Crisia eburnea*	Skomer Isl.	August	yes
			Watwick Bay	March	yes
	Crisiidae	*Crisia klugei*	Bracelet Bay	March	
	Crisiidae	*Crisidia cornuta*	Lydstep Bay	June	
			Mumbles Pier	November	
			Watwick Bay	August	
	Crisiidae	*Filicrisia geniculata*	Lydstep Bay	June	yes
			Watwick Bay	August	
	Tubuliporidae	*Tubulipora liliacea*	Pembroke Ferry	June	yes
			Watwick Bay	August	
	Tubuliporidae	*Tubulipora* sp.	Skomer Isl.	August	
			Lydstep Bay	January	
			Watwick Bay	October	
			Mumbles Pier	February	
			Dale Harbour	February	yes

Table 2 Species which were found reproducing during the seasons from October 2003 till march 2006. The crosses indicate that the given species was recorded as reproducing at this month. Blanks indicate that the species was either not reproducing or not found during this month. See text for details.

Species	MONTHS OF THE YEAR											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Alcyonidium diaphanum									X			
Alcyonidium gelatinosum		X	X						X	X		
Alcyonidium hirsutum		X	X						X			
Alcyonidium polyoum						X			X			
Bicellariella ciliata						X						
Bowerbankia citrina						X						
Bowerbankia gracilis								X				
Bowerbankia imbricata						X						
Bugula fulva						X						
Bugula neritina						X						
Bugula plumosa						X						
Bugula turbinata						X						
Callopora dumerilii					X				X			
Callopora lineata					X			X				
Callopora rylandi	X				X	X		X	X			
Cauloramphus spinifera			X									
Celleporella hyalina		X	X		X				X			
Celleporina hassallii					X	X			X			
Chorizopora brongniartii										X		
Cribrilina cryptooecium		X		X						X	X	
Crisia aculeata								X	X			
Crisia denticulata		X	X					X	X			
Crisia eburnea			X					X				
Crisidia cornuta		X										
Escharella immersa	X	X	X	X	X			X	X	X		
Escharoides coccinea	X	X	X	X	X			X	X		X	
Filicrisia geniculata						X						
Flustra foliacea		X								X	X	
Flustrellidra hispida		X	X		X	X					X	
Haplopoma graniferum				X				X	X			
Membranipora membranacea			X									
Microporella ciliata			X		X			X		X		
Omalosecosa ramulosa									X	X		
Phaeostachys spinifera		X	X		X	X						
Schizomavella linearis		X			X			X	X	X	X	
Schizomavella linearis var hastata									X	X		

Scruparia chelata						X						
Scrupocellaria reptans						X						
Tubulipora liliacea						X						
Umbonula littoralis		X	X	X				X	X		X	
Walkeria uva			X									
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec

3 MOLECULAR METHODS AND PRIMER DESIGN

This chapter is dedicated to the general methods, which were employed during the molecular part of the project as well as specifically to the primer design and optimisation. The process of primer optimisation took almost two years of this project and played one of the main parts in it. An original attempt to use the so-called "universal" 18S rRNA primers (Halanych 1995) failed. The primers thus were optimised on a per-species basis, which resulted in some delays due to the lack of DNA material. Specific attention here will be given to the issue of non-specific primers and, as a result, potential contamination of the DNA sequences and its affect on the 18S bryozoan sequences which have been submitted to the NCBI database.

3.1 DNA extraction

Many bryozoan related molecular studies (Hao *et al.* 2005; Dick *et al.* 2000; Makey et al. 1996; Giribert *et al.* 2000; Passamaneck and Halanych 2006) and studies related to many marine invertebrates use tissue DNA extraction methods. Although these methods allow extraction of a large amount of target DNA they have a potential problem related to contamination through the seawater as a carrier of bacteria, phytoplankton and other microorganisms as well as epi- or symbiotic organisms such as bacteria which could be found on Bryozoa (Porter *et al.* 2001). The problem of contamination is further exacerbated by the use of so-called "universal" primers for the polymerase chain reaction (PCR). Examples of such studies are, for instance, Halanych (1995) which used universal primers derived from those published by Hillis and Dixon (1991) (information from personal communication with A. Waeschenbach), and Hao *et al.* (2005) which used universal 16S rRNA primers for their Bryozoa study. This issue was first raised in regards to the mtDNA 16 rRNA bryozoan sequences (Dick *et al.* 2000) by Porter *et al.* (2001). In their work an attempt to use Dick's (2000) universal 16S primers for other Bryozoa failed for *Alcyonidium diaphanum* and caused multibanded PCR products. As a result an ingenious method was proposed by Porter *et al.* (2001) to use DNA in oocytes of Bryozoa; thus, instead of using DNA-rich somatic tissue, DNA is extracted from oocytes or—as it happened many times in this work—from complete bryozoan larvae. In this work the combination of the above method with the

protocol proposed by Sutherland *et al.* (1998) for the lysis of the *Mytilus edulis* larvae was used.

In this work once the colony with larvae or embryos was identified (see Figure 11 for an example of colony with embryos) they were extracted.

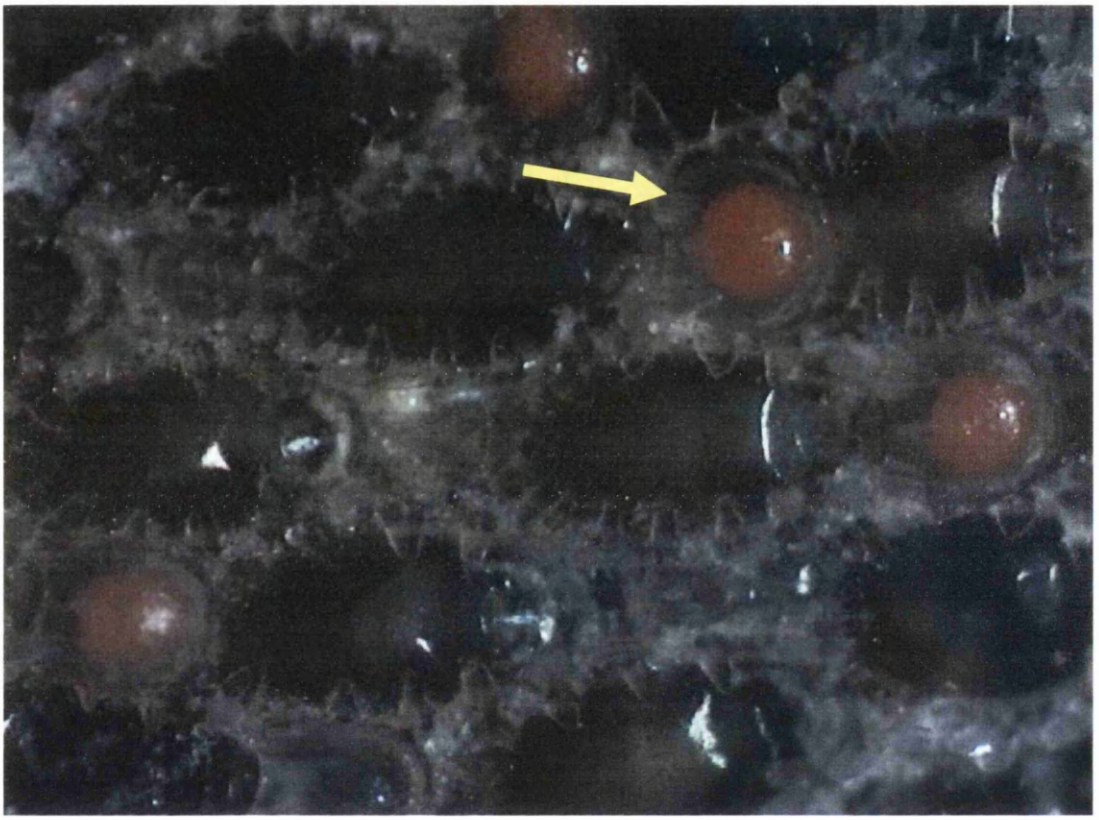


Figure 11 Example of a colony of *Callopora rylandi* with ovicells (bright pink ova can be clearly seen through the ovicell walls). Magnification *ca* x35, light microscope with digital camera attachment. Yellow arrow indicates an ovicell.

The ovicells were dissected and the larvae extracted from each individual ovicell under a stereo microscope (Olympus SZ60, magnification *ca* x60) with a pair of titanium forceps with extra fine tips (Dumont™). Breaking the sides of the ovicell released free-swimming larvae, which were picked up from the colony surface or from the water medium with a micropipette (1-10 µl) and placed into a staining watch glass.

The number of larvae extracted from each colony depended on the number of ovicells with embryos present; however, if many embryos were present then around 30 were extracted at each time in order to be used later for the PCR primer optimisation (see below). In some cases a very small number of embryos was present and although extraction was done there was not enough material for the completion or optimisation of PCR. Consequently, several species, *Alcyonidium*

diaphanum, *Bugula neritina*, *Celleporella hyalina*, *Chorizopora brongniartii* and *Omalosecosa ramulosa,* had insufficient DNA material to be included in the later analysis.

Larvae, once placed in the watch glass, were washed with filtered seawater to minimise possible contamination of the DNA sample with foreign DNA. The seawater was sterilised by filtering through a sterile 0.2 μm microporous filter (Minisart). Each larva was washed three times by transferring it through a series of watch glasses with sterile seawater.

Once the larvae had been extracted and washed a lysis reaction was performed to break up the cells and release DNA. The lysis protocol was adopted from a protocol for *Mytilus* larvae (Sutherland *et al.* 1998) for the reasons described above. Individual larvae were isolated using a 1-10 μm pipette and each larva was transferred to a microtube (50 μl tube, the same as used for PCR) containing 15 μl lysis solution and incubated at 37°C for 1 hour. The samples were then heated to 99° C for 10 minutes in a thermocycler to inactivate proteinase K. The lysis solution was prepared from:

- 7.5 mM Tris-HCl at pH 8.3;
- 3.75 mM NH_4Cl;
- 3.75 mM KCl;
- 1.5 mM $MgCl_2$ and
- 2 μg proteinase K per 15 μl lysis solution.

Transfer of larvae into the lysis solution was performed on ice, to minimise the activity of proteinase K prior to incubation. Although the initial protocol suggested that it was possible to use larvae up to 3 days old, it was found that larvae once extracted, if kept for more than 12 hours, would die and disintegrate. Consequently, lysis was performed within 2-3 hours of extraction. Once the process was completed the tubes were marked according to the date of extraction and the species from which they were extracted and placed into separate plastic containers according to the species. The lysed larvae were stored in microtubes at -20°C in a laboratory freezer. This way they could be stored for several months for further analysis.

3.2 Polymerase chain reaction (PCR)

This section describes the conditions of the PCR, and the modifications that were made to the protocol, which were necessary in order to optimise the reaction. During the optimisation process, several components were varied in concentration, using the original oligonucleotide primers set. However, once optimised, the only component which was varied depending on the DNA source and primers, was in fact the concentration of primers. This is shown in the main protocol. The details of the primer optimisation is shown in a separate section below dedicated to the primers used in this work.

3.2.1 General PCR conditions

The typical PCR protocol, which was adopted in this work, was based on the 15 µl reaction that was performed in the 50 µl mini PCR tubes.

- 5 µl of dNTP 0.1 mM (Promega)
- 1.5 µl of Buffer II (ABgene) x10 concentrated (100 mM Tris-HCl at pH 8.3, 500 mM KCl) or
- 1.5 µl of Buffer I (ABgene) x10 concentrated (100mM Tris-HCl at pH 8.3, 500 mM KCl, 15 mM $MgCl_2$)
- 0.5-1.5 µl of 100 µM primers (supplied by Sygma-Genosys) - see below for discussion
- 1 unit of *Taq* DNA Polymerase (ABgene)
- 0.5 µl of $MgCl_2$ 25 mM as supplied with Buffer II (ABgene)
- 2 µl of DNA
- H_2O (Milli-Q® purified) added to complete reaction to 15 µl when required

During PCR different buffers were used: initially Buffer I, which contained $MgCl_2$ and later Buffer II, which did not contain $MgCl_2$.

With the above concentrations a thermal cycler (PTC-225 MJ Research Peltier Thermal Cycler) was set for the initial cycle of denaturation at 95°C for 60 seconds, followed by a cycle of annealing at temperatures specific to the primers used, ranging from 40° to 70°C (see below for the temperature optimisation details) for 60 seconds for a total 33 cycles. Each annealing cycle was followed by an extension cycle at 72°C for 90 seconds, and after the last cycle the PCR reaction was terminated with a final extension step of 72°C for 10 minutes. The thermal

cycler was set to keep the samples at the end of PCR indefinitely at 4°C holding temperature. This was done to preserve PCR products after the reaction, if these were to be left unattended for some period.

3.2.2 PCR optimisation

PCR optimisation was performed at the beginning of the project for several components of the reaction, and later on a regularly basis if necessary as new primers became available.

One of the first component that required optimisation was $MgCl_2$. The concentration of Mg^{2+} ions is critical for the *Taq* DNA polymerase (ABgene) to work correctly as it influences the activity of *Taq* DNA polymerase through dNTP—Mg^{2+} complexes which interact with nucleic acids (McPherson 2000; Hillis 1996). The concentration is usually checked by performing a gradient series with the most common concentration of 0.5-3.0 mM. Therefore, the original concentration, which was fixed in Buffer I (1 mM), was adjusted using a series of PCR reactions with a gradient concentration of $MgCl_2$. As a result a switch was made from Buffer I to Buffer II as the latter was supplied without added $MgCl_2$, and thus allowed a more precise adjustment of the $MgCl_2$ concentration (Figure 12).

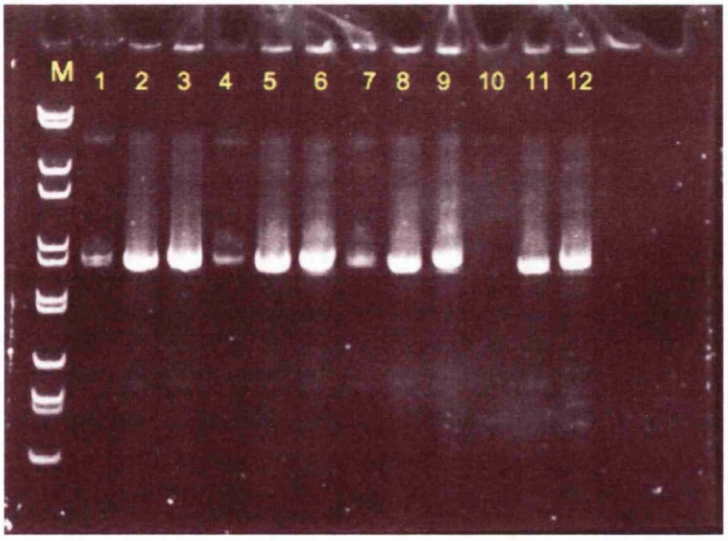


Figure 12 PAGE gel showing an example of $MgCl_2$ optimisation gradient during PCR. 4 different $MgCl_2$ concentrations were tested for three independent samples: 1,2,3 - 0.8 mM of $MgCl_2$; 4,5,6 - 1.6 mM $MgCl_2$; 7,8,9 - 2.0 mM of $MgCl_2$; 10,11,12 - 2.5 mM of $MgCl_2$. Samples 7 to 12 show secondary bands below the expected weight of 986bp and therefore this concentration of $MgCl_2$ is not used. (M) DNA Marker VI: 2176, 1766, 1230, 1033, 653, 517, 453, 394, 298, 234, 154 bp from top of the gel.

Consequently, the concentration that was favourable here was 0.8 mM. Interestingly, once $MgCl_2$ concentration was optimised there appeared to be no further need to adjust it even when new primers were introduced[12].

Temperature optimisation for annealing was performed on a primer to primer basis—an example of the temperature gradient is shown in Figure 13.

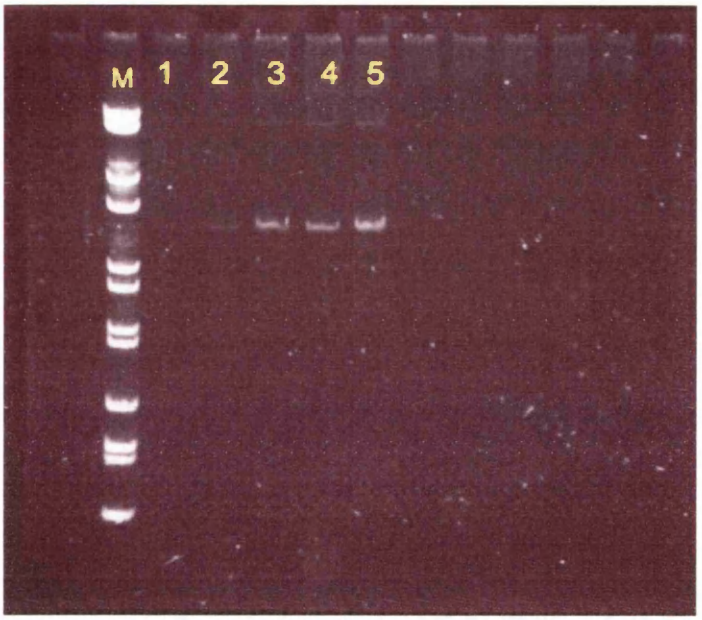


Figure 13 PAGE gel showing an example of temperature optimisation during PCR using temperature gradient: 1 - 45°C; 2 - 50°C; 3 - 55°C; 4 – 60°C; 5 – 65°C. (M) DNA Marker VI: 2176, 1766, 1230, 1033, 653, 517, 453, 394, 298, 234, 154 bp from top of the gel.

However, it was generally noted during the optimisation of multiple primers that the best temperature was approximately 3°-5°C below the empirical melting point of the primers use. The selection of as high an annealing temperature as possible also increases the chances of the primer to anneal only to its specific template, thus increasing the likelihood of amplifying only the target sequence (McPherson 2000). These temperatures were stated with the primers as received from the supplier (Sigma-Genosys). As many primers were developed and tested (see detailed discussion of this below) inevitably some pairs of primers differed considerably in the melting temperature and thus sometimes required that the optimum annealing temperature for the PCR was set just 2-3°C below the lowest melting temperature of the primer pair.

[12] With the exception of *Alcyonidium* species PCR reaction – see below in this chapter.

3.3 Gel electrophoresis and staining

3.3.1 Agarose gel

Agarose 1% gels were prepared by mixing:

- 0.4 g of Agarose
- 40 ml of TBE x1 buffer (0.13 M Tris Base, 0.075 M boric acid, 0.25 mM EDTA, pH 8.3)

Once prepared and melted in the microwave oven the gel solution was poured into the standard gel trays (Pharmacia Gel Electrophoresis Apparatus GNA-100), and once cooled used at the voltage of 60V for approximately 110 minutes or longer as required (BIO-RAD Power Pac 300 power source was used).

The DNA ladder marker used depended on the product examined, but most commonly Marker VI (Roche) (250 μg/ml ready to use solution in TE buffer) with the size of ladder 154 – 2176 bp or Marker XIV (Roche) with the size ladder of 100 – 2642 pb were used. Also in cases when cloning products were examined λ-phage DNA ladder was used (prepared by restricting λ-phage DNA by *Pst*-I enzyme in the laboratory). Agarose gels were stained in ethidium bromide (added 35 μl in concentration of 10 mg/ml) directly to the agarose preparation. The visualisation of the DNA bands on the stained gel was done using an ultraviolet transilluminator (Ultra Violet Products, TFM-20), and afterwards digitally photographed using a BIO-RAD Molecular Imager Gel Doc XR System, which uses proprietary software. With some exceptions, the images were stored digitally for later reference in TIFF format, and also printed (Mitsubishi P-91 digital b/w thermal printer) for immediate examination. Agarose gels were used for the examination of cloning products as well as during the purification of PCR products for further analysis.

3.3.2 Polyacrylamide gel electrophoresis (PAGE)

Polyacrylamide gels were prepared in the glass beakers using the following ingredients (shown for 2 gels or approximately 23 ml):

- 2.5 ml of TBE 10x buffer (1.3 M Tris Base, 0.75 M boric acid, 25 mM EDTA, pH 8.3)
- 4.0 ml of Acrylamide (37:1 30%)
- 16 ml distilled H_2O

- 220 µl APS (10% distilled water solution of ammonium persulphate $(NH_4)_2S_2O_8$)
- 22.5 µl Temed (ultra pure, Invitrogen)

The above solution was thoroughly mixed in a beaker and poured into the glass cassette casts attached to the casting stand, supplied with the PAGE tank system (BIO-RAD Mini-PROTEAN® 3 Cell). These would normally (at room temperature) set within 30 minutes and be ready to be used. If not used immediately, the cast gels were stored in the laboratory refrigerator (+4°C). This system allowed a relatively quick gel run (*ca* 60-100 min) depending on the size of the DNA product, using the power apparatus (BIO-RAD Power Pac 300) at 60V. The gels were stained using SYBR® Gold nucleic acid gel stain (Molecular Probes) at the supplied concentration by mixing 10 µl of stain in a tray containing 100 ml of TBE x1. The gels were stained in a purposely built acrylic glass tray protected from direct light by a foil–laid lid. The use of SYBR® Gold stain reduced the speed of staining (freshly prepared solution would sufficiently stain one gel within 3-5 minutes) compared to silver staining method (not described here); and increased sensitivity of DNA staining compared to ethidium bromide. The gels were visualised using a UV transilluminator and imaged in the same manner as agarose gels (see above).

3.4 DNA purification

In most cases, direct sequencing was used once a successful PCR was achieved. In some cases though, for *Alcyonidium* species, when a specific band with the target DNA weight had to be extracted from a gel, cloning was used prior to the sequencing reaction. This was mainly due to the insufficient DNA yield produced by extracting a single band from the gel. In all cases, PCR products were purified to clean out all remaining PCR reagents such as amplification primers, nucleotides, buffer components as well as co-products such as primer-dimers and non-target amplification products, which may inhibit sequencing and further work. The most commonly mentioned disadvantage of DNA purification methods is the product loss; however, here a Promega kit (Wizard® PCR Preps DNA Purification System) was used which uses minicolumns with silica membranes. These methods claim

recovery of 97%[13] of the DNA fragments with size around 1000bp (in this work the average product size was 800bp). This kit was used as per supplied protocol without the vacuum manifold. In cases when the purification of a particular band from the gel was required the PCR products were run on the agarose gel and then the required band was extracted using a scalpel and purified using GFX™ PCR DNA and Gel Purification Kit (Amersham Biosciences). This uses a similar method to the previous kit—silica membrane and the microspin columns. This method was particularly suited when bands in addition to the expected yield band were displayed on the gel and had to be extracted, for instance in cases when suitable primers were not found or not yet fully optimised.

3.5 Cloning

In several cases, in particular for optimising some primers for *Alcyonidium* species, cloning was used instead of direct sequencing. The yield of the PCR was normally sufficient for the direct sequencing of the PCR product (the average yield of a PCR reaction was around 40 ng/μl). Cloning was used when the PCR products were not specific and contained more than one band and the desired PCR product had to be extracted from the agarose gel and then had too low yield for direct sequencing. As the PCR products were amplified using *Taq* DNA polymerase the product had terminal adenine overhangs, thus making it suitable for a ligation in many commercially available vectors with 3′ terminal thymine, so-called TA cloning (McPherson 2000).

In this work an already well tested (in the same laboratory at Swansea University) and successfully working pGEM®-T Easy Vector System (Promega Corporation) was used (Figure 14).

[13] As per Promega protocol manual supplied with the kit.

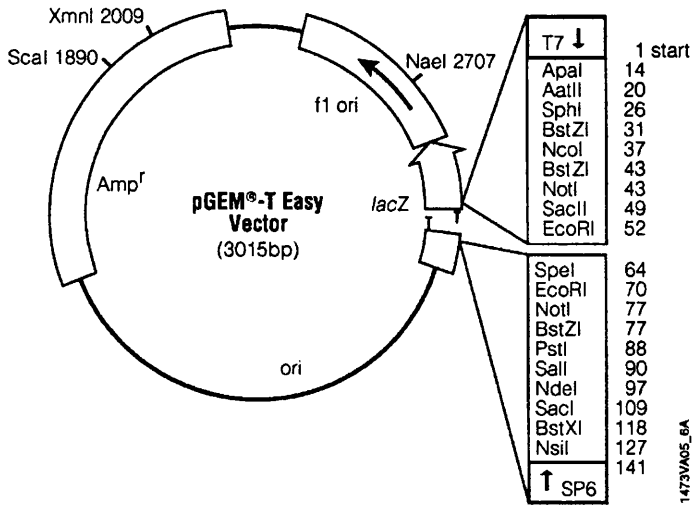


Figure 14 pGEM®-T Easy Vector circle map, with sequence reference points shown to the right of the map. Note the recognition sites for the restriction enzyme *Eco*R I, which were used for the single-enzyme restriction digestion for testing for the presence of correct inserts. Image taken from the Promega manual (TM042) accompanying the vector system.

When applicable, the protocol supplied by Promega was used. The following stages were employed during the cloning process. Prior to the first step of ligation, the PCR products were cleaned as described above using Promega PCR kit.

3.5.1 Ligation reaction

Ligation was performed using the procedure from the Promega protocol using the following ingredients (all supplied as part of the kit):

- 5 µl 2X rapid Ligation Buffer, T4 DNA Ligase
- 50 ng pGEM®-T Easy Vector
- *X* µl of PCR product (*X* - see below for details)
- 1 µl of T4 DNA Ligase (3 Weiss units/ µl)
- Deionised Milli-Q® water to a final volume of 10 µl

All the ingredients were mixed by pipetting and then the mix was incubated overnight at 4°C. The amount of PCR product (*X*) was determined based on the molar ratios (3:1) of the insert to the vector, specified in the Promega protocol and the concentration and length of the DNA product insert. The length of the insert in

this case was 1.2 kb and the concentration as determined by the spectrophotometer[14] (NanoDrop® ND-1000 UV-Vis). The calculation of the amount of DNA insert required is done using the formula:

- $\frac{50ngV \times 1.2kpI}{3.0kbpV} \times \frac{3}{1} = ngI$, where V is Vector, and I is Insert

Once incubated overnight the ligated chimeric plasmids were transformed into the host bacteria.

3.5.2 Preparation of electrocompetent cells

Electro-competent cells *Escherichia coli* JM109 were used and the transformation was done by the process of electroporation. The following protocol, which was routinely used in the laboratory was used for the preparation of the electrocompetent cells. This protocol was adopted from Sambrook *et al.* (2000).

- A 1/100 dilution of fresh overnight *E.coli* JM109 culture was incubated in 500 ml LB broth. On reaching an optical density of liquid medium (OD_{600}) equal to 0.5-0.7 the cells were chilled on ice for 20 minutes then harvested by centrifugation at 400 rpm for 15 min at 4°C.
- The supernatant was decanted and the pelleted cells carefully resuspended in 500 ml ice cold 10% glycerol, the cells were then pelleted and resuspended in 250 ml ice cold 10% glycerol.
- Finally, the cells were pelleted, resuspended in 20 ml ice cold 10% glycerol, pelleted again and resuspended into 2 ml ice cold 10% glycerol.
- The 2 ml suspension of electrocompetent JM109 cells was aliquoted out and stored at -70°C until needed.

3.5.3 Transformation by electroporation

1 mm electroporation cuvettes (HiMax EP-101 CellProjects) were used and MicroPulser™ electroporation Apparatus (BIO-RAD). The following protocol adopted from the Sambrook (2001) and BIO-RAD operating guide manual and routinely used in the laboratory for the electroporation was used:

[14] This spectrophotometer was used for all applications when DNA concentration analysis was required.

- 10 µl of ligated solution (containing plasmid vectors with ligated PCR product) was diluted to 100 µl (by adding 90 µl of PCR-sterile water).
- Electro competent cells, which are stored at -80°C were thawed on ice for 3-4 minutes prior to the transformation without removing them from tubes in which they are stored.
- 3 µl of the ligated solution dilute was transferred into a tube (1.5 ml) with electro-competent cells suspension and mixed very gently by pipetting.
- 50 µl of the above mixture containing plasmid vectors and the electrocompetent cells was added to the groove of the electroporation cuvette (which was pre-chilled first in freezer and then kept on ice).
- The cuvette was placed into the chamber slide of the pulser apparatus and pulsed once.
- Once pulsed, if the transformation was successful the apparatus displayed PLS on the screen, if it showed ARC then the electroporation did not work and the process of electroporation had to be repeated with the new cells.
- Immediately after pulsing (within 60 seconds) 500 µl sterile SOC medium was added to the electroporation cuvette and mixed gently by pipetting, then the cell suspension was transferred into a sterile Bijoux tube and incubated at 37°C for 1 hour in the culture shaker (250 rpm).

Once the above procedure was done the culture was plated in two volumes (100 µl and 400 µl) on LB agar and ampicillin plates. Different quantities were used in order to achieve a potential different density of the colonies (Sambrook 2001). The plates were prepared using a commercially available ready-mix S-Gal™ / LB Agar Blend (Sigma® S-Gal™ /LB Agar Blend C4478-6X500ML). The use of S-Gal™ (3,4-cyclohexenoesculetin-b-D-galactopyranoside) gives a much higher output of colony growth than a traditionally used X-Gal and also makes colony selection easier due to their darker colour (Heuermann and Cosgrove 2001). The mix powder was prepared as per the supplied manual protocol and then ampicillin (50µg/ml concentration) was added to the solution prior to pouring it on the plates.

Plates were then placed into an incubator (37°C) overnight for approximately 16 hours and colony growth was verified the following day. The

successful cloning was verified using "blue-white"[15] colour screening method. For an example of the plate with black and white colonies (the latter containing a successful PCR cloning), see Figure 15.

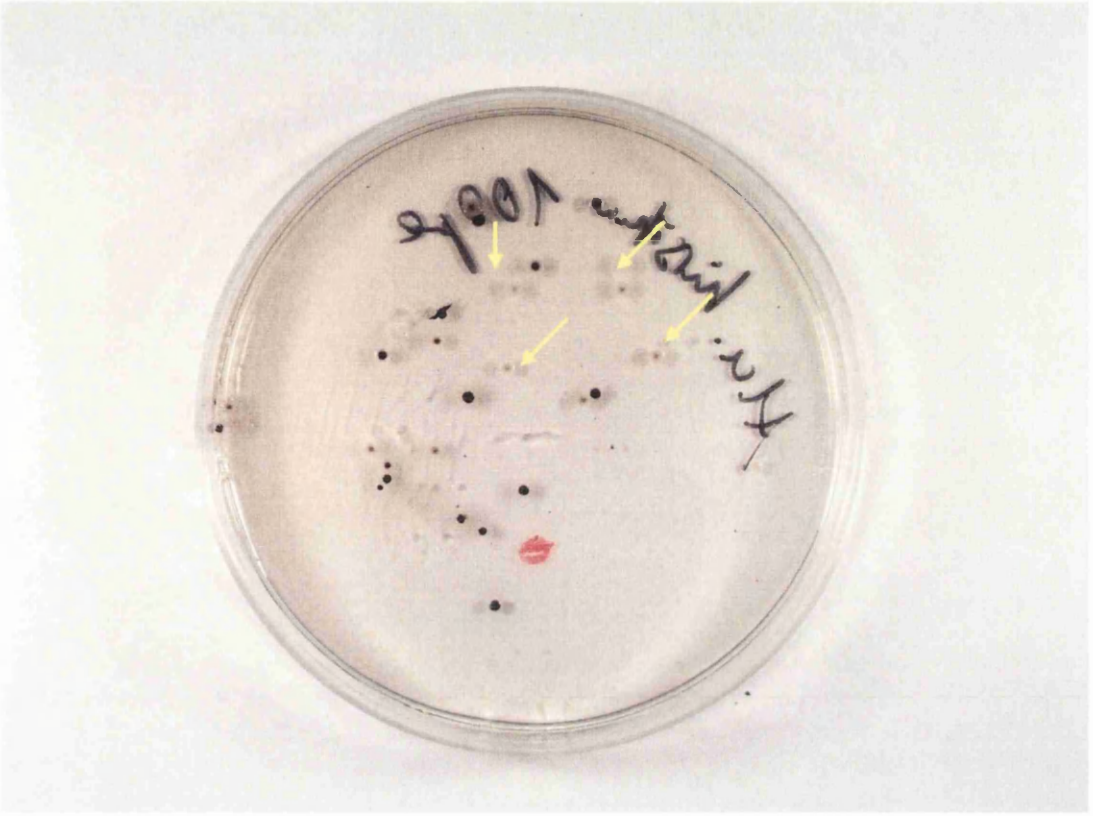


Figure 15 Cloning plate showing black and white colonies grown on S-Gal™ LB Agar medium. Black colonies can be clearly seen; white colonies can be distinguished by shadow-like spot around the colony growths (see yellow arrows).

The white colonies were picked with a P10 Gilson® micropipette from the plates and placed into universal growth tubes containing LB medium and ampicillin at 50 μg/ml concentration, and placed overnight (approximately 16 hours) in the incubator (37°C) shaking at 225 rpm. Both LB and SOC media were prepared as per standard recipe from Sambrook (2001) molecular cloning manual.

3.5.4 DNA extraction from plasmids

Plasmid extraction from the cloned cells was done using a Promega DNA purification system (Wizard® Plus SV Minipreps DNA Purification System) following the protocol provided by the manufacturer. Plasmids were checked for correct inserts prior to sequencing using a restriction enzyme single-digest by the

[15] Due to the specificity of the S-Gal™ the actual colour was black-white.

*Eco*RI enzyme (Invitrogen). *Eco*RI restriction sites are positioned conveniently on both sides of the inserts (Figure 16 also Figure 14).

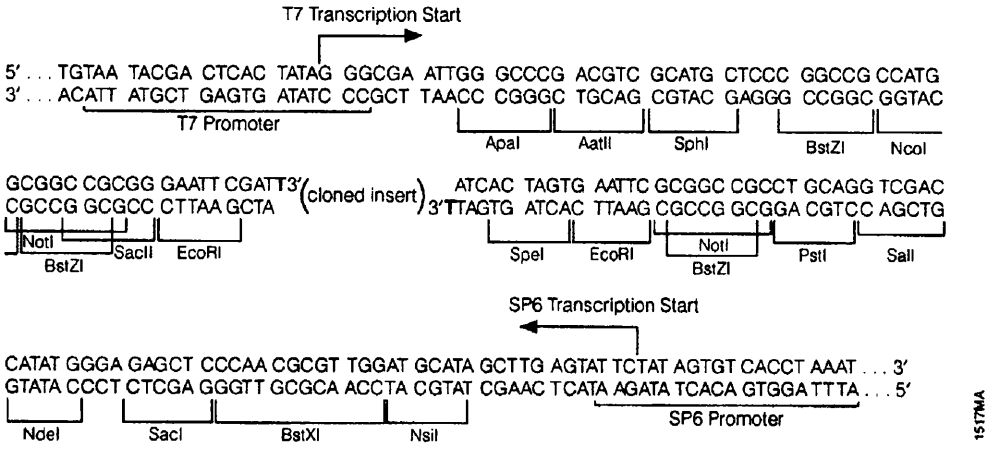


Figure 16 Multiple cloning sequence and promoter sites of the pGEM®-T Easy Vector. Primers annealing to the T7 and SP6 promoter sites were used for sequencing reaction. Image taken from the Promega manual (TM042) accompanying the vector system.

The digestion was performed as per the following protocol:

- 5 µl of DNA template
- 2 µl of 10x REact® 3 Buffer (50 mM Tris-HCl, pH 8.0; 10 mM $MgCl_2$; 100 mM NaCl)
- 1 µl of *Eco*RI Enzyme
- 12 µl of H_2O (to make up to 20 µl reaction)

The digests were incubated in a water bath at 37°C for 2-3 hours and visualised on an 0.7% agarose gel, stained with ethidium bromide (Figure 17). Expected insert was *ca* 1.2 kb, restriction results were previewed *in silico* using EnzymeX software (Griekspoor and Groothuis, mekentosj.com).

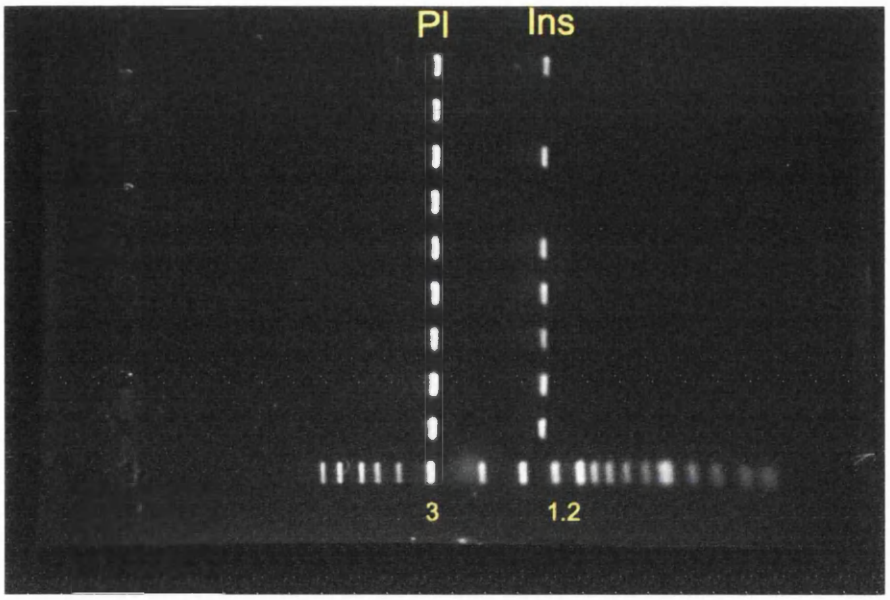


Figure 17 Restriction digest by *Eco*RI of pGEM®-T Easy Vector plasmid containing a PCR insert - marked on the gel as (PI) and (Ins) respectively. Here a 2-Log DNA ladder was used (NEB®), the length in kb is 10.0, 8.0, 6.0, 5.0, 4.0, 3.0, 2.0, 1.5, 1.2, 1.0 *etc.* The 3.0 kb and 1.2 kb are marked on the gel with numbers 3 and 1.3 respectively.

Once visualised on the agarose gel those samples which had an insert indicating successful cloning were selected and stored for the following stage – sequencing.

3.6 Sequencing

Both direct sequencing and sequencing of the cloned material was done following the same protocols. The sequencing was done on an Applied Biosystems 3730 DNA Analyzer automated DNA sequencer as per manufacture instructions. The samples containing PCR product to be sequenced were placed into 0.2 ml sterile tubes by diluting DNA product with PCR-sterile H_2O in order to achieve the concentrations required by the sequencing service[16]: for the PCR product a 500-1000 bp template was adjusted to be 5-20 ng per reaction, and for the plasmid sequencing the concentration of the template was 200-300 ng per reaction. In both instances the sequencing service required a minimum of 15 µl of template per reaction, thus in the case of direct sequencing this required diluting the samples with H_2O. The average quantity of recovery of DNA after PCR was around 40 ng/µl. The recovery of DNA after cloning was around 150 ng/µl.

[16] Sequencing was outsourced to the University of Dundee Sequencing service www.dnaseq.co.uk

In addition to the DNA template the sequencing service required primers to be supplied alongside the template. Primer quantity was standard for all reactions at 3.2 pmol per reaction. Primers were thus separately diluted to the required concentration and aliquoted to separate 0.2 ml tubes. For the direct sequencing, the same primers as those for the corresponding PCR were used (see section below). For the sequencing of the plasmids, standard library primers were used (supplied by the sequencing service). Standard primers T7 and SP6, complementary to the regions of vector flanking the PCR insert, were used (see Figure 16 page 44).

Primer	**Primer Sequence (5'->3')**
T7	TAATACGACTCACTATAGGG
SP6	AGCTATTTAGGTGACACTATAG

Both DNA template and primers were sent to the sequencing service by overnight mail, alongside electronic submission of material (references and identification for the tubes) which was done over the internet on the sequencing service website. The sequencing results were usually available within 24 hours from the internet site server run by the sequencing service. The results were supplied in the following format for each sequencing reaction separately: as individual chromatogram file (*.ab1) and ASCII text file containing the actual DNA sequence. Also an Excel spreadsheet was supplied which had all information about sample identification—this was very useful for further analysis.

Prior to the use of the Dundee University sequencing service some sequencing was performed at Swansea University using the Beckman Coulter CEQ™ 2000XL DNA sequencing system. For this, DNA samples were prepared according to the Beckman Coulter CEQ™ 2000 DNA sequencing protocol. The sequencing reaction (similar to the PCR reaction) was necessary for this process and was performed using identical primers to those which were used for the corresponding PCR, in addition to the DNA template as well as Dye Terminator Cycle Sequencing kit mix (supplied by Beckman Coulter). The DNA sequencing reaction was prepared as per the Beckman Coulter protocol (total volume 20.0 µl) and run on the thermal cycler using the following program: 96°C for 20 sec (denaturing); 50°C for 20 sec (annealing); 60°C for 4 min (extension), repeated for 30 cycles followed by 4°C holding. Following that the PCR products were

submitted to sequencing service at the Swansea University. The results were obtained in the same form as above – a chromatogram file and a sequence file for each reaction. Although use of different sequencing services was coincidental, the quality of sequences received from Applied Biosystems 3730 DNA Analyser was much higher, and thus this service was preferred.

3.7 *In silico sequence preparation*

Once received the sequences had to be checked against each other (*i.e.* forward vs. reverse) and assembled to make sure that any possible errors in the files could be corrected. This was done in several stages (Figure 18) using chromatogram files and the aligned sequence files.

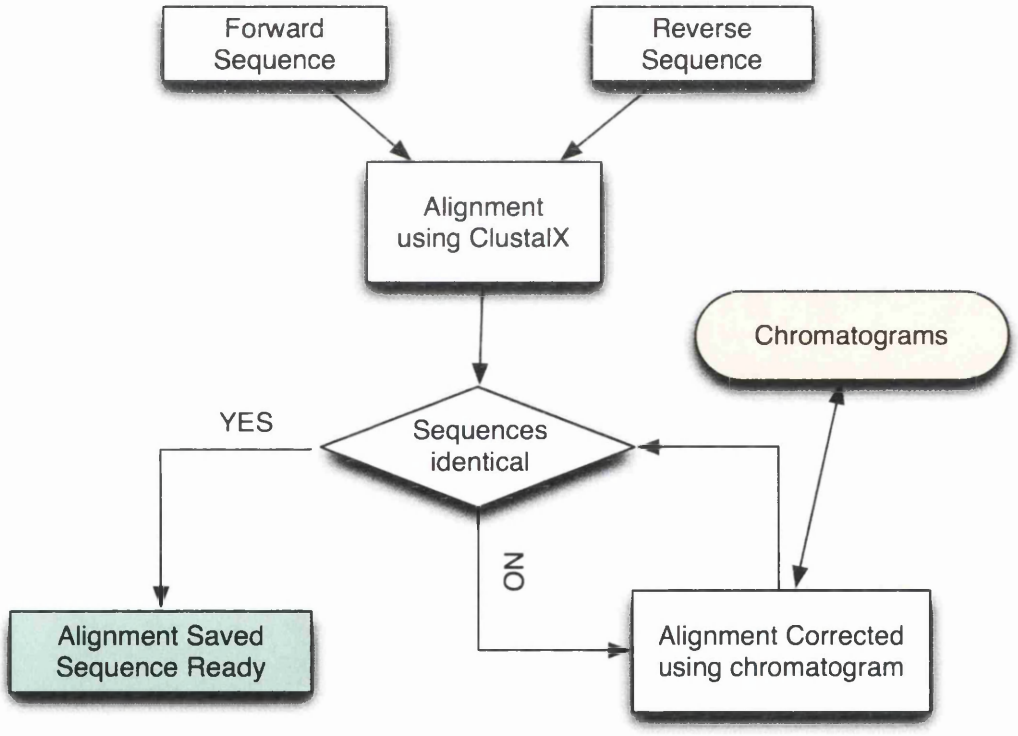


Figure 18 Diagram showing the workflow of the sequence alignment once received from the sequencing service.

When received from the sequencing service each sequence has two files – one corresponding to the forward primer and one to the reverse primer sequencing reactions as well as corresponding chromatograms files. It is very important to use both files supplied as forward and reverse sequences usually do not match fully along the entire sequence. Also the chromatograms usually have much stronger peaks at the beginning of the sequence and become weaker by the end of the

sequence, which in its turn causes possible errors in reading or ambiguities between the two files. Then the forward primer generated sequence is aligned to the reverse sequence, which has to be reverse-complement transformed. After that, once the transformation is done, the two sequences were aligned and any discrepancies investigated. In most cases when a mismatch was found between the forward and the reverse sequences, a simply reference to the chromatogram could resolve the issue – one of the sequences would usually have a very weak peak or two peaks on top of each other, which would "confuse" the reading software and a wrong nucleotide would be read off. In these cases that nucleotide corresponding to the sequence which had a strong clean peak on the chromatogram was accepted, if this was not possible then N[17] was instead placed in the sequence. By using this simple technique, many discrepancies in the sequences could easily be resolved.

Once the sequence was cleared from inconsistencies as described above it was checked using BLASTn search on NCBI. In most cases, the results of the BLAST search of the new sequence would result in a list of sequences with significant alignment similarities to those of other Bryozoa (Figure 19), usually in the first several lines, indicating high score. For "good" results the score would approach the double value of the length of the original sequence which was submitted for the BLAST query. The presence of many sequences at the top of the list (sorted by the E-value[18]) which did not belong to Bryozoa (such as "uncultured metazoan") indicated potential problems with the sequences and consequently required further attention. These problems were encountered with *Alcyonidium* sequences produced in this work – see Chapter 4 section for a detailed discussion.

[17] To stand for unknown base.

[18] E-value stands for Expect value, which describes the random occurring hits that can be observed by chance when BLAST database search is performed. This value decreases exponentially as the number of score hits increases. E-value ideally should approach zero and a lower E-value indicates a more significant sequence match (Karlin and Altschul 1990).

Sequences producing significant alignments:
(Click headers to sort columns)

Accession	Description	Max score	Total score	Query coverage	E value
AY210444.1	Crisia sp. YJP-2003 18S ribosomal RNA gene, partial sequenc	2678	2678	95%	0.0
AF119080.1	Lichenopora sp. AMNH1 18S ribosomal RNA gene, complete s	2438	2438	93%	0.0
DQ459959.1	Pristina jenkinae 18S ribosomal RNA gene, partial sequence	1951	1951	93%	0.0
AF209452.1	Bathydrilus litoreus 18S ribosomal RNA gene, partial sequenc	1949	1949	93%	0.0
DQ459960.1	Pristina proboscidea 18S ribosomal RNA gene, partial sequenc	1945	1945	93%	0.0
AF411889.1	Bathydrilus formosus 18S ribosomal RNA gene, partial sequer	1945	1945	93%	0.0
AF411882.1	Bathydrilus rohdei 18S ribosomal RNA gene, partial sequence	1945	1945	93%	0.0
AF411875.1	Pristina longiseta 18S ribosomal RNA gene, partial sequence	1945	1945	93%	0.0
U08331.1	Lingula anatina 18S ribosomal RNA gene, partial sequence	1941	1941	93%	0.0
DQ209217.1	Mesochaetopterus taylori 18S ribosomal RNA gene, partial se	1940	1940	93%	0.0
AY040699.1	Spirosperma ferox 18S ribosomal RNA gene, partial sequence	1940	1940	93%	0.0
AF411887.1	Heronidrilus gravidus 18S ribosomal RNA gene, partial sequer	1940	1940	93%	0.0
X81631.1	Lingula anatina 18S rRNA gene	1940	1940	93%	0.0
AY340433.1	Heronidrilus gravidus 18S ribosomal RNA gene, partial sequer	1936	1936	93%	0.0
DQ280316.1	Tubifex ignotus clone A 18S ribosomal RNA gene, partial sequ	1934	1934	93%	0.0
DQ459969.1	Rhyacodrilus coccineus 18S ribosomal RNA gene, partial sequ	1934	1934	93%	0.0
DQ459961.1	Pristina aequiseta 18S ribosomal RNA gene, partial sequence	1934	1934	93%	0.0
AF411879.1	Tubifex ignotus 18S ribosomal RNA gene, partial sequence	1934	1934	93%	0.0
AF360992.1	Limnodrilus hoffmeisteri 18S ribosomal RNA gene, partial seq	1934	1934	93%	0.0
DQ209221.1	Chaetopterus sarsi 18S ribosomal RNA gene, partial sequence	1932	1932	93%	0.0
U08329.1	Lingula adamsi 18S ribosomal RNA gene, partial sequence	1930	1930	93%	0.0
DQ459984.1	Dero digitata 18S ribosomal RNA gene, partial sequence	1929	1929	93%	0.0
DQ279934.1	Neocrania anomala 18S ribosomal RNA gene, partial sequenc	1929	1929	93%	0.0
AF411908.1	Bothrioneurum vejdovskyanum 18S ribosomal RNA gene, pai	1929	1929	92%	0.0
AF411880.1	Tubifex smirnowi 18S ribosomal RNA gene, partial sequence	1929	1929	93%	0.0
AY885576.1	Heterodrilus ersei 18S ribosomal RNA gene, partial sequence	1929	1929	93%	0.0
U08334.1	Neocrania huttoni 18S ribosomal RNA gene, partial sequence	1929	1929	93%	0.0
AY842018.1	Discinisca cf. tenuis BLC-2005 isolate D1504 18S small subu	1929	1929	93%	0.0

Figure 19 Results of the BLASTn search query submitted with the *Crisia denticulata* sequence produced in this work. Results sorted by the E-value (default settings).

Once a sequence was accepted it was saved in a FASTA format and "passed on" to the following stage at which the assembling of the sequences was done. As the 18S rRNA gene is relatively large (*ca.* 1.8 kb) it was usually not possible to sequence the entire gene at once, thus three overlapping parts of the gene were sequenced and then assembled, using exactly the same technique as above. This method was relatively fast and did not require[19] expensive commercial software such as CodonCode Aligner or Sequencher™.

3.8 Primer design and optimisation for overlapping segments

The oligonucleotide primers in this work were designed for the 18S nuclear gene and their design was based on some common assumptions recommended in the literature. Because of the problems with the so-called "universal" primers, due to their non-specificity, the primers were designed in such a way that they would match bryozoan species as closely as possible to avoid any possible contamination, *i.e.* picking up DNA traces from any other organisms. Some regions of the 18S gene, corresponding to the stems of the rRNA, are very conserved and are the same for such remotely related organisms as yeast and mammals. Thus, although at first

[19] The software package such as Sequencher™ would be preferable if multiple nested primers were used.

the presence of such sites among the gene sequence can be seen as an advantage for a larger study, it has a direct disadvantage that it permits amplification of non-target organisms and hence introduces contamination.

In general, the primers designed are recommended to be between 18 to 24 base pairs (Hillis 1996) or up to 30 base pairs (McPherson 2000). The length is directly linked to their melting temperature (T_m) and hence the annealing temperature (T_a) of the PCR. Excessive length not only would increase the T_m of the primer but also introduce higher risk of the primer-dimers (a process of self-priming by two primers due to the internal repeats). Also hairpin structures (secondary structures caused by internal primer self-complementarity) can affect PCR. The actual sequence content of the primer is recommended to contain an approximate equal number of each nucleotide (McPherson 2000), whilst GC-heavy primers would also increase the T_m of the primer. Finally, particular attention is recommended to be paid to the 3'-end of the primer as it is this end which once annealed is extended by the polymerase. Thus, it is recommended (Hillis 1996; McPherson 2000) that this end of the primer matches perfectly the template sequence, whilst the 5'-end of the primer can be less specific. The 3'-end of the primers is also recommended by some authors to have a so-called GC clamp (higher content of GC) (Sheffield *et al.* 1989).

As the primers are designed in pairs it is important that their T_m is roughly equal to avoid great discrepancies between annealing temperature of the PCR primers. In cases when primers T_m differ by greater than 5°C this may lead to non specific priming of the primer with higher T_a and thus produce unexpected PCR results. The melting temperature of the primer can be calculated using a simple approximation formula (McPherson 2000).

- T_m=[(number of G+C)x4°C + (number of A+T)x2°C]

Although the above formula is very useful it was found later that the empirical T_m of the primer (which was evaluated by the company that supplied oligos) was slightly different from the theoretical estimation, and thus the empirical value was used for the PCR.

In this work, the primers were designed by hand using guidelines outlined above and then further checked and assessed using specific software for the primer design. There are currently many software packages and online utilities dedicated to

primer design. Abd-Elsalam (2003) for instance lists over 40 packages, however "Net Primer" (PREMIER Biosoft International) was chosen as it allows rapid evaluation of a pair of primers and testing for the presence of the primer-dimers and hairpins and provides a quick report. Also, at the beginning of the study, for a period of one month an evaluation license was provided by PREMIER Biosoft for their commercial package – "Primer Premier 5", which was used for the design of several primers at the beginning of the study. Once primers were designed and considered to be acceptable they were further evaluated *in silico* against the bryozoan sequences using CINEMA[20] program (Parry-Smith *et al.* 1998). This interactive alignment editor allows priming any oligonucleotide sequence against a given alignment with a certain flexibility, and thus allows to evaluate if a given primer would match certain bryozoan sequences.

Several strategies were tried for primer design, originally the universal pairs of primers 18e and 18L; 18N and 18M; 18M0 and 18P and 18h were tested (see Table 5 for details of these primers). These primers were taken from Halanych *et al.* (1998) and Hillis and Dixon (1991) and are the same primers used by Passamaneck and Halanych (2006) and by Halanych (1995). The results from these primers were quite unsuccessful (Figure 21) and the situation was further aggravated by the fact that these universal primers may amplify non-target DNA and hence obtaining *bona fide* Bryozoa sequences was not guaranteed. Also the second and third pair of these primers (*i.e.* 18N/18M and 18M0/18P) were not overlapping as 18M and 18M0 were priming to the same position on the gene although in a different direction (Figure 20). Consequently, it would be impossible to assemble a complete gene without introducing a gap – in between the second and third pairs of primers.

[20] CINEMA – **C**olour **IN**teractive **E**ditor for **M**ultiple **A**lignments, is now maintained as part of the UTOPIA bioinformatics tool set at the http://utopia.cs.man.ac.uk/.

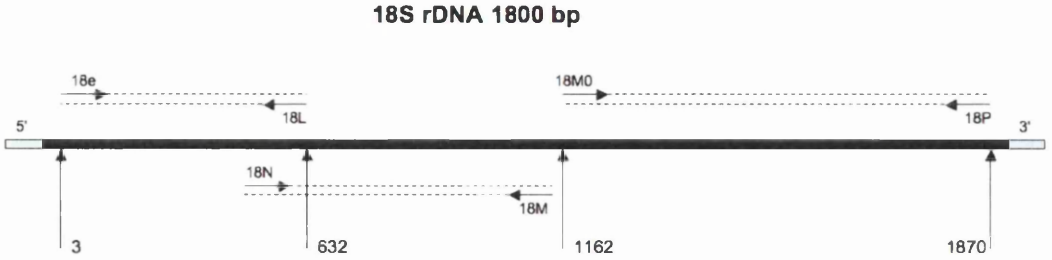


Figure 20 Relative position of the universal 18S primers (Halanych *et al.* 1998, Hillis and Dixon 1991). Note that 18M and 18M0 primers do not allow for the overlap in the sequences corresponding to the second and third segments of the gene. Base pairs are shown with arrows.

Indeed, two bryozoan sequences *Lichenopora* sp. (accession no AF119080) and *Membranipora* sp. (accession no AF119081), which were produced using the universal primers and already deposited to GenBank by Giribet *et al.* (2000) were found to be contaminants (Waeschenbach 2003). It was decided therefore that Bryozoa specific primers were necessary to insure fidelity of the sequences produced.

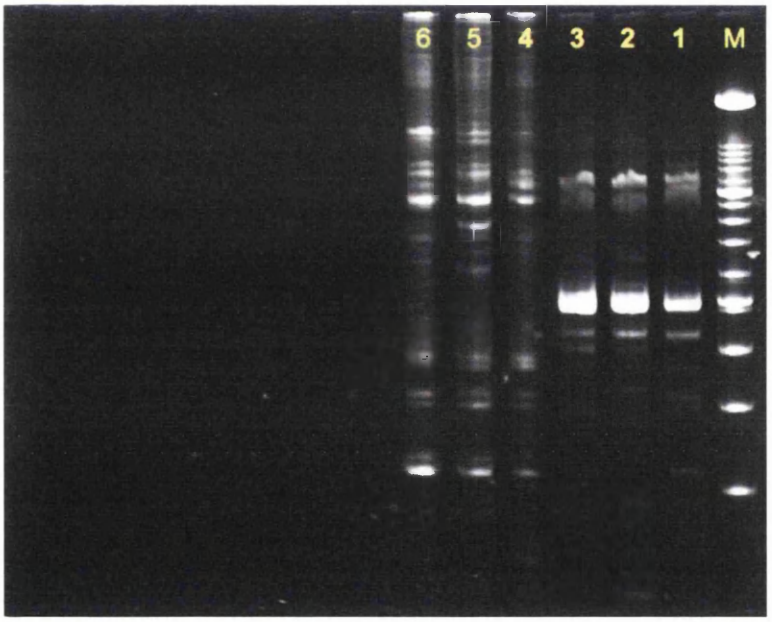


Figure 21 PAGE gel showing unsuccessful amplification of the *Microporella ciliata* (1-3) and *Schizomavella linearis* (4-6) 18S DNA using 18e and 18L primers. (M) - DNA Marker XIV: 2642,1500,1000, 500, 400, 300, 200 bp from top of the gel.

The bryozoan primer development commenced as soon as the first DNA samples were collected in the study – autumn 2003. At the time GenBank contained only 16 of 18S bryozoan sequences (Table 3). These sequences were submitted as

part of the five independent studies. Two sequences, *Lichenopora* sp. (accession no AF119080) and *Membranipora* sp., were excluded as contaminants (see above), further *Alcyonidium gelatinosum* (accession no X91403), now *A. polyoum* (Ryland and Porter 2003) was not used as it was reported to be a possible contaminant (Dr J Porter, personal communication). Out of the remaining 13 sequences, 8 (*Bugula stolonifera*, *Electra bellula*, *Smittoidea spinigera*, *Bugula neritina*, *Membranipora grandicella*, *Schizoporella erratoidea*, *Parasmittina* sp., *Celleporina* sp.) were submitted to the GenBank by Hao *et al.* (2003); however, the validity of these sequences was also questionable. For instance, when all bryozoan 18S sequences from GenBank were analysed using NJ tree to compare for sequence similarity, two *Bugula* species submitted by Hao *et al.* (2003) did not cluster together, nor were they in the same clade as *Bugula turrita* (Passamaneck and Halanych 2006) as expected (Figure 22). Also the method with which these sequences were obtained (DNA extraction and primers used) was not given in the paper.

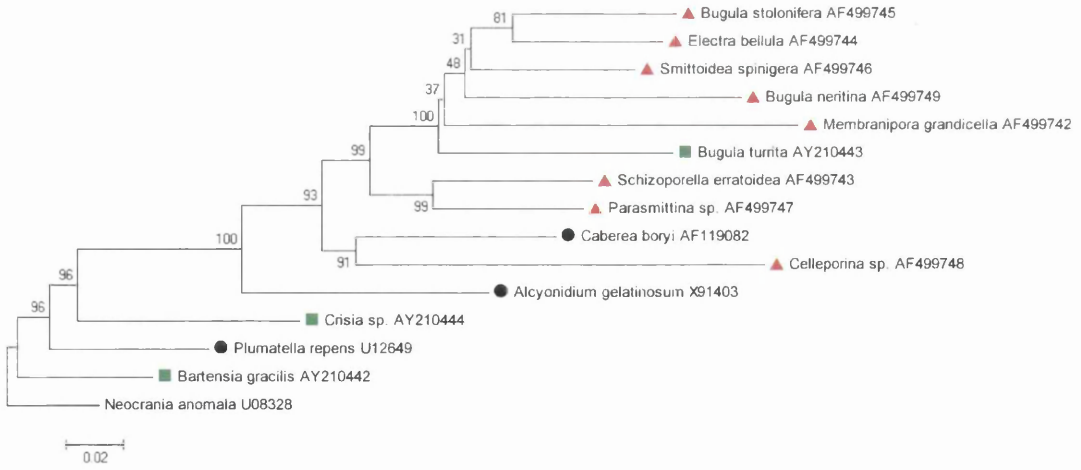


Figure 22 A NJ distance tree built using Kimura 2 parameter model, showing relative relationship between sequences of 18S Bryozoa present on NCBI. Bootstrap values (1000 replicates) shown at the base of the nodes. This tree was built to verify sequences identity and their relationship. Species are colour/shape coded by the submission author red triangle: Hao *et al.* 2003; green square - Passamaneck and Halanych 2004; black circle - all other authors, see Table 3 for details of these sequences. In this tree the sequences identified earlier (Waeschenbach 2003) as contaminants were excluded.

It was eventually decided to base the alignment for the primer design on the remaining five sequences (*Crisia* sp. accession no AY2120444, *Barentsia gracilis* accession no AY210442, *Bugula turrita* accession no AY210443, *Plumatella repens* U12649, *Caberea boryi* accession no AF119082). If new sequences were to

appear, they would eventually contribute to the alignment and thus more primers could be designed as needed.

Table 3 List of 18S primers sorted by the authors who previously submitted 18S Bryozoan sequences to GenBank. Sequences which have their accession numbers in bold font are considered to be valid, those in normal font are suggested contaminants and the status of those marked with an asterisk (*) are uncertain.

Sequence Author	Species	Accession no	Primers	Original Paper	Specificity of Primers
Passamaneck and Halanych 2006	*Crisia* sp.	**AY210444**	18e	Hillis and Dixon 1991	Universal 18S primer.
	Barentsia gracilis	**AY210442**	18P	Halanych *et al.* 1998	General Universal 18S primer derived from distantly related taxa such as yeast, mammals.
	Bugula turrita	**AY210443**	18h	Hillis and Dixon 1991	Universal 18S primer.
			18L	Halanych *et al.* 1998	General Universal 18S primer derived from distantly related taxa such as yeast, mammals.
			18M	Halanych *et al.* 1998	General Universal 18S primer derived from distantly related taxa such as yeast, mammals.
			18M0	Halanych *et al.* 1998	General Universal 18S primer derived from distantly related taxa such as yeast, mammals.

Sequence Author	Species	Accession no	Primers	Original Paper	Specificity of Primers
			18O	Halanych *et al.* 1998	General universal 18S primer derived from distantly related taxa such as yeast, mammals.
			18Q	Halanych *et al.* 1998	General universal 18S primer derived from distantly related taxa such as yeast, mammals.
			18Q0	Halanych *et al.* 1998	General universal 18S primer derived from distantly related taxa such as yeast, mammals.
Makey *et al.* 1996	*Alcyonidium gelatinosum*	X91403*	17 different primers	Winnepenninckx *et al.* 1994	Primers used for the phylogeny of Mollusca, but of universal origin.
Halanych 1995	*Plumatella repens*	**U12649**	18e	Hillis and Dixon 1991	Universal 18S primer.
			18h	Hillis and Dixon 1991	Universal 18S primer.
			18L	Halanych 1995	Designed based on distantly aligned taxa: yeast, invertebrates, vertebrates.

Sequence Author	Species	Accession no	Primers	Original Paper	Specificity of Primers
			18N	Halanych 1995	Designed based on distantly aligned taxa: yeast, invertebrates, vertebrates.
			18M	Halanych 1995	Designed based on distantly aligned taxa: yeast, invertebrates, vertebrates.
			18M0	Halanych 1995	Designed based on distantly aligned taxa: yeast, invertebrates, vertebrates.
			18P	Halanych 1995	Designed based on distantly aligned taxa: yeast, invertebrates, vertebrates.
Giribert *et al.* 2000	*Caberea boryi*	**AF119082**	1F	Giribet 1996	Arthropoda related study, but primers universal.
	Lichenopora sp.	AF119080	5R	Giribet 1996	Arthropoda related study, but primers universal.
	Membranipora sp.	AF119081	3F	Giribet 1996	Arthropoda related study, but primers are universal.
			18Sbi	Giribet 1999	Arthropoda specific 18S primer.
			5F	Giribet 1996	Arthropoda related study, but primers. are universal

Sequence Author	Species	Accession no	Primers	Original Paper	Specificity of Primers
			9R	Giribet 1996	Arthropoda related study, but primers are universal
Hao *et al*	*Bugula stolonifera*	AF499745*	Unknown	Unknown	Unknown
	Electra bellula	AF499744*			
	Smittoidea spinigera	AF499746*			
	Bugula neritina	AF499749*			
	Membranipora grandicella	AF499742*			
	Schizoporella erratoidea	AF499743*			
	Parasmittina sp.	AF499747*			

The first three sets of primers (18SF33S1 and 18SR998S1; 18SF449S2 and 18SR1346S2B; 18SF1088S3 and 18SR1871S3) were designed and tested with *Microporella ciliata* and *Schizomavella linearis*. The primers were designed in three pairs overlapping so that when three segments of the 18S gene were aligned the complete 18S sequence could be recovered. The overlapping segments of the primer pairs were designed in such a way so that they would avoid annealing to the highly variable regions of the alignment, and ideally overlap for at least 100 bp in order to simplify the eventual assembling of the three segments (Figure 23). Sometimes observed weakness of the signal during sequencing usually caused terminal ends of the sequenced segment to have more uncertain sites (recorded as N). However, when two segments overlap this allowed correction of these unresolved sites by means of comparing chromatograms of several segments.

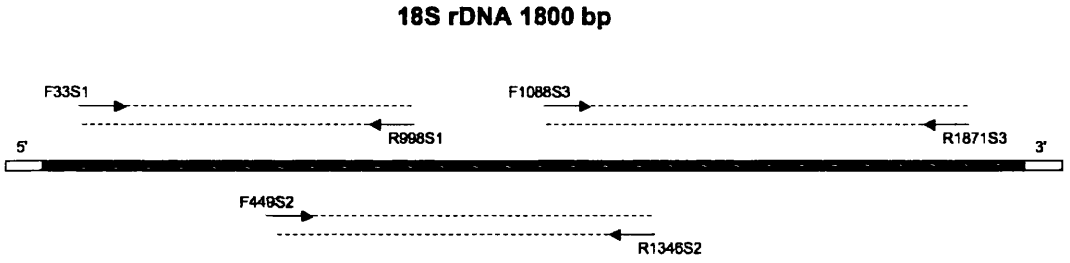


Figure 23 Three pairs of primers designed in this work and their relation to each other and the 18S gene (not drawn to scale). All three pairs were designed so that more than 100bp overlapping occurred. The name of the primer e.g. F1088S3 corresponds to the F – forward, number represents annealing site nucleotide number and S1-S3 stands for set 1 – set 3 (*i.e.* representing three overlapping sets of primers). Average length of the PCR amplified DNA fragment was 800bp. For a complete list of primers see Table 5 on page 65. The names of the primers in the main table follow the same rule, unless explicitly indicated.

The three segments were of approximate 800 bp in length so that when put together they would cover the entire length of the gene. Another possible strategy investigated was to perform a PCR for a longer segment of the gene and then use many nested primers for sequencing. However, this would require potentially too many primers as some species would not work with some primers.

Each segment pair of primers was marked as set 1, set 2 and set 3 corresponding to the beginning, middle and the rear section of the 18S sequence (in 5′- 3′ direction) respectively. Essentially, after the first two sequences were obtained a workflow was established whereby primers were tested against new DNA sequences, as such became available, and if the primers did not work new primers

were designed and tested (as described above) and then PCR performed on the new sequences (Figure 24).

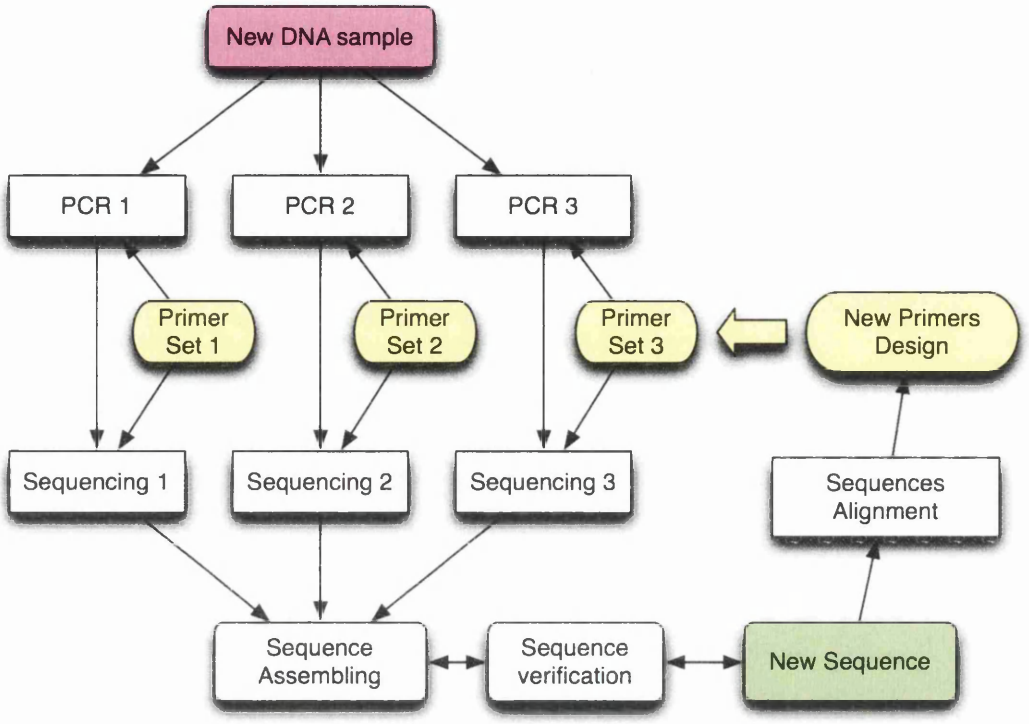


Figure 24 Primer design workflow, showing how new primers were tested against newly obtained DNA sequences of Bryozoa. The same primers were used for PCR and sequencing (unless cloning was used).

Eventually adding more sequences to the database of sequences and the sequences alignment made it possible to create more specific primers. In total 24 primers (both forward and reverse) were tested (Table 5). As the work progressed it became apparent that not all primers would work with all species and thus any attempt to create universal Bryozoa-specific primers was abandoned. No particular group of primers appeared to work with certain orders of Bryozoa, and no direct link to taxonomic affinity of primers was observed. For instance, the primers which worked perfectly with *Flustrellidra hispida* did not work at all with Alcyonidiidae species, thus no Ctenostomata specific primers were found.

Ironically, the genus with the most abundant larvae and embryos, *Alcyonidium*, appeared to cause the most problems with PCR and subsequently sequencing. All of the attempts to use the primers which belonged to set 2 (*i.e.* amplifying the middle section of 18S gene) failed. Also the beginning of the 18S gene which worked well with the standard set of primers (F33S1 and R998S1) was clearly not amplifying the right region (Figure 25).

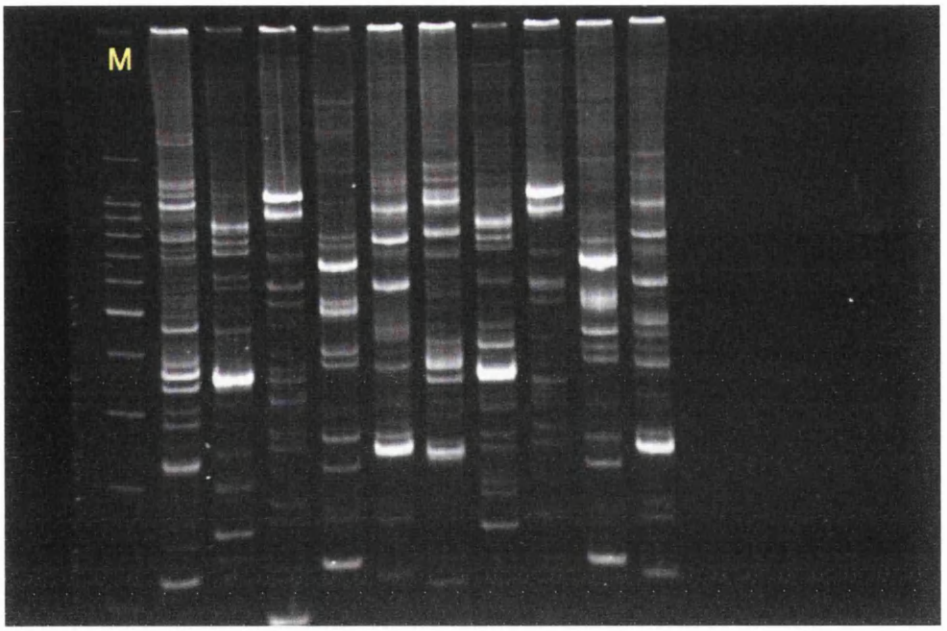


Figure 25 Unspecific amplification for *Alcyonidium gelatinosum* (first 5 lanes) and *Alcyonidium polyoum* (lanes 6-10) using F33 and R998 primers. (M) – 100 bp DNA ladder: 1500 bp, 1000 bp, 900 bp, 800 bp – 100 bp starting from top.

It was thought that using other Ctenostomata sequences as the basis of the alignment and developing primers based on these sequences would improve the specificity of the primers and help to produce the desired PCR results. However when the sequences of *Bowerbankia citrina, Bowerbankia gracilis, Bowerbankia imbricata, Flustrellidra hispida* and *Walkeria uva* (all belonging to Ctenostomata as *Alcyonidium*) obtained here were used as the basis for the primer design the results did not improve greatly. Eventually several pairs of primers were tested in different conditions and a group of primers was selected which covered two slightly overlapping regions (beginning and end of the gene) and worked well with all *Alcyonidium* species (Figure 26).

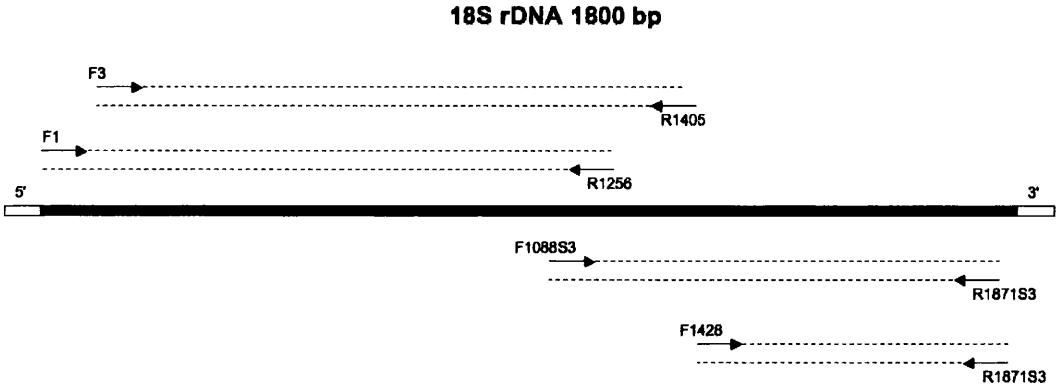


Figure 26 *Alcyonidium* specific primers (not all *Alcyonidium* primers are shown). From the map it can be clearly seen that not all combinations of primer pairs if sequenced would overlap well. For instance the segment of primer pair F3 – R1256 does not overlap with the segment of F1428 – R1871S3 primers, see text for more discussion.

Primer combinations which worked best with the *Alcyonidium* species are marked with an asterisk (*) in the main Table 5 which shows all primers designed. Unfortunately, although these primers seemed to work well and produced the desired results, when these sequences were assembled and tested using BLASTn search, for instance, they did not match any bryozoan sequences. The validity of these sequences is discussed in detail in the next Chapter. However, because these sequences appeared to be from contaminants the use of the primers which worked with *Alcyonidium* species tested here should be cautioned as they are clearly not specific enough to amplify *Alcyonidium*. It is clear from the results here and below that special attention to Alcyonidiidae has to be paid and more sequences are urgently required from Alcyonidiidae. In addition, some changes to the PCR conditions had to be implemented for the *Alcyonidium* species. For instance when *Alcyonidium polyoum* set 1 was optimised for the pair of primers F1 and R1256 the amount of $MgCl_2$ had to be increased to 2.5 mM instead of the usual 0.8 mM (these changes are indicated in the "Special notes" column of Table 5).

In total, 28 Bryozoan sequences were obtained in this work (15 Cheilostomata, 8 Ctenostomata, 5 Cyclostomata). In addition, sequences of two more Cheilostomata species *Celleporina hassallii* and *Cribrilina cryptooecium* were supplied by Dr Joanne Porter – these species were collected from the same site locations as the species collected for this work (Table 4). The complete sequences are also given in Appendix A.

Table 4 List of all species and sequences which were obtained in this work. Source column indicates AT – the author or JP – Joanne Porter (see text above)

Species	Family	Order	Source
Bicellariella ciliata	Bicellariellidae	Cheilostomata	AT
Bugula fulva	Bugulidae	Cheilostomata	AT
Bugula plumosa	Bugulidae	Cheilostomata	AT
Bugula turbinata	Bugulidae	Cheilostomata	AT
Callopora dumerilii	Calloporidae	Cheilostomata	AT
Callopora lineata	Calloporidae	Cheilostomata	AT
Callopora rylandi	Calloporidae	Cheilostomata	AT
Celleporina hassallii	Celleporidae	Cheilostomata	**JP**
Cribrilina cryptooecium	Cribrilinidae	Cheilostomata	**JP**
Escharella immersa	Escharellidae	Cheilostomata	AT
Escharoides coccinea	Exochellidae	Cheilostomata	AT
Haplopoma graniferum	Hippothoidae	Cheilostomata	AT
Microporella ciliata	Microporellidae	Cheilostomata	AT
Phaeostachys spinifera	Escharinidae	Cheilostomata	AT
Schizomavella linearis	Bitectiporidae	Cheilostomata	AT
Scruparia chelata	Scrupariidae	Cheilostomata	AT
Umbonula littoralis	Umbonulidae	Cheilostomata	AT
Alcyonidium gelatinosum	Alcyonidiidae	Ctenostomata	AT
Alcyonidium hirsutum	Alcyonidiidae	Ctenostomata	AT
Alcyonidium polyoum	Alcyonidiidae	Ctenostomata	AT
Flustrellidra hispida	Flustrellidridae	Ctenostomata	AT
Bowerbankia citrina	Vesiculariidae	Ctenostomata	AT
Bowerbankia gracilis	Vesiculariidae	Ctenostomata	AT
Bowerbankia imbricata	Vesiculariidae	Ctenostomata	AT
Walkeria uva	Walkeriidae	Ctenostomata	AT
Crisia aculeata	Crisiidae	Cyclostomata	AT
Crisia denticulata	Crisiidae	Cyclostomata	AT
Crisia eburnea	Crisiidae	Cyclostomata	AT
Filicrisia geniculata	Crisiidae	Cyclostomata	AT
Tubulipora liliacea	Tubuliporidae	Cyclostomata	AT

3.9 Primer database

Below is the table showing the complete list of primers which were tested and which worked in this study. The main primers which appeared to work well with most of the species are marked in bold. These main primers alongside the additional primers can be used to obtain 18S sequences of Cheilostomata, Ctenostomata and Cyclostomata for any future work. The additional primers (not highlighted in bold) could be used to substitute those main primers if any problems appear during the PCR. Series of sets are also indicated as set 1, set 2, set 3 and these should be followed if possible in order to produce the most overlap of sequences. The guidelines outlined for the PCR conditions should be sufficient to obtain strong clear bands of the desired weight. However, as noted above, *Alcyonidium* species appeared to have very specific response to the primers used in this study and more work is required and possibly new primers have to be tested.

Additional discussion of the *Alcyonidium* species obtained here is given in Chapters 4 and 5, and possible future work is discussed in the last chapter.

Table 5 List of all 18S rDNA primers used in this work, including universal primers from other authors. Main primers (*i.e.* those which worked with most bryozoan species) are marked in bold font. *Alcyonidium* specific primers are marked with an asterisk (*).

Primer Name	Empirical T_m °C	18S section	Direction	Nucleotide Sequence 5'-3'	Special Notes
18e	61.3	SET 1	For	CTGGTTGATCCTGCCAGT	Universal primer. Hillis and Dixon 1991
18L	77.3	SET 1	Rev	GAATTACCGCGGCTGCTGGCACC	Universal primer. Halanych *et al.* 1998
18N	52.8	SET 2	For	GTAATTGGAATGAGTCCA	Universal primer. Halanych *et al.* 1998
18M	61.3	SET 2	Rev	GAACCCAAAGACTTTGGTTTC	Universal primer. Halanych *et al.* 1998
18M0	61.3	SET 3	For	GAAACCAAAGTCTTTGGGTTC	Universal primer. Halanych *et al.* 1998
18PR	70.9	SET 3	Rev	TAATGATCCTTCCGCAGGTTCACCT	Universal primer. Halanych *et al.* 1998
18h	71	SET 3	Rev	AGGGTTCGATTCCGGAGAGGGAGC	Universal primer. Hillis and Dixon 1991

Primer Name	Empirical T_m °C	18S section	Direction	Nucleotide Sequence 5'-3'	Special Notes
F449	60.7	SET 2	For	TCTAAGGAAGGCAGCAGG	
R1105	60.8	SET 2	Rev	CCAGTCGGCATCGTTTA	
F33*	52.4	SET 1	For	TGTCTCAAAGATTAAGCC	
R637	62.4	SET 1	Rev	ACGCTATTGGAGCTGGAATT	
F1088	57.6	SET 3	For	GTAAACGATGCCGACTG	
R1871*	50.9	SET 3	Rev	AACCTTGTTACGACTTTT	
R1346	61.5	SET 2	Rev	CACCACCAACCACTGAATC	
F1*	66.1	SET 1	For	TACCTGGTTGATCCTGCCAGTAG	This primer when used in combination with R1256 primer required 2.5mM of $MgCl_2$ for *A.gelatinosum*.
R998	64.4	SET 1	Rev	CTTGGCAAATGCTTTCGC	
R1405*	55.9	SET 2	Rev	CGTTCGTTATCGGAATT	This primer worked well with Set 1 primers for *Alcyonidium* spp.
ALCF1049*	51.8	SET 3	For	GTTCTGACCATAAACGAT	This primer was developed specifically for *Alcyonidium* spp.

Primer Name	Empirical T_m °C	18S section	Direction	Nucleotide Sequence 5'-3'	Special Notes
ALCR1824	51.3	SET 3	Rev	GAAACCTTGTTACGACTT	developed for Alcyonidiidae only but did not work well.
ALCF28	51.7	SET 1	For	ATGCTTGTCTCAAAGATT	developed for Alcyonidiidae only.
ALCR807	51.7	SET 1	Rev	CTGCTTTGAACACTCTAAT	developed for Alcyonidiidae only.
ALCF422	59.2	SET 2	For	AACGGCTACCACTTCCA	This primer was developed specifically for *Alcyonidium* spp.
ALCR1347	57.5	SET 2	Rev	CCAGACAAATCGCTCC	This primer was developed for Alcyonidiidae.
R1977MBRY	68	SET 3	Rev	GATCCTTCCGCAGGTTCACC	This primer worked very well as an alternative to R1871.
F460	56	SET 2	For	GATCCTTCCGCAGGTTCACC	
F1015	75.5	SET 3	For	CCGAAGACGCCCTACTGCGAAAGC	
F1031*	67	SET 3	For	GCGAAAGCATTTGCCAAGAAT	This primer worked well with *Alcyonidium* in pair with R1871, but at T_a = 50°C.

Primer Name	Empirical T_m °C	18S section	Direction	Nucleotide Sequence 5'-3'	Special Notes
F1428*	68.2	SET 3	For	ACGAACGAGACTCTTGCCTGCTA	This primer was used only for *Alcyonidium* spp.
R1103	65.6	SET 2	Rev	CGGTATCTGATCGCCTTCG	
R1256*	65.4	SET 2	Rev	CCTTCCGTCAATTCCTTTAAGTTTC	This primer worked with Set 1 *Alcyonidium* primers.
R1736	62.2	SET 3	Rev	CCACTCAATCGGTAGTAGCG	

4 SEQUENCE ALIGNMENT AND SECONDARY STRUCTURE RECONSTRUCTION

4.1 Reasons of selection of the method which was used

This chapter is concerned with the secondary structure alignment of the 18S rRNA sequences obtained in this work. In several studies preceding this dedicated to bryozoan phylogeny, 18S rRNA has been used as the gene for phylogenetic reconstruction. However, to the best knowledge of the author, so far there has not been any published secondary structure for Bryozoa, no previous attempts have been made to reconstruct the bryozoan SSU rRNA secondary structure for the purpose of the alignment. Due to the fact that secondary structure reconstruction is a tedious, mostly manual process, the details of this are given below. At the end of this chapter the secondary structure of *Bugula turbinata* is presented.

4.2 Methods used for SSU secondary structure reconstruction

The topic of the secondary structure assisted alignment of 18S rRNA was actively discussed at the Evolutionary Directory (EvolDir) forum[21]: and a consensus was that most authors tended to use manual methods for the alignment. One particular of such interest is jRNA project (http://hymenoptera.tamu.edu/rna/index.php), which gives a rather detailed tutorial on how to align the sequences using secondary structure. This tutorial is based on Kjer (1995) alignment strategy alongside with the email communications from Kjer—available at the jRNA website and through Google search engine, which gave some indications on how the alignment should be performed. The procedure of alignment in essence boils down to a manual arrangement of all sequences in a text editor capable of colour coding of individual nucleotides, starting with completely unaligned sequences. This method may well be preferable for smaller RNAs molecules such as those described on the jRNA web site tutorial, but using this completely manual method for alignment of over 30 sequences with length around 2000 bp was not very practical due to time limitation. Consequently, an alternative method was sought, which would if not automate fully at

[21] http://evol.mcmaster.ca/brian/evoldir.html

least would allow a partial automation during the initial sequence alignment. The method, which was adopted here is believed to be employed by other researches (EvolDir forum, personal communication) although here some modifications were made. An extrapolation of the Kjer (1995) method was offered (Telford *et al.* 2005) where data from the European ribosomal RNA database (ERRD) were used as a secondary structure alignment.

4.3 Detailed description of the alignment method used

Secondary structure can be predicted in different ways and several methods have been proposed and discussed. In this work the method adopted is an amalgamation of several methods described above, including the one which uses software for DCSE file manipulation, developed by Telford (see above).

For aiding the secondary structure alignment, an already published reference alignment was used. The species, which were used for the reference alignment, were selected based on their relatedness to Bryozoa from the ERRD[22] (Van de Peer 1999; 2000). Obviously, sequences of other Bryozoa species were sought, as well as ones of those taxa which are believed to be closely related to Bryozoa.

4.4 Description of ERRD database

The ERRD database was built on the sequences acquired from GenBank and EMBL. Support for this project appears to have been discontinued at the time when few bryozoan sequences were deposited into GenBank. The database was last updated in 2002 (Jan Wuyts, personal communication), which would explain the absence of many newer sequences. Personal communication with the authors confirmed that it is no longer maintained and instead just kept due to its high popularity and demand from the scientific community. Consequently, the sequences which are present in the ERRD are those which were deposited into NCBI before 2001. The sequences are retrievable in the form of an alignment, which also contains secondary structure encoded into the alignment file. Apart from the sequences and secondary structure the database stores identification codes such as accession number (identical to that of

[22] The ERRD was originally based in University of Antwerp (Belgium) and has now moved to a different virtual location. Its URL has changed from that of the published links. The new link could be located through Google search by using keywords: "european ribosomal database".

NCBI). The sequences can be selected through a menu from the list of various taxa and downloaded separately or in combination with any selected taxa.

The database had grown from just over 50 sequences (Huysmans and Dewachter 1986) to well over 3000 sequences by 2000 (Wuyts *et al.* 2000). In order to add sequences in the database and prepare secondary structure alignment the researchers used the software package DCSE[23] (DeRijk and Dewachter 1993)—Dedicated Comparative Sequence Editor—now defunct (de Rijk, personal communication) and only available by contacting the author. This package although still possible to acquire is essentially useless as it cannot be easily compiled and installed on any UNIX/Linux clones and if installed on Microsoft DOS environment proves to be very difficult to use due to its rather outdated interface and complete lack of support for the package. It also lacks any conversion utilities and hence any potential use is restricted to itself, without any possibility to extract aligned sequences for further analysis in other packages.

The ERRD has a limited number of species in it from the selection of SSU section from Metazoa. Several available species were downloaded in DCSE format. These were two Bryozoa species: *Plumatella repens* (accession no. U12649); *Alcyonidium gelatinosum* (accession no. X91403); three entoprocts: *Barentsia benedeni* (accession no. U36272), *Barentsia hildegardae* (accession no. AJ001734), *Pedicellina cernua* (accession no. U36273). In addition, two Brachiopoda species were selected from Lophotrochozoa: *Neocrania anomala* (accession no. U08328), *Neocrania huttoni* (accession no. U08334). A complete list of taxa used in the alignment is given in Table 6 on page 99. The sequences being in DCSE file format could not be read by the commonly used software such as MacClade, PAUP* and ClustalW and therefore had to be converted to a common format accessible to all these applications such as Nexus or similar. The software applications Xstem and Ystem (Telford *et al.* 2005) ware used for data conversion and general negotiation with DCSE file format and ERRD.

Sequences, once downloaded, were converted using Ystem utility to Nexus format for editing in MacClade and ClustalW applications. For the purpose of the alignment, the sequences were imported into MacClade and then one by one my Bryozoa sequences were added and aligned to the reference alignment of sequences

[23] See Appendix B for details of DCSE file format and associated software problems.

from the ERRD database using the pairwise aligner embedded in MacClade—MacClade does not allow multiple sequence alignment. This preliminary alignment was necessary to speed up further with manual alignment. The use of the embedded pairwise aligner in MacClade is simple and gave good results.

Once aligned the sequences were exported in Nexus format and imported into a word processor—in this case Microsoft Word was used due to its ability to colour code individual nucleotides—following the method which was suggested by Kjer (see above). Subsequently, the sequences were manually aligned using the reference alignment as a template together with the published secondary structure models from ERRD (such as *Daphnia pulex*).

A helix was examined manually to be aligned between all species: if all sequences were identical to the reference alignment then the pairing nucleotides were underlined in pairs both in the 5′–3′ part of the helix and its counterpart. If any disagreement in basepairs was found it was checked using Mfold web server software (Zuker 2003; Mathews *et al.* 1999) for hydrogen bonding. Mfold is a widely used program which employs an algorithm of RNA secondary structure prediction by means of calculating a minimum free energy (ΔG) required for folding and maintaining a certain base pair in the secondary structure (Zuker 2003). For instance: GUC would be checked to pair with GAC thus C:G, U:A, G:C pairs would form. Some helices had compensatory[24] substitutions. Non Watson–Crick pairs were examined manually with the aid of Mfold program, and if the secondary structure built with Mfold corresponded to that of the reference alignment, formation of non-canonical pairs could preserve the proposed secondary structure and therefore was accepted. The base pairing once checked and accepted to be correct was underlined (this technique was proposed by Kjer and later by jRNA). In cases where the secondary structure of the reference sequences, acquired from ERRD, did not match sequences obtained here, each helix was examined using Mfold software and the secondary structure folding was accepted as per the Mfold results. Using the above methods all helices were manually examined and their alignment adjusted in the text file.

[24] A compensating substitution in an alignment is defined as a complete replacement of one base pair in a sequence by another complementary base pair at the same position in the sequence. For example a substitution of A:U pair by G:C is compensating, whereas A:U by G:U is not as only one base pair changes (Wuyts *et al.* 2000).

Subsequently, the sequence was exported back into MacClade for final visual examination and export to either Ystem utility to convert it back to DCSE format; or for a direct export to Nexus format for further phylogenetic analysis. This step was necessary as sequences had to be converted to NBRF format in order for them to be recognised by Ystem. An export to DCSE format was necessary in order to preserve secondary structure. When Ystem performs parsing of an original DCSE file (*i.e.* the one which contained the reference alignment with secondary structure symbols) and newly aligned sequences, the utility introduces the same helices into the file mask[25] which were present in the original ERRD extracted file. This includes extra helices, which are not present in the actual alignment and have to be manually removed. In case of alignment used here the following helices had to be removed: E8_1, E8_1′; E23_5, E23_5′; E23_6, E23_6′; E23_15, E23_15′; E23_16, E23_16′; E23_17, E23_17′; E45_1, E45_1′. Once in DCSE format, the file was manually edited to insert all DCSE specific secondary structure elements (as per DCSE format) for all sequences; this is a tedious process, which requires manual editing of a large DCSE text file. See Figure 27 for a comparison of DCSE file before and after manual editing. Once done this file could finally be converted—using another utility Xstem—to a Nexus format, which included separate blocks of stems and loops based on the secondary structure. The above lengthy procedure is essentially a manual way of processing data, which could be done using software (such as DCSE) if such was still available.

[25] File mask is an additional line of text, which is placed under the alignment in DCSE file format. This line of text contains markings, which identify individual stems of helices by numbers. For details of DCSE file format please see Appendix B.

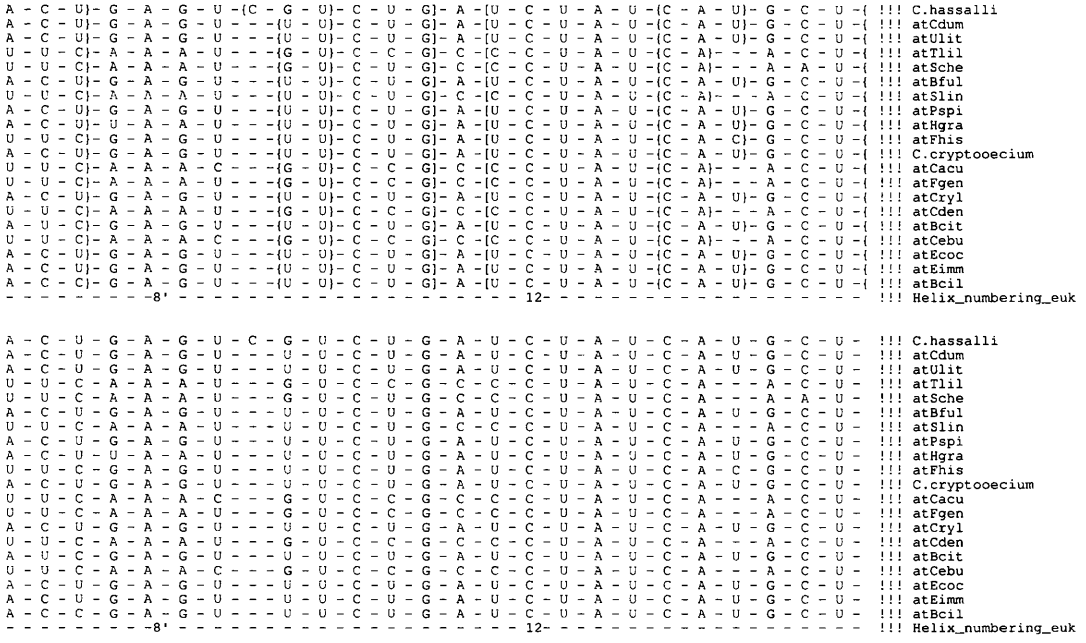


Figure 27 An example of the DCSE file after (above) and before (below) the manual entering of secondary structure symbols required for DCSE file format. Symbols [ ] represent beginning and end of a new helix; symbols { } represent beginning and end of an internal loop or a bulge loop interrupting a helix. Note a helix numbering line underneath the nucleotides columns.

4.5 Principles of hierarchical structure and organisation of RNA

4.5.1 Types of nucleotide pairing

Ribosomal RNA exhibits complex interactions between paired nucleotides. Apart from the canonical Watson–Crick base pairing also found elsewhere, more complex non–canonical interactions are observed. Whilst tRNA X–ray crystallography allows non–canonical types of interaction to be inferred this appears to be not possible for rRNA molecules which are much larger (Gutell *et al.* 1994), therefore a multiple sequences comparison technique using the covariation analysis is used to predict base pairing. Base pair covariation is defined as a base substitution in one column of the alignment, which is influenced by another base in a different column of the same sequence (Wuyts 2000). As a result, base covariation can show presence of secondary and even tertiary interactions.

A very detailed treatment of this method is given by Gutell *et al.* (1992), but here it is sufficient to mention that the covariance Cramer's φ index is estimated using the following formula:

- $\varphi = \sqrt{\frac{\chi^2}{n(k-1)}}$

Where: χ^2 is calculated from a 4x4 table (which contains all possible nucleotide combinations); n- number of sequences; k- number of columns or rows in the table. φ assumes values between 0 and 1 with 1 being the strongest possible covariation (Wuyts *et al.* 2000; Gutell *et al.* 1992; Sokal and Rohlf 1995). In the case of the Wuyts *et al.* (2000) publication, the table which gives base pair interactions for V4 area shows Cramer's index alongside with other statistical estimations of potential base pairing.

The following non–canonical base pairing has been reported for rRNA:

(1) G:U pairs are frequently observed in rRNA and sometimes reflect highly conserved regions (Gutell *et al.* 1994; Leontis *et al.* 2002). This type of pair carries functional importance and was reported to be responsible for helix stacking and found in specific locations (Buckley *et al.* 2000).

(2) Another less common non-canonical pairing is C:A, which is a subclass of G:U pairing and was observed to interchange with U:G pairs (Gutell *et al.*1994; Hickson *et al.* 1996).

(3) Additionally, G:A , G:G, and A:A pairs have been reported by some authors (Gutell *et al.*1994; Leontis *et al.*2002). Although reported to be less frequent, they are found terminally in loops or in the interior of some helices and believed to play a key role in rRNA structure (Gutell *et al.*1994). These less common pairings are observed in a different orientation to the Watson-Crick pairs and called trans orientation. Although this topic is beyond this work, a very detailed treatment is given in Leontis *et al.* (2002). It is important to be aware of this type of base pairing in order to assess the validity of some proposed secondary structure components and to aid in alignment.

4.5.2 Types of secondary structure interactions

rRNA forms complex secondary and tertiary structures, the correctness of which is facilitated by the r–proteins, which allow the rRNA to be folded in the ribosome (Lafontaine and Tollervey 2001; Tinoco and Bustamante 1999). Secondary structure is energetically more stable than tertiary structure and thus can sustain itself. The difference between secondary and tertiary structure is empirical and could be shown

by drawing folded rRNA on a 2D plane and connecting all pairings by lines (see Figure 28), if there is any crossing of these lines the interaction is considered to be tertiary, thus pseudoknots belong to the tertiary interactions (Chastain and Tinoco 1991).

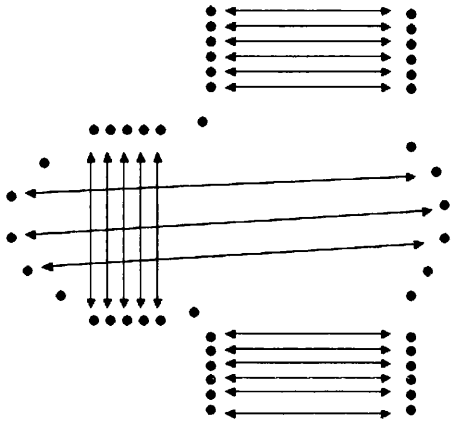


Figure 28 Two-dimensional interpretation of the tertiary interactions such as pseudoknots. On this diagram, each dot represents a potential nucleotide, which is paired to its counterpart (marked as a double-arrowed line). Tertiary interactions though cross ordinary pairing lines (also marked with arrowed lines). See text for further discussion. Image modified from Chastain and Tinoco (1991)

The formation of secondary structure is influenced by the thermodynamics of bonds and temperature as well as ion concentration. Several programs of thermodynamic folding exist but for this work Zuker's program Mfold (Zuker 2003; Mathews *et al.* 1999) was used. The algorithm of Zuker's program is based on experimental data (Tinoco and Bustamante 1999) and therefore was favoured by the author. In addition, this program is accessible through the Internet and thus does not require any specific operating system. Several secondary structure elements have been shown to exist in rRNA.

Duplexes (Figure 29) are the main components of Watson–Crick type interactions and form double helix interactions between base pairs. The rest of the interactions are placed into single–stranded regions (Kjer 1995; Tinoco and Bustamante 1999).

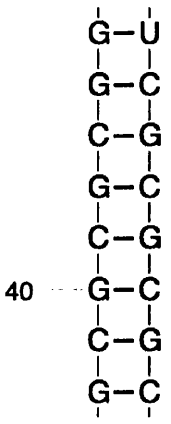


Figure 29 Example of a duplex (see text for details)

Hairpins (Figure 30) consist of a double helix terminating in a loop of unpaired nucleotides, which are known to bond proteins (Chastain and Tinoco 1991; Tinoco and Bustamante 1999). Some researches identify tetraloops, hairpins formed by four nucleotides, as a specific subclass of hairpins (Gutell *et al.* 1994).

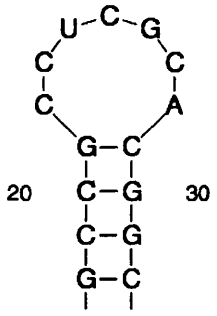


Figure 30 Example of a hairpin (see text for details)

Tetraloops (Figure 31) are remarkable as they are commonly constrained and only a few nucleotide sequences with specific types of interaction appear to dominate tetraloops: UUCG, CUUG, GAAA or GCAA.

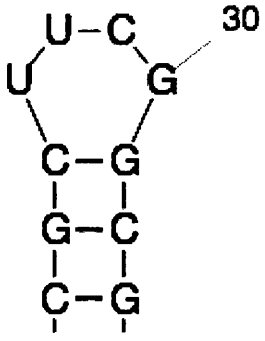


Figure 31 Example of a tetraloop (see text for details)

This is linked to their hydrogen bonding interactions. Several types of loops could be identified: bulges with unpaired nucleotides, which appear on one side of a double

helix, with a single nucleotide bulge being a subclass. These structures can affect the shape of the secondary structure by introducing bends.

Internal loops, also called mismatches, are very important as they also include all non Watson–Crick base pairing, thus G:U for example, being part of a mismatch. Internal loops with three or more nucleotides are also mismatches and are very common in the secondary structure. Internal loops could be symmetrical, with an identical number of nucleotides on each side, or asymmetric. Crucially for secondary structure it is not known what causes some loops to stay open or closed by forming non Watson–Crick interactions (Chastain and Tinoco 1991). Presence of internal loops may case formation of some non-canonical base pairing if these loops are to be closed and raises a question on how they could be interpreted. Should they be treated as an open loop or can they be treated as a new base pairing? In the alignment which is described below there have been several instances when certain internal loops could be closed or made smaller if some non–canonical base pairing were to be accepted.

4.5.3 Numbering system

The originally defined secondary structure for 16S–like rRNA was done by Gutell *et al.* (1985) with definition of a universal core (and hence universal helices), which was common to all 16S–like rRNA secondary structures. The phylogenetically constrained core was similar in most groups of eukaryotic and prokaryotic organisms. However the number of universal helices was discovered to be highly variable from 25 helices in flagellates to 53 helices in plants in mitochondrial rRNA. Because of such variability of numbers of universal helices, a numbering system was used which would correspond to most taxonomic groups and hence 48 universal helices were defined for eukaryotic rRNA.

The numbering system for the helices is identical for both prokaryotic and eukaryotic rRNA and was originally proposed by Nelles *et al.* (1984) based on the secondary structure rRNA models of Woese *et al.* (1983) and Gutell *et al.* (1985). The system essentially stayed the same with the exception of new helices being added as the knowledge about sequence variability increased and the number of recognised helices consequently changed from 40 to 50.

Below is a slightly more detailed description on how the system has evolved, which is important to this study as the alignment was based on a certain type of secondary structure model. The original numbering for eukaryotic organisms was for

Artemia salina (Nelles *et al.* 1984) and contained 40 helices. The numbering was based on the order of occurrence of the helices from the 5′ proximal strand. Each individual number was given to an area such as uninterrupted helix or set of helical segments, which could be connected by bulges or internal loops. Different numbers were given to helices if they were separated by a multibranched loop, a pseudoknot, or a single stranded area that does not form a loop (Nelles *et al.* 1984; Van de Peer 1999). The number of universal helices has grown from 40 (Nelles *et al.* 1984) to 50 (Neefs *et al.*1991). Some helices are specific only to eukaryotic organisms and in order to distinguish them they were labelled with `Ea_b` (**E** for **E**ukaryotic as opposed to **P** for **P**rokaryotic helices) where "`a`" shows the preceding universal helix and "`b`" an order number of the specific E helix. The example of E23_1 (see Figure 32) shows Eukaryotic specific helix 1, which follows universal helix 23.

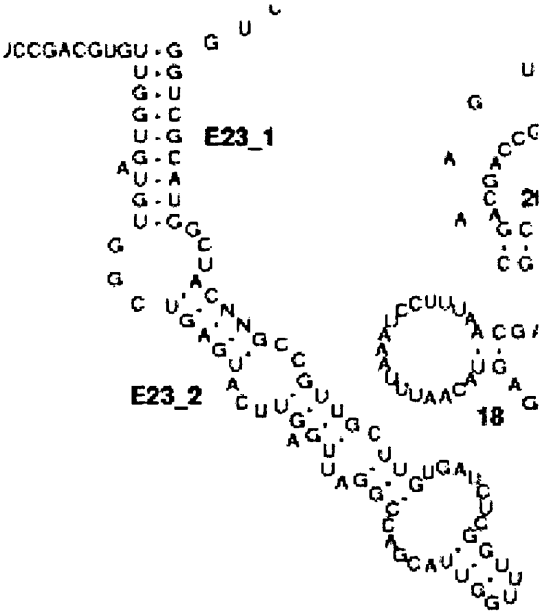


Figure 32 Eukaryotic specific helix E23_1 (see text for details)

Not all 50 universal helices have to be present in rRNA sequences—some taxonomic groups were reported to have shown anomalies and lack certain helices, for instance microsporidians and trichomonads (Van de Peer 1999).

4.5.4 Features which could be identified in the rRNA

Conserved regions within rRNA vary based on their relative position in the RNA or a complete region. These can be linked to the function of the region and its interaction within the ribosome and with the surrounding proteins.

4.6 Number of helices described for the secondary structure has changed

The 18S rRNA alignments based on the secondary structure have "evolved" and changed with new sequences becoming available to researchers. During this period several changes occurred as described below.

Change occurred in the number of helices due to the accumulation of data. For instance the first description of 18S secondary structure of *Artemia salina* and the proposal of its secondary structure (Nelles *et al.* 1984) shows only 40 helices for 18S as well as E9_1 and E19_2 helices which correspond to eukaryotic only regions. It is notable that at the time of the above publication there did not seem to be much consensus on the general secondary structure model for 18S rRNA. Further work was somewhat based on Nelles *et al.* (1984) publication, for instance Ellis *et al.* (1986) which showed *Caenorhabditis elegans* 18S secondary structure. A very important work is that of Gutell *et al.* (1985) where a good review of up to date information on 16S–like rRNA is given with over 20 species and secondary structures shown. Many authors based their work on this publication (Hepperle *et al.* 1998; Abouheif *et al.* 1998; Flynn and Nedbal 1998; Brown and Pestano 1998; Carranza *et al.* 1997; Alvesgomes *et al.* 1995).

Another update of helix numbering happened with the publication of over 30 18S rRNA sequences and secondary structures by Dams *et al.* (1988). This review described 48 universal helices and included eukaryotic only helices E10_n and E21_n. The number shift from E9_n to E10_n was due to the discovery of a pseudoknot in helix 1, which split it into helices 1 and 2. This work also identified nine V1–V9 hypervariable regions (see below for detailed description). This specific publication is especially important as the secondary structure model proposed for the eukaryotes has been widely used by many researches even though newer and revised models (see below) were proposed. Of special importance is helix 19 (Figure 33), which has persistently been used in many alignments. In another major work a revision of more than 452 small rRNAs (Neefs *et al.* 1991) was made and as a result the interaction previously recorded as helix 10 was rejected, previously named helix E10_1 was renamed helix 10. Because of that, the old helices E10_2 and E10_3 shifted their numbering and became E10_1 and E10_2 respectively. This however did not change the overall number of helices, which still remained at 48 for 18S rRNA. In

the above study the first questions were asked about the validity of region V4 (i.e. helices E21_1 to E21_9) and a tentative structure was proposed. Finally, Neefs *et al.* (1993) changed the number of helices in the secondary structure from 48 to 50 due to the separation of the universal helix 19 into helices numbered 19, 20, 21. This was done due to the pseudoknot structure loop, which was described previously by Woese and Gutell (1989).

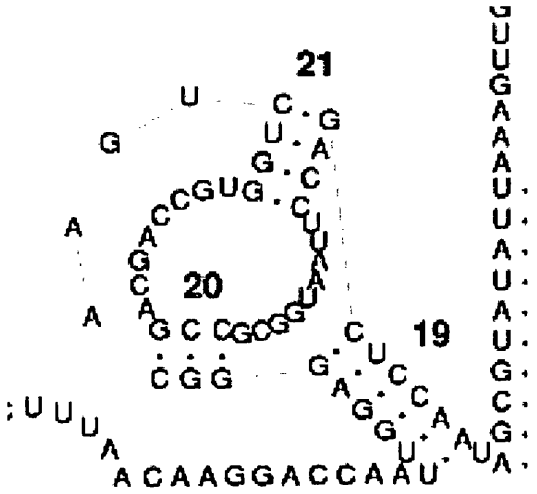


Figure 33 Helix 19, which was later split into Helices 19, 20, 21 due to the pseudoknot.

Unfortunately, many publications failed to notice the amendment made by Neefs *et al.* (1993) and persisted in using the old folding for helix 19 (*i.e.* without a pseudoknot). After the above amendments the secondary structure of eukaryotic 18S rRNA takes the numbering as it is presented here and as used as a reference for the purpose of the alignment of bryozoan sequnces. Although given the history of changes it is possible that the secondary structure may change again, when over 3000 sequences were recently examined by Wuyts *et al.* (2000) they did not make any changes to the universal helices. However, they proposed a change to the hypervariable area V4 and its corresponding helices: E23_1 to E23_9. This change of V4 area is treated below in more detail as it affected the alignment of sequences in this work.

4.7 Hypervariable regions

Hypervariable areas have always presented many problems both for the primary structure alignment as well as for the secondary structure. Because of the difficulty of

aligning these areas, researchers most often disregard these areas when using phylogenetic reconstruction.

The number of hypervariable regions has remained largely unchanged with the addition of more sequences. Originally, Nelles *et al.* (1984) showed seven hypervariable areas for eukaryotic rRNA (V1–V7), five of which corresponded to those of *Escherichia coli* (Spenser *et al.* 1984) and wheat mitochondrial rRNA. Later Dams *et al.* (1988) distinguished eight variable areas V1–V9 for eukaryotic cytoplasmic rRNA (but area V6 is missing in eukaryotes and only present in prokaryotic rRNA). In his work it is emphasised that area V4 is missing from prokaryotic rRNA and being highly variable in eukaryotic organisms, the author also adds that further addition of sequences to area V4 should pour some light on the number of helices in that area. The V4 hypervariable area is defined by E21_1 to E21_7 helices. The number of variable areas is confirmed by other researchers (Neefs *et al.* 1991, 1993; De Rijk *et al.* 1992; Kjer 1995). Neefs (1991) pointed out that many insertions and deletions are observed in the areas V2, V4 and V8 corresponding to helices 10, E21, and 43 respectively.

Finally, it was shown based on more than 3000 sequences (Wuyts *et al.* 2000) that area V4 corresponding to the helices E23_n (former helices E21_n – see above) contains two pseudoknots (instead of the previously reported one) and hence the number of helices in this area was proposed to be changed to fourteen (E23_1 to E23_14) in some organisms – up from the previously reported nine. The analysis of these areas was made using covariation Cramer's φ index, also called coefficient of association or phi coefficient (Sokal and Rohlf 1995). This coefficient allows determination of the strength of the relationships between two variables; Wuyts *et al.* (2000) analysis showed the presence of several new helices in eukaryotic organisms, with most eukaryotes having the following helices in the E23_n V4 region: 1,2,4,7–14. In this work the reference alignment, which was downloaded from the European Small Ribosomal database, utilises the above secondary structure model in relation to the V4 region.

4.8 Individual helices description for Bryozoa

Below follows a more detailed treatment of individual helices and their comparison to the reference alignment. Each helix has a 5′–3′ stem and a 3′–5′ stem; for the purpose

of the description of individual helices the 5′–3′ stem is named as n-stem and 3′–5′ named n′-stem respectively (where n refers to the number of the helix).

4.8.1 Helices 1–3

These helices were not recorded for all species. Although well conserved when present they were only observed for *Bugula turbinata*, *Cribrilina cryptooecium*, *Bowerbankia gracilis*, *Microporella ciliata* and *Escharella immersa* with some nucleotides missing or recorded as N – during sequencing. When present and complete they were identical to the reference alignment. In the case of *Escharella immersa*, 1′–stem was GCAG instead of GCCAG, therefore it was recorded as GDCAG, with D standing for missing nucleotides. This file format, described above, which is created for DCSE aligner was adopted here for the purpose of the compatibility with the Xstem/Ystem conversion tools. The beginning of the sequences was not available for many species because they overlapped with the primers used (see Chapter 3 for description of primers).

4.8.2 Helix 4

The 4-stem of the helix was identical between species for those seven sequences (out of the total of 30 species) where nucleotides were present; the remaining sequences did not have this part of the helix or appeared to have scrambled sequences. See above section for helices 1-3 for possible explanation. The 4′–stem of this helix was identical for all species. For location of helices in relation to the whole 18S rRNA see Figure 38, which shows a complete gene folded for *Bugula turbinata*.

4.8.3 Helix 5

The 5–stem of this helix was identical for all sequences when present but missing or scrambled for seven species. *Bowerbankia imbricata* had an insertion GC **A** C (shown in bold) which probably was caused by scrambling of the sequence. The 5′–stem was identical for all species.

4.8.4 Helix 6

The 6–stem of this helix was not complete for all sequences and was missing or scrambled for ten species. Although not completely identical to the reference sequences some strong similarity was observed among stems. The length of the

terminal loop varied between species. Thermodynamic folding when constrained to the reference sequences failed with Mfold error: "job aborted'. Consequently, folding was accepted as per the Mfold folding. See Figure 38 for examples of folding, most sequences were similar to those of *Bowerbankia gracilis*.

4.8.5 Helix 7

The 7′–stem of this helix was identical for all species whereas the 7–stem had two species different: *Celleporina hassallii* and *Haplopoma graniferum*, UAU and UUC respectively. In the case of *Haplopoma graniferum* the changes were compensatory in relation to 7′–stem. However, for *Celleporina hassallii* the changes could not be accepted as per standard base pairing. Notably, this species had many differences in its entire sequence such as insertions and deletions, which could have been specific to this species, or possibly, could have been caused by sequencing errors. As this sequence was not sequenced by myself—supplied to me by Dr J Porter—the validity of assembling this sequence could not be verified.

4.8.6 Helix 8

Although the sequences aligned well to the reference sequences it was difficult to fold this helix using Mfold constraints. All mutations in the stems of the helix were compensatory. For instance, *Callopora dumerilii* and *Scruparia chelata* had CUG in the 8′–stem, but CGG and CAG in the 8–stem respectively, thus forming the following pairs: C:G, A:U, G:C, for *Scruparia chelata* and C:G, G:U, G:C, for *Callopora dumerilii*. An example of secondary structure thermodynamic folding using Mfold is given for *Bowerbankia gracilis* (Figure 34). Some sequences had insertions in the bulge region (*e.g. Bowerbankia imbricata*, *Celleporina hassallii*), or deletions (*e.g. Schizomavella linearis*).

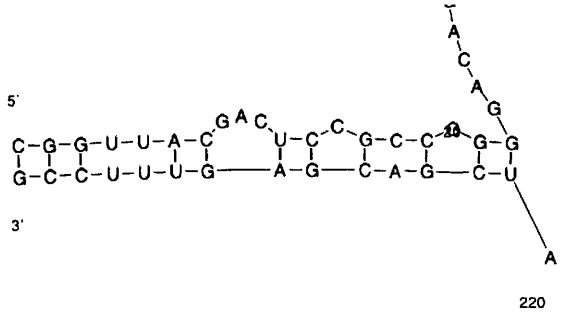


Figure 34 Helix 8 of *Bowerbankia gracilis* folded using Mfold with necessary restrictions.

4.8.7 Helix 9

The 9–stem was almost identical to the reference alignment with some insertions in the internal bulges and hairpins (*e.g. Celleporina hassallii*); all substitutions were compensatory and did not change the secondary structure. The only problematic species was *Celleporina hassallii* due to insertions in 9–stem constraining folding of Mfold resulted in aborting of the computation. Therefore, the folding was done without constraining the internal bulge, which is reflected in the images (Figure 35). Two alternative structures are given. Neither of the structures appears to agree with the reference secondary structure.

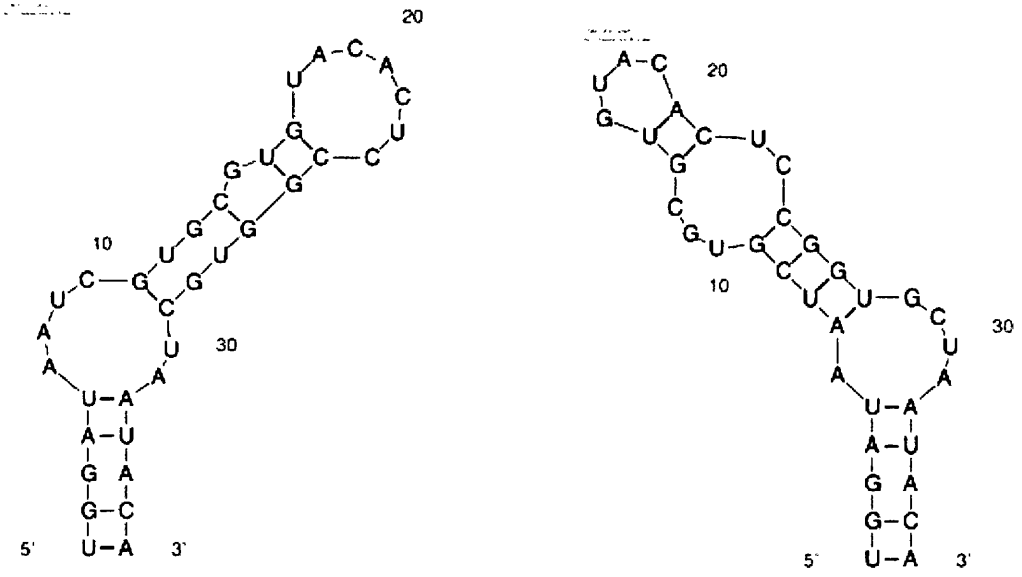


Figure 35 Two alternative holdings for Helix 9 of *Celleporina hassallii*.

4.8.8 Helix 10

The proximal parts of this helix (both 10–stem and its corresponding 10′–stem) were identical to the reference sequences. By contrast, the distal part of this helix for both 10 and 10′–stems differed from the reference sequences considerably, and showed some expansion and contraction. However all sequences preserved the central bulge and showed similarity in their secondary structure. Due to the differences from the reference sequences the alignment was based on the thermodynamic folding with Mfold.

4.8.9 Helix E10_1

This eukaryote–specific helix could not be aligned against the reference sequences, and differed considerably especially in the distal area, showing, as in the previous helix, expansion/contraction towards its distal part. As the folding could not be guided by the reference alignment, it was done as per Mfold algorithm. Secondary structure was persistent among most of the species, and showed two internal bulges and a long distal stem. The terminal loop was four nucleotides in all but one of the species (*Schizomavella linearis*).

4.8.10 Helix 11

This helix showed almost no variation for both 11 and 11′–stems. The secondary structure corresponded to that of the reference alignment, with some compensatory substitutions present in both stems. The secondary structure was identical for all species with the exception to *Scruparia chelata*; there was some variation in the terminal loop among all sequences.

4.8.11 Helix 12

Both 12 and 12′–stems showed some differences from the reference alignment, but all differences were compensatory and did not alter the secondary structure, which was in agreement with the reference alignment. Thermodynamic folding based on Mfold, using reference constraints was not possible and resulted in failure of the program to complete. Consequently, the folded structures differ both in the stems and in the loop area from those recorded in the reference alignment. For instance for *Bowerbankia gracilis* (Figure 36) the terminal loop, as predicted by Mfold, was **GCUGU**, however the reference alignment placed it as …G)CUG(U… that is the terminal G & U nucleotides were included in the stems and not in the terminal loop.

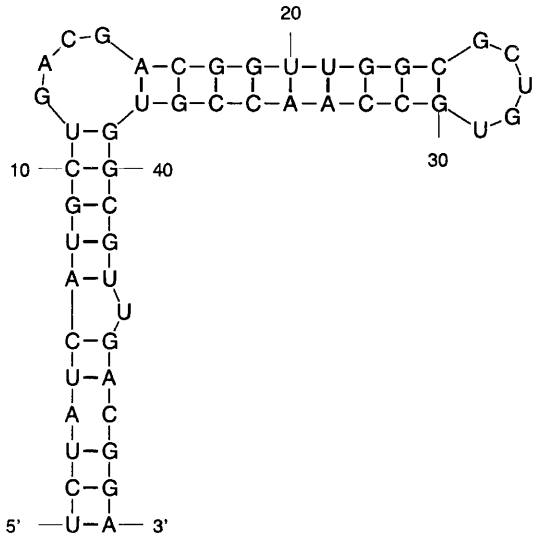


Figure 36 Helix 12 folding of *Bowerbankia gracilis* see text for details.

4.8.12 Helix 13

All species showed similar secondary structure with the exception of *Celleporina hassallii*, which had an insertion in the 13–stem.

4.8.13 Helix 14

All secondary structures were identical with some compensatory mutations present in *Bowerbankia gracilis*, *Bowerbankia imbricata*, *Schizomavella linearis* and *Flustrellidra hispida*.

4.8.14 Helix 15

This is a very short helix, with almost all species displaying identical sequences in both stems. The exceptions were *Scruparia chelata* and *Schizomavella linearis*, which had compensatory mutations in the 15–stem.

4.8.15 Helix 16

This helix was identical to the reference alignment and had similar secondary structure for all species. Both stems were identical except for a point insertion in the internal bulges in *Callopora lineata* and *Celleporella hastata*. For an example of this helix see a main secondary structure image with *Bugula turbinata* on Figure 38.

4.8.16 Helix 17

Although several substitutions were present in both stems of this helix it corresponded well to the proposed secondary structure in the reference alignment.

4.8.17 Helix 18

All species had identical 5′–3′ and 3′–5′ stems with the exception of *Tubulipora liliacea* which had a substitution: U:G pair instead of U:A. *Flustrellidra hispida* had an A:A pair, which does not pair according to conventional base pairing.

4.8.18 Helix 19

All species had identical secondary structure; both stems had compensatory mutations (species: *Tubulipora liliacea*, *Scruparia chelata*, *Crisia denticulata* and *Crisia eburnea*), which did not affect the secondary structure.

4.8.19 Helices 20–21

Both helices (part of the pseudoknot, see above) were identical for all species.

4.8.20 Helix 22

This short helix was identical in all species with some compensatory substitutions present.

4.8.21 Helix 23

All species have identical nucleotide sequences, identical to those of the reference alignment.

4.8.22 Helix E23_12

This helix had considerable variation in the terminal loop area, however the secondary structure was consistent throughout the set of sequences.

4.8.23 Helix E23_1

The sequences analysed in the present study did not match the reference sequence alignment at all. All folding and alignment was done based on the model suggested by the Mfold thermodynamic algorithm. There was a good structural agreement within sequences produced in this work. The following species were clearly different from the rest: *Tubulipora liliacea*, *Scruparia chelata*, *Crisia aculeata*, *Filicrisia geniculata*

and *Crisia eburnea.* Because of inconsistency with the reference sequences and the fact that parts of the helix were separated by another helix it was not possible to infer definitive secondary structure using Mfold. Any attempts to constrain the folding as per the alignment matrix resulted in Mfold bringing up an error. Consequently, the sequences were folded using Mfold for thermodynamic folding and the reference alignment as scaffolding for the potential helix borders. As an example *Bugula turbinata* is given (Figure 38) for this region.

4.8.24 Helix E23_2

The stem of this helix was joined together with the previous one and examples of secondary structure are given above (under E23_1 helix). This helix in some other taxonomic groups also adjoins the E23_3 helix, which was not observed in sequences obtained here.

Restricting the folding algorithm to the site positions similar to those in the published alignment caused Mfold program to abort the calculation and return a run error. Therefore, the alignment was based on the conserved motifs as produced by ClustalX, reference alignment and the results of thermodynamic folding using Mfold whenever possible.

4.8.25 Helix E23_4

This helix did not match the sequences of the reference alignment and as with the previous helix was predicted using Mfold. The border between E23_4 and E23_7 helices becomes arbitrary since these two helices are adjacent to each other and when predicted using Mfold there appears to be no transition from one helix to another. When possible the reference alignment was used to aid in this decision. This helix was folded together with the following helix E23_7.

4.8.26 Helix E23_7

None of my sequences (apart from that of *Scruparia chelata*) align against the reference sequences. The secondary structure of this helix appeared similar in all of my sequences. No alignment adjustment was made as this would not improve the actual alignment in any way. Therefore ClustalW alignment was used with a secondary structure mask on top for later references. The terminal loop of this helix was of similar length among all sequences and consisted of six nucleotides. For an

example of folding see this helix on the main image of secondary structure *Bugula turbinata* (Figure 38).

4.8.27 Helices E23_8; E23_9; E23_10; E23_11; E23_12; E23_13

These helices were very small (2–4 nucleotides each stem) and well conserved among my species. Because of the pseudoknot (described in the V4 area discussion above) which is formed by these helices it was not possible to use Mfold server to predict secondary structure of this area. However my sequences matched well those of the reference alignment. The correctness of these helices was reconfirmed when *Bugula turbinata* sequence was folded into secondary structure.

4.8.28 Helix E23_14

This helix showed great uniformity among all sequences, with differences only in bulges. It is also part of the second pseudoknot described for this region and therefore very difficult to estimate using Mfold algorithm. Manual alignment with the aid of Mfold showed many similarities in the secondary structure of this helix, especially in the 5′–3′ stem.

4.8.29 Helix 24

This helix was identical to the reference alignment and among my sequences. The secondary structure was preserved as per the reference model.

4.8.30 Helix 25

This helix was uniform among all species and the reference alignment, except for *Callopora lineata*, which had an insertion of three extra nucleotides in the internal bulge, not present in the other species of *Callopora*, see Figure 37 for details.

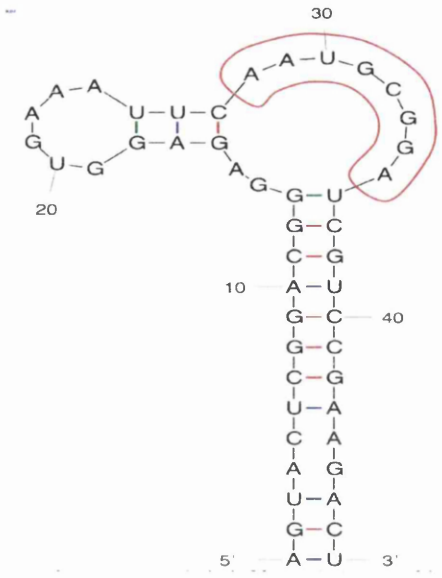


Figure 37 Helix 25 of Callopora lineata. Extra nucleotides inserted in the internal loop are shown by the red line.

4.8.31 Helix 26

This helix was conserved among all my sequences and the reference alignment.

4.8.32 Helix 27

This helix was conserved among all sequences. Some substitutions were present in the internal bulge. Thermodynamic folding with Mfold showed two types of secondary structure with one identical to that of the reference alignment (present in *Tubulipora liliacea*, *Scruparia chelata*, *Flustrellidra hispida*) and the other (present in the rest of the sequences obtained here) which had a slightly shorter internal bulge.

4.8.33 Helix 28

All sequences were identical with some compensatory substitutions and one non–typical base pairing C:U (in *Bugula fulva*).

4.8.34 Helix 29

The 29–stem of this helix matched well the reference alignment. It was also conserved among most of the sequences. The main differences were in the internal bulges which connected 29, 29′, and 30–stems. This helix has to be viewed in connection with helix 28 and 30, when done this way it fits the proposed secondary structure and allows easier folding using Mfold.

4.8.35 Helix 30

The sequences of this helix were identical for all species.

4.8.36 Helix 31

The secondary structure of this helix was conserved among all species (including the reference alignment). Some species, *Tubulipora liliacea, Crisia aculeata, Filicrisia geniculata, Crisia denticulata* and *Crisia eburnea,* showed compensatory substitutions.

4.8.37 Helix 32

This helix was identical for all sequences, with the exception of an insertion in *Bowerbankia imbricata* in the 32–stem.

4.8.38 Helix 33

All sequences were identical between species obtained here and the reference alignment. *Celleporina hassallii* had an insertion in the 5′–3′ stem, which did not alter the secondary structure.

4.8.39 Helix 34

Sequences were identical for all species, except for *Flustrellidra hispida,* which has a single substitution in the internal bulge.

4.8.40 Helix 35

This helix was identical for all species.

4.8.41 Helix 36

This helix is identical for all species except for *Scruparia chelata* (single substitution in the internal bulge of the 5′–3′ part) and *Phaeostachys spinifera* (insertion in the forward stem).

4.8.42 Helix 37

There were several compensatory substitutions which did not affect the secondary structure. Some species (*Crisia aculeata, Crisia denticulata, Crisia eburnea, Filicrisia geniculata, Scruparia chelata, Tubulipora liliacea*) could be folded as per the reference alignment using Mfold. However, other sequences could not be folded

using Mfold constraints. The second type of folding is represented by *Walkeria uva*. This secondary structure was typical for all other species that did not fold as per the reference alignment.

4.8.43 Helix 38

This helix was identical for all species; some differences were present in the internal bulge, which did not affect the secondary structure.

4.8.44 Helices 39–42

These helices are very short and identical for all species and for the reference alignment.

4.8.45 Helix 43

The sequences obtained here were similar to the sequences of the reference alignment at the proximal part of the stems. The distal part together with the terminal loop showed considerable expansion – contraction for most sequences. This helix is part of the V7 region and is expected to have much variation between species. The folding of this region was guided by the reference alignment in the proximal parts of the stems and by thermodynamic folding using Mfold for the proximal stems and the external loop. Secondary structure showed some similarity among the species: terminal hairpin showed expansion – contraction and varied between 4 to 8 nucleotides; also one or two internal bulges were present in all sequences. For an example of this helix folding see complete secondary structure diagram of *Bugula turbinata* (Figure 38).

4.8.46 Helix 44

This helix had the same secondary structure for all species. Some species *e.g. Microporella ciliata* and *Scruparia chelata* had some compensatory substitutions. The terminal loop was of the same length for all species (5 nucleotides) but varied in the nucleotide composition.

4.8.47 Helix 45

This helix had similar secondary structure between all species, some variation was present due to the compensatory substitutions, which has not affected the secondary structure. This helix formed a connection with helix 46 and was folded together with this helix. Constraining Mfold to match the reference secondary structure was not

possible, hence the above image shows a slightly larger internal bulge than that proposed for the secondary structure.

4.8.48 Helix 46

All species showed identical secondary structure, the variation was only in the nucleotide composition where mutations were all compensatory. *Flustrellidra hispida*, when folded directly using Mfold, had a slightly larger hairpin. However, if C:A pairing is to be accepted as valid, then this species would have exactly the same secondary structure as the other species.

4.8.49 Helices 47–48

Sequences of most species are identical to that shown for the main structure of *Bugula turbinata*. The secondary structures of these helices were identical in all species, however some compensatory mutations were present in the stems and terminal loops.

4.8.50 Helices 49 – 50

The last two helices correspond to the variable region V9. These sequences as expected were rather different among species. The problem of alignment was further complicated by the fact that not all sequences were of identical length due to the sequencing and as a result some sequences extended for another 300 basepairs beyond the helix 50 (*e.g. Escharella immersa*). Others on the contrary did not have helix 50 present at all due to incomplete sequencing and finished around the position of 49′–stem of the helix 49.

The sequences of the helix 49 showed strong conservation at the proximal end of the helix, the terminal loop showed much variability. The sequences of this helix were aligned where possible against the reference alignment and to each other first using ClustalX and then manually adjusted based on the conserved motifs. The internal bulges were drawn based on the Mfold thermodynamic folding.

The 50–stem was partially recovered in some species, where present it was conserved. The 50′–stem was present (thus completing helix 50) only in three species (*Microporella ciliata*, *Bugula plumosa* and *Bugula turbinata*). The sequences of the remaining species had this part of the helix missing or if some nucleotides were present they were scrambled and would not align to each other or to the secondary structure reference alignment and were clearly the artefacts of the sequencing, this

often happened at the end of the sequence files received from the sequencing service. Consequently, this helix was excluded from alignment.

4.9 *Secondary structure model of Bugula* turbinata

The secondary structure model of Bryozoa was built based on the helices defined in the previous section. Unfortunately, out of the whole alignment presented in this work only *Bugula turbinata* had a complete sequence. This species had the only sequence which spanned from the 1-stem of helix 1 to the 50′-stem of helix 50, thus covering the entire 18S rRNA gene, and therefore the model is presented only for this species.

The most difficult and time consuming exercise for the secondary structure prediction is building the correct alignment and establishing exact position of the helices, especially for those helices which are part of the pseudoknots.

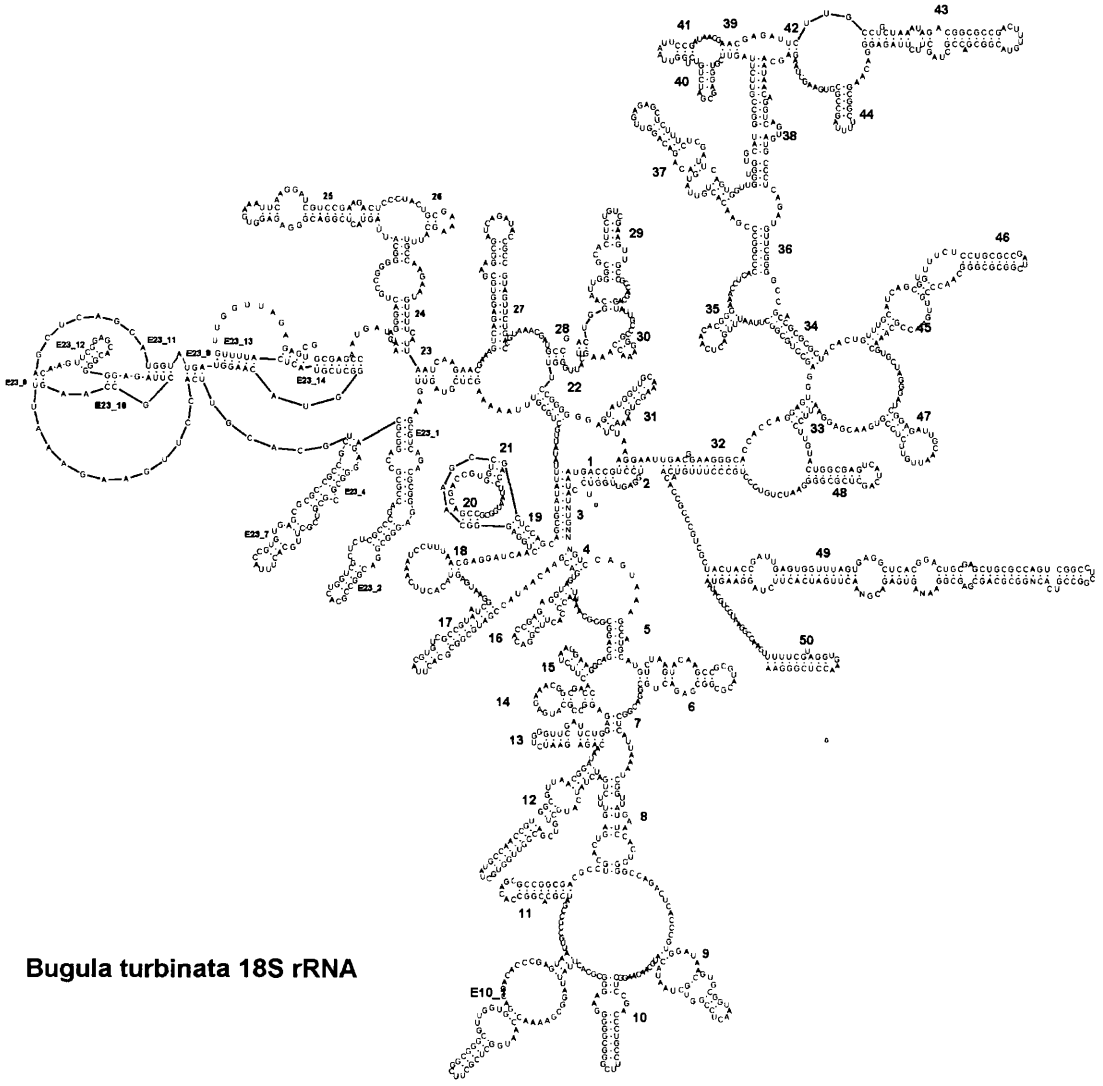


Figure 38 Secondary structure model of *Bugula turbinata* 18S rRNA. The helix numbering as per Van de Peer (2000), with the specific numbering of helices of E23_1 to E23_14 as per Wuyts *et al.* (2000).

Once the correct secondary structure was identified for all sequences in the alignment the *Bugula turbinata* sequence was extracted from the alignment file and used for drawing the secondary structure model for the species. Whilst programs like Mfold, Vienna RNA structure (Hofacker 2003) and GeneBee (Brodsky *et al.* 1995) can predict and draw local foldings of the rRNA they are not meant to predict correctly the secondary structure of the entire rRNA (Zuker, personal communication), partly because neither of these programs can predict tertiary interactions—pseudoknots. For instance, helix 19 region of rRNA can be folded using Mfold, but the resulting structure lacks a pseudoknot, which in fact splits helix 19 into helices 19, 20, 21 (see Figure 38 for the helices 19,20,21). Unfortunately, some authors (Goffredi *et al.* 2006) disregard that and still attempt to fold the entire 18S rRNA gene.

In this work, the file containing *Bugula turbinata* sequence with secondary structure DCSE characters was imported into RNAViz[26] software (de Rijk 2003). This program although it cannot predict secondary structure can draw it (including the pseudoknots) if supplied with correct data. In order to speed up drawing RNAViz uses a so called "skeletal file", which is essentially an *a priori* folded rRNA model stored in an RNAViz recognisable file format. For *Bugula turbinata* a skeletal file of *Saccharomyces cerevisiae* SSU was used. This file is supplied with the software. After the secondary structure model was drawn using the skeleton file and DCSE file containing target sequence the image required some adjustment. The pseudoknot structure in the helices 19-21 was automatically recognised, however the pseudoknots of the variable region V4 were not and had to be manually re-arranged according to the alignment.

The final image of 18S rRNA secondary structure model of *Bugula turbinata* is given on Figure 38. This is the first image of bryozoan 18S rRNA. The file created using RNAViz can be used as a skeleton file for rapid folding of other Bryozoan 18S rRNA, and if requested it could be supplied by the author.

4.10 Alcyonidium specific issues

As can be seen from the list of the sequences acquired in this work (see Chapter 3) three sequences of species of *Alcyonidium* were obtained; however, these

[26] RNAViz can be freely downloaded from the Internet through sourceforge.net werbsite.

were not included in the structural RNA alignment due to their strong differences from the rest of the sequences.

Originally, the sequences obtained here were aligned using ClustalX for similarity evaluation and comparison reasons; also, the process of structural alignment (described above) required this procedure. Unfortunately three sequences of *Alcyonidium* — *A.gelatinosum*, *A.hirsutum* and *A.polyoum* — differed considerably from the rest of the Bryozoan and non-bryozoan 18S rRNA sequences. Initially the length difference was noted – see Table 6. Average sequence length for *Alcyonidium* sequences was 2168 nucleotides versus 1797 nucleotides for the remainder of the sequences examined in this work.

Table 6 Nucleotide composition and sequence length of the sequences used in this study. Unusually high values of *Alcyonidium* sequences marked with an asterisk.

Species	T(U)	C	A	G	Total nucleotides
Neocrania anomala	25.5	22.1	25.6	26.8	1753
Neocrania huttoni	25.7	21.9	25.5	26.9	1753
Scruparia chelata	25.2	22.5	24.8	27.4	1804
Barentsia hildegardae	25.4	21.7	25.8	27.2	1759
Pedicellina cernua	25.4	21.9	25.7	27	1720
Barentsia benedeni	25.6	22.3	25	27	1734
Plumatella repens	24.3	22.9	25.1	27.7	1755
Crisia aculeata	22.6	24.7	23.2	29.5	1755
Crisia eburnea	23.4	24.1	23.2	29.3	1817
Filicrisia geniculata	23.9	24	23.5	28.7	1799
Crisia denticulata	23.8	23.9	23.4	28.9	1746
Tubulipora liliacea	24.4	23.5	23.8	28.3	1755
Microporella ciliata	21.5	25.4	23.8	29.2	1867
Escharella immersa	21.3	25.4	23.7	29.6	1868
Bugula fulva	21.4	25.9	23.1	29.7	1858
Haplopoma graniferum	21.3	25.8	23.5	29.5	1828
Schizomavella linearis	22.9	23.9	23.9	29.3	1776
Escharoides coccinea	21.1	25.9	22.8	30.1	1766
Phaeostachys spinifera	20.8	25.7	23.4	30.1	1787
Callopora lineata	21	26.2	22.8	30	1815
Callopora rylandi	21.3	25.8	22.8	30.1	1793

Species	T(U)	C	A	G	Total nucleotides
Callopora dumerilii	21.1	25.5	23.4	30.1	1862
Bugula turbinata	21.4	25.6	23.1	29.8	1820
Cribrilina cryptooecium	21.4	25.6	23.3	29.7	1832
Celleporina hassallii	20.9	26	23.9	29.2	1730
Umbonula littoralis	20.9	26	23.7	29.4	1853
Bicellariella ciliata	20.5	26.2	22.7	30.6	1822
Walkeria uva	20.4	27	22.3	30.3	1859
Bowerbankia citrina	20.1	26.9	22.5	30.5	1769
Bowerbankia gracilis	19.9	27.2	22.2	30.7	1815
Bowerbankia imbricata	19.6	27.1	22.1	31.2	1823
Flustrellidra hispida	20	27.3	22.2	30.4	1799
Alcyonidium polyoum	19.1	27.5	21.6	31.8	2125*
Alcyonidium hirsutum	19.7	26.9	21.5	31.9	2216*
Alcyonidium gelatinosum	19.6	27	22.3	31.1	2164*

Kjer (2004) mentions that using secondary structure alignment can be a good indicator to detect chimeras or any errors in the sequences. Indeed, strong differences between *Alcyonidium* sequences and the rest of the structurally aligned sequences as well as their length raised a possibility that these sequences contained errors and inserts. The problems of chimeras specifically due to cloning were excluded as the same sequences were obtained for *Alcyonidium* for both cloning and direct sequencing (see Chapter 3 - Molecular methods). Also, all three *Alcyonidium* sequences were processed separately. That is, they were collected from different locations, DNA was extracted during different sessions, and sequenced separately, which lowered the chances of cross contamination. Despite the above process these three sequences, when aligned against each other, had very high similarity (Figure 39). In addition as can be seen from the alignment below, the three sequences are almost the same, which would make the possibility of cloning chimeras even less likely.

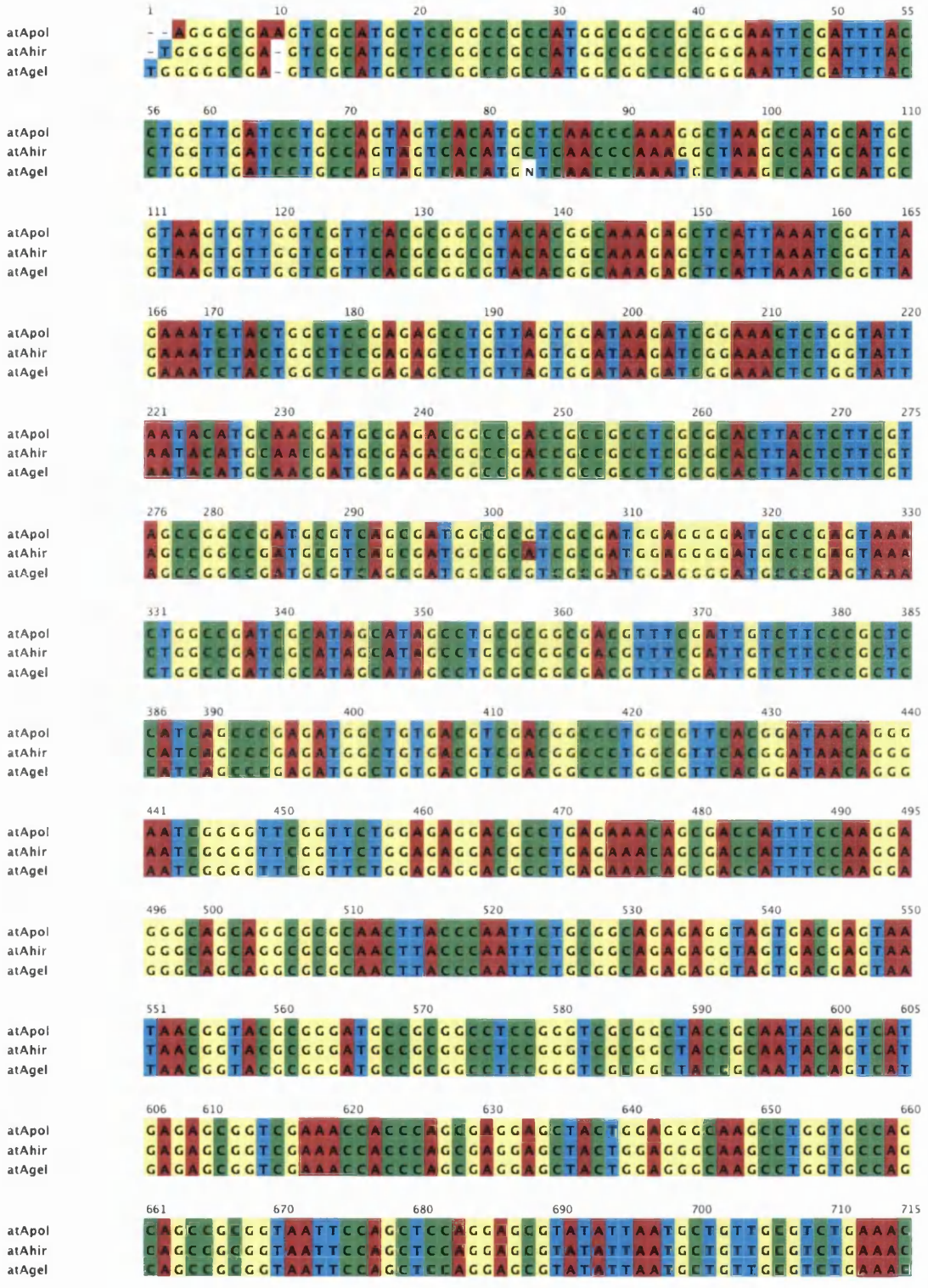
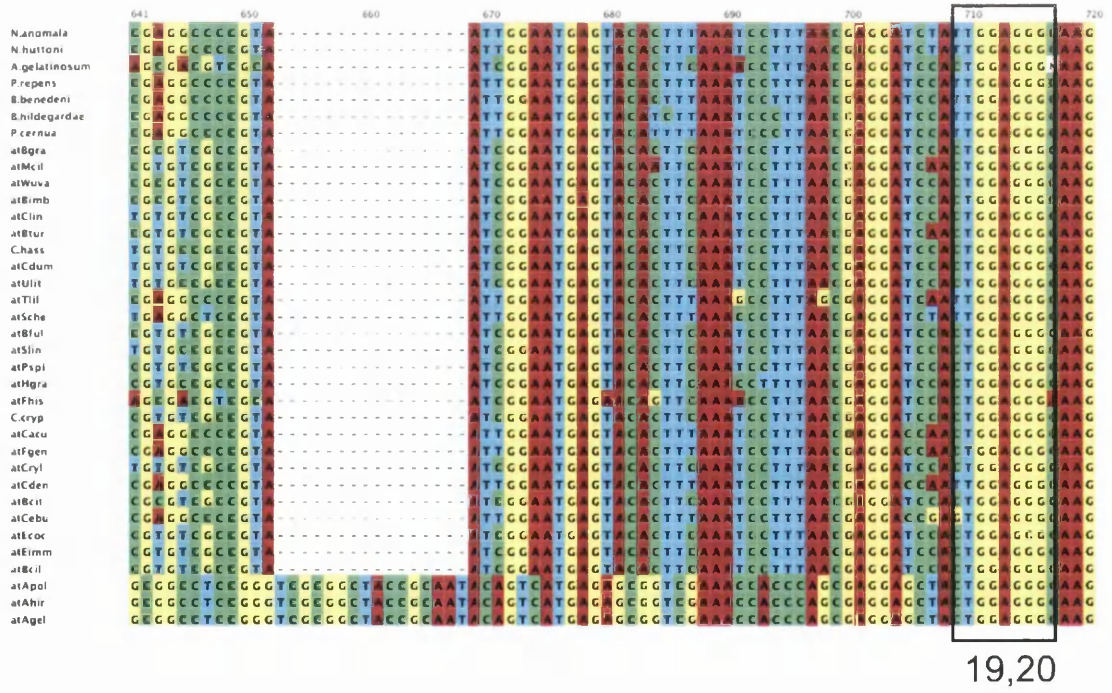


Figure 39 Sequence alignment of three *Alcyonidium* sequences obtained in this work. *Alcyonidium polyoum* (atApol) *Alcyonidium hirsutum* (atAhir) and *Alcyonidium gelatinosum* (atAgel) are aligned using ClustalX, using default parameter settings. Only the first 715 nucleotides shown as an example.

It appeared therefore that the only plausible explanation was that the primers which were developed specifically for *Alcyonidium* species "picked up" some kind of contaminant when DNA extractions were made. Because of the high similarity of these sequences, it is possible that a contaminant DNA with which the primers reacted was of an organism of a symbiotic nature. Certainly morphological similarities of *Alcyonidium* species may suggest that their colonies could have a common symbiotic organism should such symbionts exist. Also, as discussed in the Chapter 3, *Alcyonidium* species appeared to require a slightly different set of primers from other bryozoan species (in particular from other Ctenostomata).

The alignment produced by MAFFT also emphasised the problem – this software opened several large gaps. The conserved motifs, which aligned well against other bryozoan sequences obtained here, were exclusively located in the regions corresponding to the stems of RNA. For instance, Figure 40 shows three segments of the multiple alignment of all bryozoan sequences including *Alcyonidium* species.



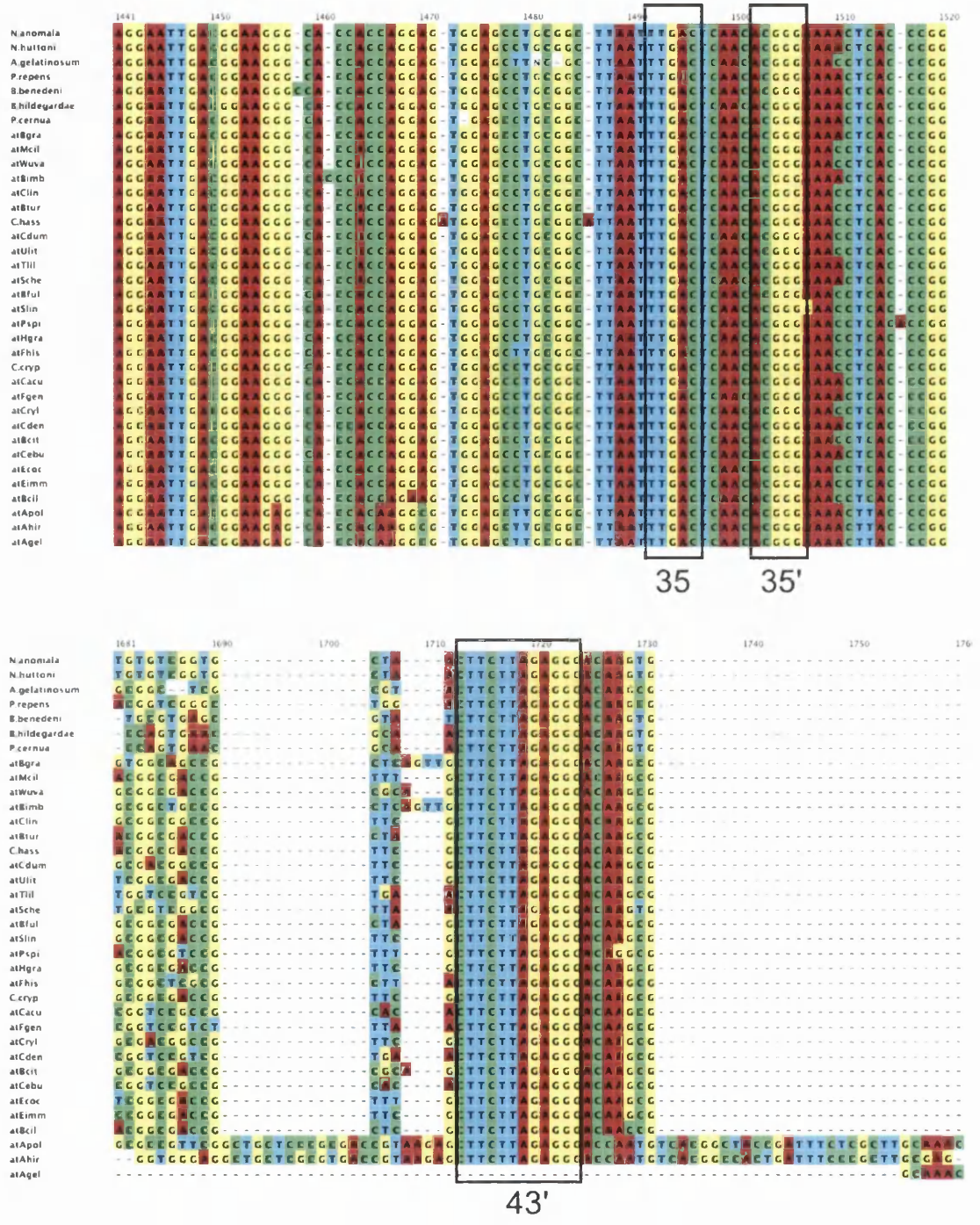


Figure 40 Segments of computer alignment (MAFFT) of *Alcyonidium* sequences (marked as atApol, atAhir, atAgel – the last three ones) obtained in this work against other sequences. Black squares indicate location of corresponding helices of 18S rRNA.

Examination of the similarities of sequences by using BLASTn search (from NCBI) gave unusual results – see Figure 41 for example. BLAST results did pick up some similarity to bryozoan sequences, but they corresponded to a very short transcript of the sequence. For instance, the *Alcyonidium gelatinosum* sequence obtained here had

its first hit against the *Membranipora grandicella* (accession no. AF499742) sequence with the score of 217, that is a very low score for a 1800 nucleotides sequence.

Accession	Description	Max score	Total score	Query coverage	E value	Max ident
AY172989.1	Uncultured metazoan 18S ribosomal RNA gene, partial sequence	303	513	14%	1e-78	96%
AF499742.1	Membranipora grandicella 18S ribosomal RNA gene, complete sequence	217	217	8%	2e-52	88%
AF119082.1	Caberea boryi 18S ribosomal RNA gene, complete sequence	217	418	17%	2e-52	88%
AY210433.1	Bugula turrita 18S ribosomal RNA gene, partial sequence	217	461	27%	2e-52	88%
X91977.1	S.ventrolineatus 18S ribosomal RNA	215	345	20%	7e-52	88%
AY582119.1	Pseudechiniscus islandicus 18S ribosomal RNA gene, partial sequence	211	407	16%	9e-51	92%
AY830844.1	Arabella semimaculata 18S ribosomal RNA gene, partial sequence	211	417	23%	9e-51	95%
AY525624.1	Arabella iricolor 18S ribosomal RNA gene, partial sequence	211	409	20%	9e-51	88%
EU164974.1	Scutellospora castanea 18S ribosomal RNA gene, partial sequence	206	206	8%	4e-49	87%
EU164968.1	Glomus deserticola 18S ribosomal RNA gene, partial sequence	206	206	8%	4e-49	87%
DQ839601.1	Macrobiotus sapiens 18S ribosomal RNA gene, complete sequence	206	391	25%	4e-49	87%
AJ852598.1	Glomus etunicatum 18S rRNA gene, isolate UFPE06	206	394	23%	4e-49	95%
EF136915.1	Glomus sp. NBR PP1 clone PP1-11 18S small subunit ribosomal RNA ge	206	389	23%	4e-49	95%
EF136914.1	Glomus sp. NBR PP1 clone PP1-10 18S small subunit ribosomal RNA ge	206	318	21%	4e-49	87%
EF136912.1	Glomus sp. NBR PP1 clone PP1-8b 18S small subunit ribosomal RNA ge	206	394	23%	4e-49	95%
EF136911.1	Glomus sp. NBR PP1 clone PP1-8 18S small subunit ribosomal RNA gen	206	394	23%	4e-49	95%
EF136908.1	Glomus sp. PM1.2 clone PM1-2-4 18S small subunit ribosomal RNA gen	206	389	23%	4e-49	93%
EF136907.1	Glomus sp. PM1.2 clone PM1-2-3 18S small subunit ribosomal RNA gen	206	389	23%	4e-49	95%
EF136906.1	Glomus sp. PM1.2 clone PM1-2-2 18S small subunit ribosomal RNA gen	206	394	23%	4e-49	95%
EF136903.1	Glomus sp. NBR8.7 clone NBR8-7-27 18S small subunit ribosomal RNA	206	389	23%	4e-49	95%
EF136902.1	Glomus sp. NBR8.7 clone NBR8-7-25 18S small subunit ribosomal RNA	206	394	23%	4e-49	95%
EF136901.1	Glomus sp. NBR8.7 clone NBR8-7-5 18S small subunit ribosomal RNA ge	206	394	23%	4e-49	95%
EF136896.1	Glomus sp. NBR3.1 clone NBR3-1-43 18S small subunit ribosomal RNA	206	396	23%	4e-49	95%
EF136895.1	Glomus sp. NBR3.1 clone NBR3-1-42 18S small subunit ribosomal RNA	206	389	23%	4e-49	95%
EF033121.1	Glomus sp. PM1.2 18S small subunit ribosomal RNA gene, complete sec	206	394	23%	4e-49	95%
EF033120.1	Glomus sp. NBR3.1 18S small subunit ribosomal RNA gene, complete se	206	394	23%	4e-49	95%
DQ085262.1	Uncultured Glomus clone JPC091 JP7 sequence type 18S ribosomal RNA	206	323	21%	4e-49	87%
DQ085261.1	Uncultured Glomus clone JPC090 JP7 sequence type 18S ribosomal RNA	206	323	21%	4e-49	87%
DQ085260.1	Uncultured Glomus clone JPC089 JP7 sequence type 18S ribosomal RNA	206	323	21%	4e-49	87%

Figure 41 Results of BLASTn search for the sequence of *Alcyonidium gelatinosum*.

Notably, all three *Alcyonidium* sequences when BLASTn searched gave the highest similarity to the same particular sequence—*Uncultured metazoan* sequence (accession no. AY172989). This sequence is deposited on the NCBI database with the following description: "Environmental sample of uncultured metazoan obtained from suspension feeding invertebrate such as Bryozoa" A neighbour joining tree, containing all the sequences produced when a BLASTn search was performed on the *Alcyonidium gelatinosum* sequence, is given on Figure 42. In addition to the *Alcyonidium gelatinosum* (atAgel) sequence, sequences of *Alcyonidium hirsutum* (atAhir) and *Alcyonidium polyoum* (atApol) were added to this tree for comparison. This tree clearly shows that *Alcyonidium* sequences (clade marked on the tree in red point) are much more similar to the *Uncultured metazoan* sequence (in the same clade), and equally well distanced from the rest of the sequences, which included representatives of Metazoa (*e.g. Anodonta alba*, *Astarte sulcata*), Plantae (*e.g. Spermatozopsis similes*), algae (*e.g. Rhodella* sp.), and many fungi (*e.g. Glomus* sp.; Uncultured soil fungi). Interestingly, another separate clade can be seen on this tree which contains five species of Bryozoa: *Bugula turrita* (accession no. AY210433), *Caberea boryi* (accession no. AF119082) and *Membranipora grandicella* (accession no. AF499742), *Smittoidea spinigera* (accession no. AF499746) and *Bugula stolonifera* (accession no. AF499745). These sequences were reviewed earlier in the Chapter 3 when the development of oligonucleotide primers was discussed. The first

two are considered to be valid sequences, however the validity of the last three, submitted by the same group of authors Hao *et al.* (2002) cannot be verified. The presence of these sequences in the BLASTn search results most likely corresponds to the matching conserved regions of the 18S rRNA between *Alcyonidium* sequences and the rest of the Bryozoa (as can was shown above). However, the fact that only these five bryozoan sequences were picked up by BLASTn search is difficult to explain and suggests that further investigation is needed.

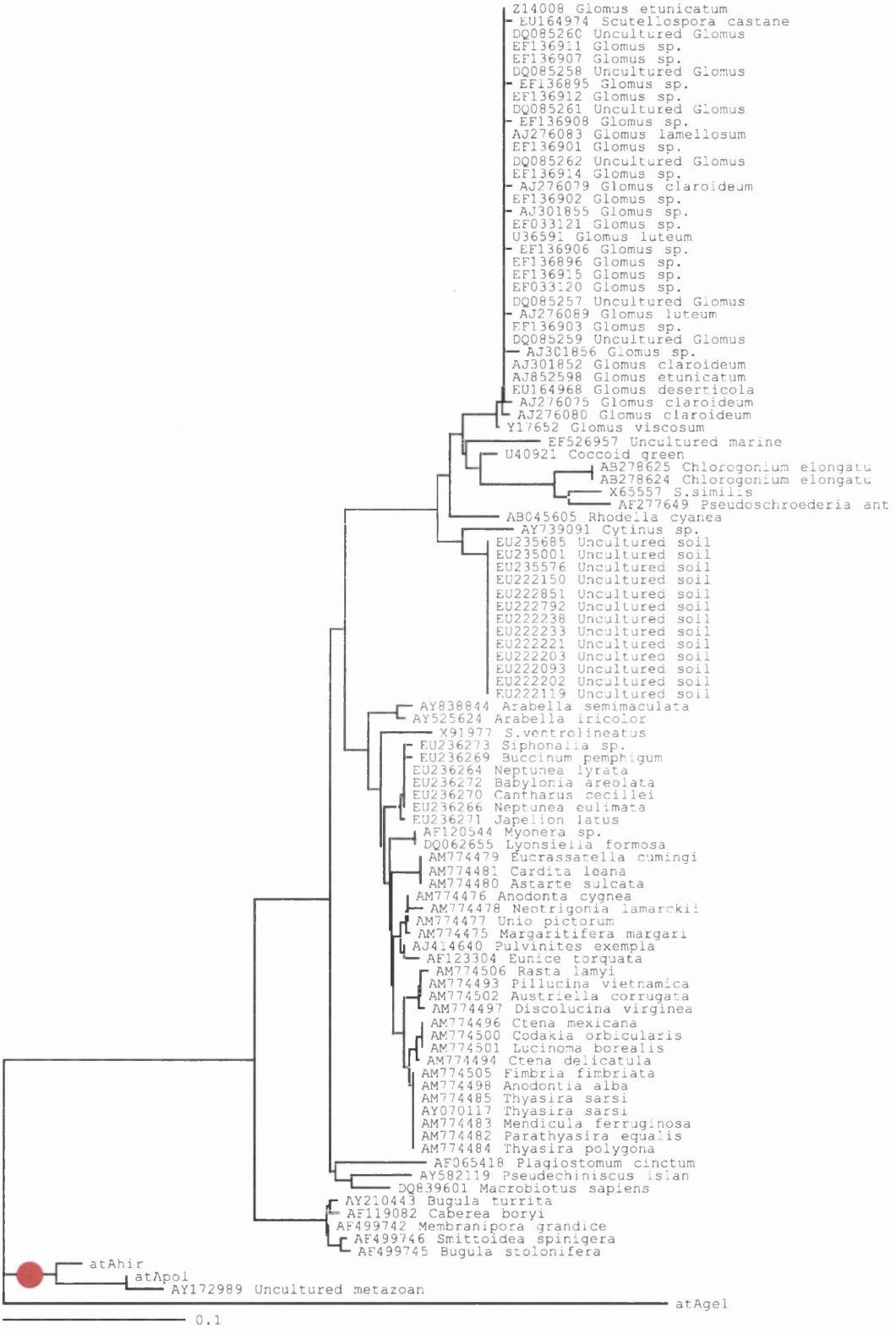


Figure 42 A NJ tree (K2P model) of BLASTn search results and *Alcyonidium* sequences obtained in this study. A red point shows a clade in which all three *Alcyonidium* sequences cluster together with the sequence of "Uncultured metazoan". NCBI sequences are preceded by their corresponding accession no. Bar at the bottom of the tree shows substitutions per site.

The above finding gives support to the hypothesis that the *Alcyonidium* sequences contain some kind of chimeric sequence from a contaminant picked up during the DNA extraction from the larvae. The first possibility is that the contamination could have occurred when freely swimming larvae were collected by a micropipette to be placed into the digestion solution tubes and thus contained in the surrounding seawater. Alternatively, it is possible that the larvae themselves had the contaminant attached to their surface. The former explanation would make the method of larval DNA extraction less robust, but at the same time this cannot explain why extractions of larvae of other species of Bryozoa were not contaminated in the same way. Therefore, it is more likely that the contaminant had some kind of affiliation to the specific *Alcyonidium* species, and most likely, their larvae.

The question remains open about the causes of these sequence anomalies and requires further investigation.

5 PHYLOGENETIC ANALYSIS OF BRYOZOAN 18S rRNA SEQUENCES

5.1 Secondary structure based alignment advantages vs. conserved motifs alignment

There has been a long history of discussion in the literature about the importance of using secondary structure for rRNA sequence alignment. In addition, many arguments have been stated that secondary structure can improve phylogenetic reconstruction in general. Below some of the advantages and criticism of this approach are discussed. In general, secondary structure being highly constrained and relatively universal can greatly aid in alignment of variable segments of rRNA (Gutell *et al.* 1994; Kjer 1995; Woese and Gutell 1987).

An improvement of the alignment using secondary structure was achieved in many phylogenetic studies (Wilmotte *et al.* 1993; Winnepenninckx and Backeljau 1996; Winnepenninckx *et al.* 1994; Van de Peer *et al.* 2000; Kjer 1995; Telford *et al.* 2005; Kjer *et al.*1994; Kjer 1995; Kjer 2004; Ouvard *et al.* 2000; Xia *et al.* 2003; Morrison and Ellis 1997; Hudelot *et al.* 2003; Xia *et al.* 2003; Page 2000; Gillespie *et al.* 2005; Passamaneck and Halanych 2006).

Some authors (Kjer 1995; Winnepenninckx and Backeljau 1996; Hudelot *et al.* 2003; Xia *et al.* 2003) emphasised the importance of an accurate alignment of 18S for successfully retaining the homologous characters within the aligned sequences. Also some researchers found that secondary structure based alignment improved the analysis findings and therefore strongly advocate the use of these methods (Winnepenninckx and Backeljau 1996; Xia *et al.* 2003).

An example of such improvement can be seen in the work of Xia *et al.* (2003) who analysed anecdotal evidence of grouping of birds and mammals based on 18S rRNA analysis for nearly a decade of published works. This paradox of abnormal grouping was solved by using a different alignment method that took into account secondary structure of 18S rRNA, as well as reconsidered which data are discarded during the analysis. Some authors discard those regions of 18S rRNA gene which are difficult to align (usually hypervariable regions). The study conducted by Xia *et al.*

(2003) is definitely a very important example of how using incorrect methods can influence our understanding of phylogenies.

In many studies in which Bryozoa were used as a taxonomic group, secondary structure was not used during an alignment, and some recent publications concerning Bryozoa and Lophophorates (Giribert *et al.* 2000; Dick *et al.* 2000; Halanych 1995; Okuyama *et al.* 2006; Zrzavy *et al.* 1998; Hao *et al.* 2005) seemed to fail to take secondary structure into consideration as well. Certainly a recent study of bryozoan phylogeny which used rRNA data (Dick *et al.* 2000) created more questions than answers and challenged our common knowledge of bryozoan phylogeny (see discussion section of this chapter).

The advantages of accounting for secondary structure are evident, but the alignment of these highly variable regions – loops, bulges and stems – can be very difficult. An obvious choice would be to use software, which could aid in the alignment as is done with coding genes. But because of the difficulty of assessing homologous sites in rRNA it is much more difficult to automate the alignment of rRNA. Most of the alignment computer programs rely on a gap penalty, which is assigned by the investigator or, as in many cases, left to be the default settings. And this is when certain regions of 18S rRNA gene can clearly mislead the alignment programs.

Most importantly for rRNA sequence is that each region within the gene would have to have different gap penalties as loops, bulges and stems vary in size and have to be treated in their own way. Any alignment program, which looks into similarities between different sequences, will also fail to take into account homologous sites within the same sequences and thus secondary structure. In fact when several commonly used alignment programs were evaluated the results of their alignment were only 25%-34% similar to the alignment based on secondary structure (Hickson *et al.* 2000).

Further, different lengths of sequences in some regions can also confuse the alignment program. This can be clearly seen on any loop-stem border: some species have larger loops, others shorter (*i.e.* fewer nucleotides). These regions are sometimes called regions of expansion and contraction. This happens when there are segments of the gene that have large inserts. When other species are brought into the alignment that have shorter sequences, the software can move some segments of the shorter sequences freely to match those of other species. Unfortunately, this rarely refers to

the homologous sites (personal findings; see also Hickson *et al.* 2000). Thus, a manual alignment of difficult to align regions (especially regions of expansion and contraction and hypervariable regions) is required to make a confident overall alignment.

Nevertheless there are some software packages which attempt to automate secondary structure alignment such as POY (Wheeler *et al.* 2003; Varón *et al.* 2007) which utilises a direct optimisation method using a dynamic homology algorithm during the phylogenetic reconstruction without the use of an alignment. But this is not an alignment tool *per se*. Another well documented software package is PRAGA (Notredame *et al.* 1997) which uses a genetic algorithm for secondary structure alignment. However, this software package appears to be limited by the total number of nucleotides which can be processed. The total length of any submitted alignment file is limited to 2kb nucleotides in any combination—such as four sequences of length 500 nucleotides for instance or two of 1kb *etc.* Therefore it cannot be used with several nearly 2kb 18S rRNA sequences (which add up to nearly 60 kb nucleotides) presented here.

The other interesting software project worth attention is ARB (Ludwig *et al.* 2004). This software has an automated secondary structure aligner integrated into the package. Yet, there are some limitations in this software as well. Primarily, it was created for 16S bacterial rRNA sequences; the main strength of its integrated software aligner is that it can incorporate a large number of sequences and thus align mostly different structures. However, the large database of already aligned sequences (incorporating secondary structure information) on which the software relies when the aligner is invoked is built for 16S rRNA sequences and different, especially in the variable regions, from those of 18S rRNA. For instance, 16S does not form pseudoknots in the V4 region, whereas 18S does. The applicability of this software for use with 18S rRNA is weakened by the fact that 18S secondary structure is still under discussion and open to different interpretations (Wuyts *et al.* 2000; Gillespie *et al.* 2005a; Gillespie *et al.* 2005b). Therefore manual alignment is needed for the time being.

5.1.1 Treatment of hypervariable regions

The other part of the issue, which was also mentioned by Xia *et al.* (2003), is which regions to exclude from the alignment. Sometimes so called hypervariable regions

(those usually corresponding to eukaryote specific helices as well as those which include pseudoknots) are excluded from the phylogenetic analysis even if the secondary structure was used for the alignment of the sequences (Cohen *et al.* 1998; Hao *et al.* 2005; also see Xia *et al.* 2003 for more examples).

The exclusion of the hypervariable regions is a "two-sided sword" as on the one hand it simplifies the alignment procedure: hypervariable regions are sometimes impossible to align using software methods. On the other hand the omission of the hypervariable regions can take away crucial informative sites (Kjer 1995). Also, inclusion of these regions increases resolution of the clade support (Hudelot *et al.* 2003).

One of the proposed ways to deal with ambiguously aligned variable regions is a procedure of unequivocally coding of these regions by coding the sequences which have similar ambiguity using ASCII characters for each ambiguous region (Lutzoi *et al.* 2000) but this type of coding was not considered in this work.

Hillis and Dixon (1991) refers to the fact that highly variable regions can produce unreliable results due to homoplasy and difficulty of alignments, and consequently recommends removing these regions from the analysis. The amount of conservation of the different regions can clearly be seen from comparison of the 16S and 18S gene from different regions (Figure 43)

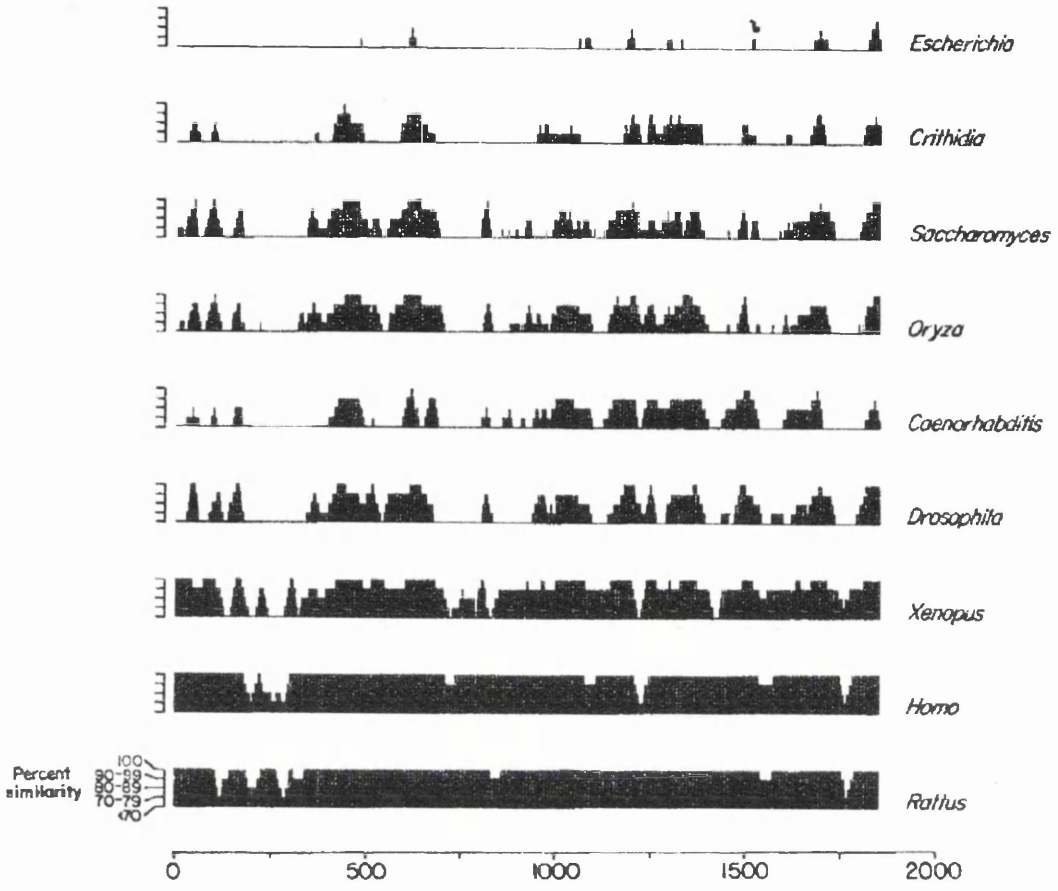


Figure 43 This image shows similarity comparisons of the rRNA genes 16S and 18S between different taxa. The vertical axis shows percentage of the similarity of different species, mapped to *Mus*; the scale on the x-axis shows nucleotide positions on the *Mus* sequence (from Hillis and Dixon 1991: p414).

Also, as much as ten-fold difference in the rate of observed number of nucleotide substitutions between different regions of the 18S rRNA was observed (Abouheif *et al.* 1998). The same rate was found to be correct not only within the entire molecule, but also within secondary structure classes: loops, stems, bulges and therefore preserving different levels of information stored in the molecule. Thus exclusion of highly variable regions, if done, should be done with a certain care in order not to make aligned sequences uninformative.

Kjer (1995) also mentions that a secondary structure alignment is important for more distantly related taxa, because substitution and length variations tend to accumulate with the increase of the divergence time. This presumably affects conserved and highly variable regions at a different rate and hence there is more chance of misalignment if an automated method is used.

5.1.2 Treatment of paired regions

Some controversy exists (Hillis and Dixon 1991) regarding the use of so-called paired regions, i.e. stems which evolve in a constrained way correlated through the hydrogen bonds. These regions may produce less agreeable results with those of the analysis of the unpaired regions (i.e. loops and bulges). On the other hand the same authors give an example of their own study where the opposite effect was observed, in other words paired regions produced "better" phylogeny—more agreeable with current dogma, and morphological data. Hillis and Dixon (1991) concludes that the paired conserved regions must be retained for analysis and it is these regions that contain most of the phylogenetically informative sites.

5.1.3 Structural alignment of Bryozoa

Kjer (1995) postulated that the correct alignment and presentation of data could allow for multiple methods of phylogenetic reconstruction as well as different hypotheses (i.e. models) to be tested at any time. And subsequently it would allow a re-assessment of the results to be performed even if computational methods and hypotheses were to change in the future. The above approach was used in this work: that is, based on secondary structure alignment, different models were evaluated.
The list of all sequences used in the structural alignment is given in Table 7 on page 116. In one case, however, because of the marked difference of *Alcyonidium hirsutum*, *Alcyonidium polyoum*, *Alcyonidium gelatinosum* from the rest of bryozoan sequences obtained in this work, these species were not included in the secondary structure alignment and consequently, excluded from analysis, which used RNA-specific partitioning of the data set and RNA-specific evolutionary models.

Although there are good reasons to believe that the secondary structure should be taken account of, it was decided to sacrifice this principle in the case of *Alcyonidium* spp. sequences (which were not possible to align using secondary structure) in order to use this sequence data for phylogenetic reconstruction. Therefore, computer assisted alignment without a consideration of the secondary structure was used for a separate analysis involving *Alcyonidium* sequences. This in its turn prevented the secondary structure motifs to be recorded and therefore affected the models that could be used.

A detailed treatment of model selection procedure for all alignments used in this work is given below, but here it is sufficient to say that the alignment of the data

set which included the three *Alcyonidium* species was done using a relatively newly developed multiple sequence alignment software—MAFFT (Katoh and Toh 2007; Katoh *et al.*2005; Katoh *et al.* 2002).

Some features of the secondary structure based alignment are subjective—that is some nucleotides could be placed in different positions because loops could make the whole concept of secondary structure alignment useless due to their high variability. Placement of many ambiguous sites in the alignment in this work was actually improved because of the secondary structure refinements.

Additionally, it is quite important to use the most up-to-date secondary structure model available. Winnepenninckx and Backeljau (1996) during examination of phylogenetic data using more up-to-date secondary structure models found different phylogenetic results from those obtained using older models. Thus, different secondary structure models can potentially yield different results and therefore secondary structure based alignments also affect phylogenetic reconstruction. This however merely indicates the fact that any different alignment would affect, to some degree, the outcome of the phylogenetic reconstruction. This in itself is notable as it strongly suggests that alignment plays one of the crucial roles in phylogenetic analysis!

Some other reasons, not affecting phylogenetic reconstruction directly, for using secondary structure alignment were mentioned in literature. For instance, Kjer (2004) argued in favour of secondary structure based alignment as a method of screening for contaminated sequences – sequences containing contaminants or chimeras would not align well and would not follow the general secondary structure models. This has certainly become evident in this work as an attempt was made to align *Alcyonidium* spp. sequences, which were considerably different from the rest of the sequences of Bryozoa obtained here. Secondary structure aided alignment has direct impact upon the models that can be used with the analysis. Careful consideration of the secondary structure during the alignment process allows partitioning of data so that RNA-specific models can be used during the phylogenetic analysis.

5.2 Non structural alignment of sequences

5.3 Selection of software

As mentioned above in this chapter the alignment of some *Alcyonidium* sequences were not possible using secondary structure due to large differences between the reference alignment and the *Alcyonidium* sequences. Therefore a decision was made to use a software-based alignment, which would allow alignment of all sequences based on the conserved motifs. By doing so the valuable information from three *Alcyonidium* species will be preserved and some additional models and methods could be tested simultaneously.

A large number of computer packages have been designed for protein sequences, these programs do not perform as well for the rRNA data though. The effect of parameter settings also affects the outcome of how these programs perform in relation to non-coding sequences. In fact, robustness to the effect of parameter change was suggested as a more important criteria than the program itself (Hickson *et al.* 2000).

In a comparison of several commonly used alignments programs, ClustalW had the highest relative alignment score even with different gap costs parameters (Hickson *et al.* 2000). However, these parameters performed best with small gap cost (both opening and extending the gap). In recently repeated tests of several computer programs for ability to align RNA sequences, software aligner MAFFT (Katoh *et al.* 2002; Katoh *et al.* 2005) utilizing L-INS-I algorithm scored highest in all tests (Wilm *et al.* 2006). MAFFT has also persistently scored best in several tests concerned with protein alignments (Ahola *et al.* 2006; Pei and Grishin 2006; Nuin *et al.* 2006).

Finally, in the tests performed in this study, MAFFT had no difficulty coping with large differences between sequences and presence of inserts in some sequences in the alignment, thus enabling the opening of large gaps where necessary to accommodate inserts in *Alcyonidium* spp. sequences. ClustalW however, could not detect the inserts in the sequences. MAFFT produces sequence alignments based on the iterative refinement method, which allows detection of homologous sequence segments. It uses a staged approach whereby an initial alignment is done using a progressive method and then an iterative refinement of the alignment is performed using fast Fourier transform. The software offers different settings for the alignment

depending on speed and precision. The slowest, but most accurate, method was chosen, in this case – L-INS-i. This procedure is specifically recommended by the software author (see MAFFT website[27]) for the alignment of RNA-like sequences that may require opening of large gaps.

This alignment, which can be supplied upon request, consisted of 37 sequences (Table 7), the same sequences included in the structure specific alignment plus three *Alcyonidium* species: *A. gelatinosum, A. hirsutum, A. polyoum.* The sequence of *Bugula plumosa* was later excluded from the alignment and consequently from the analysis due to the suspected error with the sequence identification. See below the analysis section of this chapter for more details.

Table 7 All species which were used both in the structure based (*) and software-based alignments. Source of sequences: AT - the author; JP - Joanne Porter; number – accession number from NCBI; Species, which were used as references for structural alignment and downloaded from the European ribosomal RNA database.

Classification		Scientific name	Source
Phylum	Brachiopoda		
Class	Craniata		
Order	Craniida		
Family	Craniidae	*Neocrania anomala**	U08328
		*Neocrania huttoni**	U08334
Phylum	Bryozoa		
Class	Gymnolaemata		
Order	Cheilostomata		
Family	Bugulidae	*Bicellariella ciliata**	AT
Family	Bugulidae	*Bugula fulva**	AT
		*Bugula plumosa**	AT
		*Bugula turbinata**	AT
Family	Calloporidae	*Callopora dumerilii**	AT
		*Callopora lineata**	AT
		*Callopora rylandi**	AT
Family	Celleporidae	*Celleporina hassallii**	JP
Family	Cribrilinidae	*Cribrilina cryptooecium**	JP
Family	Escharellidae	*Escharella immersa**	AT
Family	Exochellidae	*Escharoides coccinea**	AT

[27] http://align.bmr.kyushu-u.ac.jp/mafft/software/

Classification		Scientific name	Source
Family	Hippothoidae	*Haplopoma graniferum**	AT
Family	Microporellidae	*Microporella ciliata**	AT
Family	Escharinidae	*Phaeostachys spinifera**	AT
Family	Bitectiporidae	*Schizomavella linearis**	AT
Family	Scruparidae	*Scruparia chelata**	AT
Family	Umbonulidae	*Umbonula littoralis**	AT
Order	Ctenostomata		
Family	Alcyonidiidae	*Alcyonidium gelatinosum*	AT
		Alcyonidium hirsutum	AT
		Alcyonidium polyoum	AT
		*Alcyonidium gelatinosum**	X91403
	Flustrellidridae	*Flustrellidra hispida**	AT
	Vesiculariidae	*Bowerbankia citrina**	AT
		*Bowerbankia gracilis**	AT
		*Bowerbankia imbricata**	AT
	Walkeriidae	*Walkeria uva**	AT
Class	Stenolaemata		
Order	Cyclostomata		
Family	Crisiidae	*Crisia aculeata**	AT
		*Crisia denticulata**	AT
		*Crisia eburnean**	AT
		*Filicrisia geniculata**	AT
	Tubuliporidae	*Tubulipora liliacea**	AT
Class	Phylactolaemata		
Family	Plumatellidae	*Plumatella repens**	U12649
Phylum	Entoprocta		
Order	Pedicellinida		
Family	Pedicellinidae	*Barentsia benedeni**	U36272
		*Barentsia hildegardae**	AJ001734
		*Pedicellina cernua**	U36273

5.4 Bayesian phylogenetics

Bayesian methods are closely related to the ML methods through the use of the likelihood as well as a specific model (Felsenstein 2004; Archibald *et al.* 2003). They differ in the use of priors distribution, of what is being inferred – in this case a tree. In

addition, a Bayesian method arrives at a sample of trees rather than one. Bayesian methods were first proposed to be used in phylogenetics in 1995 by Kass and Raftery (1995) see Huelsenbeck *et al.* (2002) for more details. The only possible way to calculate the posterior probabilities of the tree is the Markov chain Monte Carlo method (MCMC). Bayesian probability is calculated based on the general formula of the Bayes theorem where the inferences of phylogeny are based on the posterior probabilities of phylogenetic trees (Huelsenbeck and Ronquist 2001; Alfaro and Holder 2006):

- $$f(\tau_i \mid X) = \frac{f(X \mid \tau_i) f(\tau_i)}{\sum_{j=1}^{B(s)} f(X \mid \tau_i) f(\tau_i)}$$

where

- $$f(X \mid \tau_i) = \int_{\upsilon} \int_{\Theta} f(X \mid \tau_i, \upsilon, \Theta) f(\upsilon, \Theta) d\upsilon d\Theta$$

In the above formulae, the posterior probability of the *i*-th phylogenetic tree (τ_i) is conditional on all parameters of the aligned DNA sequences (X). The summation is done over all B(s) trees that are possible for s species, where (υ) is combination of branch lengths and (Θ) combination of all substitution parameters.

A good overview of the use of Bayesian methods in phylogenetics is given by Felsenstein (2004), and although the method's description dates back to the 1970s its full adoption in phylogenetics was restricted by the computational power of computers.

A possibility of using MCMC methods to draw samples from the posterior probabilities distribution sped up the use of Bayesian methods in phylogenetics. Further, Metropolis coupled algorithm with a random tree at the beginning and step by step evaluation of the neighbouring trees is embedded in MrBayes program (Huelsenbeck *et al.* 2002; Ronquist and Huelsenbeck 2003), which was used in this work and discussed in more details below.

Although some controversy exists regarding Bayesian methods—linked to the priors which are assumed *a priori* in the method—this problem is a philosophical issue rather than statistical (Felsenstein 2004).

Bayesian method is based on the Maximum Likelihood (ML) methods and the latter has been shown to perform very well even if the model violation is present

(Huelsenbeck *et al.* 2002) thus outperforming other methods. The Bayesian method therefore appears to be the best choice here especially due to its robustness towards model violation and relative ease of implementation.

Other advantages of the Bayesian method include better robustness against being stuck at a wrong local maximum of posterior distribution. MrBayes incorporates Metropolis coupling (MC) to improve MCMC sampling of the posterior probabilities distribution and lower the chance of being stuck in a localised "pseudo" maximum by means of use of several chains in each independent run, hence sometimes referred to as $(MC)^3$. These chains are controlled by heating parameter and therefore called hot (or heated) and "cold" chains. A swap is attempted after each generation step between two randomly chosen chains. Heated chains act like "scouts" to look for remote maxima and if such are found swap themselves with a cold chain and their states are switched (Ronquist 2004; Huelsenbeck and Ronquist 2001; Huelsenbeck *et al.* 2002, Lewis 2007). However, inferences are only made based on the "cold" chains (Huelsenbeck and Ronquist 2001).

Philosophically the Bayesian analysis is similar to a general path an experienced systematist employs—they base decisions about placement of taxa in different groups based on their own previous experience with the similar taxa and problems (Huelsenbeck *et al.* 2002).

5.4.1 Posterior probabilities

Another issue which as been given an extensive coverage in literature recently (see Simmons *et al.* 2004; Alfaro and Holder 2006;Huelsenbeck *et al.* 2002; Svennblad *et al.* 2006; see Bergsten 2005 for extensive discussion) is the interpretation of posterior probabilities. The debate in the literature is extensive, with points of view being from as simple as posterior probabilities being "equal to bootstrap values" (Hall 2004:p128) to that they seriously overestimate support values and perform poorly (Simmons *et al.* 2004). Generally however, the consensus in the literature appears to be that posterior probabilities tend to overestimate compared to bootstrap values, with the latter in their turn underestimating the support (Reed *et al.* 2002; Alfaro and Holder 2006, Simmons *et al.* 2004; Taylor and Piel 2004). It is also clear that bootstrap values cannot be used as a reference against which other tree support can be measured. Certainly interpretation of which is correct can lead to disputing of whether bootstrap is likely to cause Type I error (*i.e.* fail to support a correct true node) and

posterior probabilities to cause Type II error (*i.e.* fail to reject false tree) (Archibald *et al.* 2003; Huelsenbeck *et al.* 2001). Interpretation of posterior probabilities is therefore best to be left to understanding that a tree with a certain posterior probability has a chance equal to that probability of being true given priors, data and model rather than attempting to assess an absolute "trueness" of this tree.

Conversely, the finding and assertion that Bayesian values show overestimate of posterior probabilities of the branches was criticised as these probabilities are incomparable to bootstrap values because Bayesian analysis instead takes into consideration both data and the nucleotide substitution model (Huelsenbeck *et al.* 2002). Also a higher sensitivity of the Bayesian method to the model misspecification was suggested as one of the possible explanations of the difference between bootstrap and Bayesian posterior probabilities values (Huelsenbeck *et al.* 2002). However, the same authors state that there are no reasons to believe that a Bayesian method is more sensitive to the model parameters. This issue and debate is covered by several publications (Alfaro and Holder 2006; Ronquist 2004; Huelsenbeck 2002;Simmons *et al.* 2004; Svennblad *et al.* 2006; Yang and Rannala 2005) and will not be covered further here. Huelsenbeck *et al.* (2002) suggests several explanations to the problem of overestimation as well as some critique. He offers some explanation of possible reasons for overestimation of posterior probabilities relating to the underlying methodology and statistical interpretation of likelihoods with respect to statistical bias. Also, posterior probabilities are actually sometimes higher because the Bayesian method is more sensitive to the model settings. In a simulation study, Svennblad *et al.* (2006) found a considerable difference between ML and Bayesian methods and as a result they found these differences influenced the outcome of bootstrap values and posterior probabilities.

Ronquist (2004) stated that a branch with a posterior probability has an equivalent percentage chance of being there given that the model and priors are correct, and thus an incorrect posterior probabilities are essentially caused by the models being over simplified (Ronquist 2004). In this respect MrBayes and Bayesian MCMC methods are superior because they are capable of handling more complex models. Therefore, when an appropriate model is specified the Bayesian approach is superior to bootstrapping (Ronquist 2004). Bayesian methods also are much faster in general than ML methods (Archibald *et al.* 2002) especially in that non-parametric bootstrapping is not required to be performed.

5.4.2 MrBayes

In this work MrBayes was selected as the main phylogenetic analysis software as it has been in use for several years and has an extensive discussion and support information both on the Internet and in the literature. This program can be relatively easily compiled for the available multiprocessor cluster at Swansea and finally the author of this work had some prior knowledge of the program through the direct contact with the authors of MrBayes and training experience. This program has also "survived" several major revisions and updates (currently in version 3.2) and still is under constant development, improvement and research—version 4 is being currently developed (MrBayes WIKI; F.Ronquist personal communication). MrBayes allows the executing simultaneously of several independent runs for the same dataset. Thus starting with a completely different prior tree for each individual run—the individual runs can be spread through separate processors to speed up computation.

Another critical point about MrBayes is a decision about after how many generations to stop the analysis—so called convergence time. In this work it was in many cases limited by the computational power and allowed allocated time access to the supercomputer cluster[28], although a convergence was always required for the analysis results to be accepted. A direct correlation between number of generations and the computer power available to the researchers can be observed, if available publications are examined. The use of high performance computing was suggested (Sanderson and Shaffer 2002) as one of the best ways to deal with this problem. Indeed the only limitation appears to be that of computing hardware, which was mostly overcome in this work by utilising a 2 teraflop supercomputer cluster (UNIX IBM Blue-C). This allowed the spread of independent runs and individual chains throughout the cluster node (see below and Appendix C for more details). However, even an MPI[29] version of MrBayes is not capable of multithreading and thus is limited by the individual CPU performance (in case of MrBayes, each chain could be allocated to an individual CPU).

[28] For a more detailed description of the supercomputer cluster hardware architecture, please see Appendix C.

[29] Message Passing Interface (MPI) is computer software configuration that allows several nodes or computers within a computing cluster to communicate with one another thus allowing parallel computation; it is used in the cluster supercomputers.

There are a variety of methods to determine the length of run required however there appear to be no consensus or commonly accepted method.

5.4.3 Convergence diagnostics

One of the pitfalls of using Bayesian methods software is that it is impossible to know when the chain sampling has converged and this is considered to be one of the greatest practical problems of the MCMC methods (Huelsenbeck *et al.* 2001; Huelsenbeck *et al.* 2002; Ronquist and Huelsenbeck 2003). However, some methods are generally employed to evaluate the results of convergence.

Assessment of the correlation between the posterior probabilities of the individual clades found in separate chains and runs was suggested as one of the methods of checking for convergence (Huelsenbeck *et al.* 2001). Another method to evaluate convergence is to examine the behaviour of parameters such as posterior probability through the duration of the run. Log likelihoods would initially change in value but eventually, after so called burn-in time, would level off and fluctuate around a certain value. Although log likelihood plots are a very common tool for estimating convergence they are reported to be unreliable due to sudden change in values after an apparently reached plateau (Ronquist 2003; Huelsenbeck *et al.* 2002). To solve this problem a comparison of several independent runs was proposed (Huelsenbeck *et al.* 2002). In this work all analyses were performed with four independent runs. Inclusion of several hot chains in the analysis is one of the methods which increases the possibility of convergence. Monitoring of individual parameters of the evolutionary model in independent runs is also another way of detecting convergence. All of the above methods are now included in the MrBayes default setting and after the analysis runs are completed a summary is displayed, which allows assessment of the convergence. In addition, in this work a separate plot of log likelihood values and the posterior probabilities of splits (i.e. taxon bipartitions) over an entire MCMC analysis run were evaluated using AWTY (Wilgenbusch *et al.* 2004) online utility.

5.5 Models used in this work

Compensatory substitutions are well known to happen in rRNA. Different types of compensatory substitutions have been described and sometimes subdivided into compensatory and semi-compensatory (Ouvrard *et al.* 2000). Semicompensatory

substitutions are those, which do not disrupt helical structure, such as A-U being replaced by G-U.

Most models of molecular evolution assume independent substitutions (Felsenstein 2004) and thus are not quite suitable to the stem regions of the rRNA. Therefore models that consider pairs of sites—so called doublet models—are specifically designed for these types of interactions. The treatment of these RNA-specific models has seen some detailed attention in the literature recently with the availability of more advanced computational methods (Notredam 2000; Telford *et al.* 2005; Hudelot *et al.* 2003; Jow *et al.* 2002; Sullivan and Joyce 2005; Smith *et al.* 2004; Schoniger and Von Haeseler 1994). One of the fundamental aspects of rRNA is that the helical regions representing stems are conserved in order to preserve secondary structure and even tertiary structure. In other words, the interpretation and influence is unidirectional here from primary structure to secondary structure and eventually to the tertiary structure with secondary structure being the most energetically stable (Larsen 1992; Woese and Gutell 1989; Tinoco and Bustamante 1999). Compensatory mutations thus lie in the heart of structural formation of rRNA and they determine the stability and preservation of the structural helical units. Therefore, reliable estimates of RNA models are required to be used in phylogenetic reconstruction. The majority of RNA models are based on the 16 possible pairs of nucleotides and thus form 16x16 matrices. Models based on 16x16 matrices have been suggested by several authors (Schöniger and Von Haeseler 1994; Rzetski and Nei 1995) and a very detailed review of many other models (eighteen in total) is given by Savill *et al.* (2001). Savill *et al.*'s system of model numbering appears to have been used and accepted by many authors.

Not all of the 16 pairs occur all the time – the most frequently recorded are AU, GC, GU, UA, UG and CG. The rest of pairs are less frequent and are sometimes referred to as "mismatches". The above arrangements—with the six most common pairs— are referred to as 6-state models (Savill *et al.* 2001; Tillier and Collins 1995).

As has been seen with non-RNA substitution models, there is always a trade-off between overparameterisation and exact model fitting. In addition, more parameter-rich models take much longer to calculate and thus from a purely pragmatic point of view they are less advantageous.

The classification of models used in this work falls under 16 parameter models RNA16A-RNA16H by Savill *et al.*'s classification, where the last letters distinguish

the models by the number of free parameters employed by each model (i.e. number of frequency parameters, plus the number of rate parameters minus the number of constraints).

In simplified models (such as any RNA7) the first six states (1-6) of paired nucleotides are AU, GU, GC, UA, UG, CG with the remaining 10 referred to as "mismatches" and simply coded as MM. These seven states give number 7 to the name of the model[30].

Of particular interest here are the RNA16A and RNA16B models, which are simplified 16-state models (simplified from the RNA16 general reversible model which has 120 frequency rate parameters for each possible mismatch + 15 free parameters). The RNA16 general reversible model although described is not actually implemented in software and not used in phylogenetics reconstruction due to its complexity (Savill *et al.* 2001; Hudelot *et al.* 2003).

The RNA16A model (Table 8), which was found to be superior to many RNA models by Telford (2005) and Kosakovsky-Pond *et al.* (2007), includes 16 frequency parameters and 5 rate parameters, giving it a total of 19 free parameters.

Another variation on the 16 state model is called RNA16B (Savill *et al.* 2001). This model (Table 9) was originally described and proposed by Schöniger and Von Haeseler (1994). It is a simplification of the RNA16A model as it reduces the exchangeability parameters of a more complex RNA16A model to one, thus having 16 frequency parameters and one rate parameter μ. This model can also be described as a F81-like model for doublets of nucleotides.

The above model (RNA16B) is implemented, with slight modifications, in MrBayes - the software package that was used in this work. The model implemented there is a General Time Reversible (GTR) like modification of the RNA16B model: sometimes referred to as RNA16I (Gowri-Shankar and Jow 2006) or RNA16GTR (Telford *et al.* 2005). In MrBayes the number of rate parameters can be fixed to six, two or one via nset=x command line option and corresponds to RNA16GTR

[30] In most cases when an RNA model name is given the number in the name refers to the number of states or frequency parameters. For instance RNA7A model has 7 frequency parameters: π_1, π_2 ,π_3 ....π_7. However this is not always true as for instance model RNA6D has only three frequency parameters π_1, π_2 ,π_3 .

(RNA16I)[31], RNA16HKY (RNA16K) or RNA16F81 (RNA16B) models respectively. These models each estimate 20, 16 and 15 parameters respectively.

A very important difference between the RNA16A and RNA16B models is that the RNA16B model does not consider changes of pairs of nucleotides as does the RNA16A model, but instead it evaluates a single nucleotide change within the pair (stem pairs here) and if no change has occurred it treats it as zero (Savill *et al.* 2001; Telford *et al.* 2005; Gowri-Shankar and Jow 2006). Any compensatory change (i.e. a replacement of two nucleotides – one to compensate for another one being substituted) in the stem is thus evaluated as a simple two-step process of one nucleotide substitution and then the second nucleotide in exactly the same manner, each step with its own 4x4 model of nucleotide substitutions.

In an evaluation of different models (Savill *et al.* 2001; Telford *et al.* 2005) using log-likelihood and AIC statistics, best scores appeared to have achieved by the 6A, 7A, and 16A general reversible models based on well-known phylogenies. Telford *et al.* (2005) used a permutation test to select for the best fit models out of several RNA16-based models specifically for the stem regions of the partitioned dataset, specifying different models for stems and loops. In comparing RNA16AGTR, RNA16B (both RNA16BHKY and RNA16BGTR) and GTR models they found correlation in the nucleotide changes in the stems and therefore showed superiority of RNA16-based models over GTR models for the stems. Kosakovsky-Pond *et al.* (2007) in their comparison of genetic algorithm (GA) derived models to structural RNA models described by Savill *et al.* (2001) also found that the RNA16A model had the best AIC score out of several RNA structural models (although it performed considerably worse against GA-derived models). Telford *et al.* (2005) also found an improvement in the likelihood when the RNA16A model was used as opposed to RNA16B-type models. However, importantly for the model choice in this work there was a negligible improvement from the use of RNA16BGTR over the RNA16BHKY model (likelihood values of -2820.78 and -2823.82 respectively).

In the current work a stem-specific model was chosen based on several factors. The first is the previous findings of the superiority of 16-state RNA specific

[31] The names in brackets represent a commonly used system, based on Savill *et al.* (2001), such as the one used in PHASE software (Hudelot *et al.* 2003). This is given here to avoid ambiguity as different sources appear to refer to the same models with different names. Further, in this work, the RNA16B model and all its derivatives are referred to.

models over GTR models when used for stems. Secondly, the choice was limited by the models implemented in the software which was used here. As mentioned above MrBayes employs the RNA16B model and its derivatives as doublet type models. Because the RNA16BGTR model appears to be only negligibly better than RNA16BHKY, this model was eventually chosen as the stem model. Its transition matrix is shown on Table 10. The model choice eventually affected the calculation simplicity as the dataset which has the RNA16GTR model specified for stems would take significantly longer to calculate on MrBayes and of course RNA16HKY reduced the effect of possible overparameterisation, which was shown to introduce extra "noise" in data (Huelsenbeck *et al.* 2002).

$$Q = m_r \times \begin{Bmatrix}
 & AU & GU & GC & UA & UG & CG & AA & AG & AC & GA & GG & CA & CC & CU & UC & UU \\
AU & * & \pi_{GU}\alpha_1 & \pi_{GC} & \pi_{UA}\alpha_2 & \pi_{UG}\alpha_2 & \pi_{CG}\alpha_2 & \pi_{AA}\alpha_3 & \pi_{AG}\alpha_3 & \pi_{AC}\alpha_3 & 0 & 0 & 0 & 0 & \pi_{CU}\alpha_3 & 0 & \pi_{UU}\alpha_3 \\
GU & \pi_{AU}\alpha_1 & * & \pi_{GC}\alpha_1 & \pi_{UA}\alpha_2 & \pi_{UG}\alpha_2 & \pi_{CG}\alpha_2 & 0 & 0 & 0 & \pi_{GA}\alpha_3 & \pi_{GG}\alpha_3 & 0 & 0 & \pi_{CU}\alpha_3 & 0 & \pi_{UU}\alpha_3 \\
GC & \pi_{AU} & \pi_{GU}\alpha_1 & * & \pi_{UA}\alpha_2 & \pi_{UG}\alpha_2 & \pi_{CG}\alpha_2 & 0 & 0 & \pi_{AC}\alpha_3 & \pi_{GA}\alpha_3 & \pi_{GG}\alpha_3 & 0 & \pi_{CC}\alpha_3 & 0 & \pi_{UC}\alpha_3 & 0 \\
UA & \pi_{AU}\alpha_2 & \pi_{GU}\alpha_2 & \pi_{GC}\alpha_2 & * & \pi_{UG}\alpha_1 & \pi_{CG} & \pi_{AA}\alpha_3 & 0 & 0 & \pi_{GA}\alpha_3 & 0 & \pi_{CA}\alpha_3 & 0 & 0 & \pi_{UC}\alpha_3 & \pi_{UU}\alpha_3 \\
UG & \pi_{AU}\alpha_2 & \pi_{GU}\alpha_2 & \pi_{GC}\alpha_2 & \pi_{UA}\alpha_1 & * & \pi_{CG}\alpha_1 & 0 & \pi_{AG}\alpha_3 & 0 & 0 & \pi_{GG}\alpha_3 & 0 & 0 & 0 & \pi_{UC}\alpha_3 & \pi_{UU}\alpha_3 \\
CG & \pi_{AU}\alpha_2 & \pi_{GU}\alpha_2 & \pi_{GC}\alpha_2 & \pi_{UA} & \pi_{UG}\alpha_1 & * & 0 & \pi_{AG}\alpha_3 & 0 & 0 & \pi_{GG}\alpha_3 & \pi_{CA}\alpha_3 & \pi_{CC}\alpha_3 & \pi_{CU}\alpha_3 & 0 & 0 \\
AA & \pi_{AU}\alpha_3 & 0 & 0 & \pi_{UA}\alpha_3 & 0 & 0 & * & \pi_{AG}\alpha_4 & \pi_{AC}\alpha_4 & \pi_{GA}\alpha_4 & 0 & \pi_{CA}\alpha_4 & 0 & 0 & 0 & 0 \\
AG & \pi_{AU}\alpha_3 & 0 & 0 & 0 & \pi_{UG}\alpha_3 & \pi_{CG}\alpha_3 & \pi_{AA}\alpha_4 & * & \pi_{AC}\alpha_4 & 0 & \pi_{GG}\alpha_4 & 0 & 0 & 0 & 0 & 0 \\
AC & \pi_{AU}\alpha_3 & 0 & \pi_{GC}\alpha_3 & 0 & 0 & 0 & \pi_{AA}\alpha_4 & \pi_{AG}\alpha_4 & * & 0 & 0 & 0 & \pi_{CC}\alpha_4 & 0 & \pi_{UC}\alpha_4 & 0 \\
GA & 0 & \pi_{GU}\alpha_3 & \pi_{GC}\alpha_3 & \pi_{UA}\alpha_3 & 0 & 0 & \pi_{AA}\alpha_4 & 0 & 0 & * & \pi_{GG}\alpha_4 & \pi_{CA}\alpha_4 & 0 & 0 & 0 & 0 \\
GG & 0 & \pi_{GU}\alpha_3 & \pi_{GC}\alpha_3 & 0 & \pi_{UG}\alpha_3 & \pi_{CG}\alpha_3 & 0 & \pi_{AG}\alpha_4 & 0 & \pi_{GA}\alpha_4 & * & 0 & 0 & 0 & 0 & 0 \\
CA & 0 & 0 & 0 & \pi_{UA}\alpha_3 & 0 & \pi_{CG}\alpha_3 & \pi_{AA}\alpha_4 & 0 & 0 & \pi_{GA}\alpha_4 & 0 & * & \pi_{CC}\alpha_4 & \pi_{CU}\alpha_4 & 0 & 0 \\
CC & 0 & 0 & \pi_{GC}\alpha_3 & 0 & 0 & \pi_{CG}\alpha_3 & 0 & 0 & \pi_{AC}\alpha_4 & 0 & 0 & \pi_{CA}\alpha_4 & * & \pi_{CU}\alpha_4 & \pi_{UC}\alpha_4 & 0 \\
CU & \pi_{AU}\alpha_3 & \pi_{GU}\alpha_3 & 0 & 0 & 0 & \pi_{CG}\alpha_3 & 0 & 0 & 0 & 0 & 0 & \pi_{CA}\alpha_4 & \pi_{CC}\alpha_4 & * & 0 & \pi_{UU}\alpha_4 \\
UC & 0 & 0 & \pi_{GC}\alpha_3 & \pi_{UA}\alpha_3 & \pi_{UG}\alpha_3 & 0 & 0 & 0 & \pi_{AC}\alpha_4 & 0 & 0 & 0 & \pi_{CC}\alpha_4 & 0 & * & \pi_{UU}\alpha_4 \\
UU & \pi_{AU}\alpha_3 & \pi_{GU}\alpha_3 & 0 & \pi_{UA}\alpha_3 & \pi_{UG}\alpha_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \pi_{CU}\alpha_4 & \pi_{UC}\alpha_4 & *
\end{Bmatrix}$$

Table 8 RNA16A model transition matrix. This model is simplified from an RNA16 16-state model. In it α_1 is a rate of single transition, α_2 a rate of double transition, α_3 a rate of mismatch to non-mismatch transition requiring only one substitution, α_4 a rate of mismatch to mismatch transition requiring only one substitution (details of the model taken from PHASE manual).

$$Q = m_r \times \begin{Bmatrix}
 & AU & GU & GC & UA & UG & CG & AA & AG & AC & GA & GG & CA & CC & CU & UC & UU \\
AU & * & \pi_{GU} & 0 & 0 & 0 & 0 & \pi_{AA} & \pi_{AG} & \pi_{AC} & 0 & 0 & 0 & 0 & \pi_{CU} & 0 & \pi_{UU} \\
GU & \pi_{AU} & * & \pi_{GC} & 0 & 0 & 0 & 0 & 0 & 0 & \pi_{GA} & \pi_{GG} & 0 & 0 & \pi_{CU} & 0 & \pi_{UU} \\
GC & 0 & \pi_{GU} & * & 0 & 0 & 0 & 0 & 0 & \pi_{AC} & \pi_{GA} & \pi_{GG} & 0 & \pi_{CC} & 0 & \pi_{UC} & 0 \\
UA & 0 & 0 & 0 & * & \pi_{UG} & 0 & \pi_{AA} & 0 & 0 & \pi_{GA} & 0 & \pi_{CA} & 0 & 0 & \pi_{UC} & \pi_{UU} \\
UG & 0 & 0 & 0 & \pi_{UA} & * & \pi_{CG} & 0 & \pi_{AG} & 0 & 0 & \pi_{GG} & 0 & 0 & 0 & \pi_{UC} & \pi_{UU} \\
CG & 0 & 0 & 0 & 0 & \pi_{UG} & * & 0 & \pi_{AG} & 0 & 0 & \pi_{GG} & \pi_{CA} & \pi_{CC} & \pi_{CU} & 0 & 0 \\
AA & \pi_{AU} & 0 & 0 & \pi_{UA} & 0 & 0 & * & \pi_{AG} & \pi_{AC} & \pi_{GA} & 0 & \pi_{CA} & 0 & 0 & 0 & 0 \\
AG & \pi_{AU} & 0 & 0 & 0 & \pi_{UG} & \pi_{CG} & \pi_{AA} & * & \pi_{AC} & 0 & \pi_{GG} & 0 & 0 & 0 & 0 & 0 \\
AC & \pi_{AU} & 0 & \pi_{GC} & 0 & 0 & 0 & \pi_{AA} & \pi_{AG} & * & 0 & 0 & 0 & \pi_{CC} & 0 & \pi_{UC} & 0 \\
GA & 0 & \pi_{GU} & \pi_{GC} & \pi_{UA} & 0 & 0 & \pi_{AA} & 0 & 0 & * & \pi_{GG} & \pi_{CA} & 0 & 0 & 0 & 0 \\
GG & 0 & \pi_{GU} & \pi_{GC} & 0 & \pi_{UG} & \pi_{CG} & 0 & \pi_{AG} & 0 & \pi_{GA} & * & 0 & 0 & 0 & 0 & 0 \\
CA & 0 & 0 & 0 & \pi_{UA} & 0 & \pi_{CG} & \pi_{AA} & 0 & 0 & \pi_{GA} & 0 & * & \pi_{CC} & \pi_{CU} & 0 & 0 \\
CC & 0 & 0 & \pi_{GC} & 0 & 0 & \pi_{CG} & 0 & 0 & \pi_{AC} & 0 & 0 & \pi_{CA} & * & \pi_{CU} & \pi_{UC} & 0 \\
CU & \pi_{AU} & \pi_{GU} & 0 & 0 & 0 & \pi_{CG} & 0 & 0 & 0 & 0 & 0 & \pi_{CA} & \pi_{CC} & * & 0 & \pi_{UU} \\
UC & 0 & 0 & \pi_{GC} & \pi_{UA} & \pi_{UG} & 0 & 0 & 0 & \pi_{AC} & 0 & 0 & 0 & \pi_{CC} & 0 & * & \pi_{UU} \\
UU & \pi_{AU} & \pi_{GU} & 0 & \pi_{UA} & \pi_{UG} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \pi_{CU} & \pi_{UC} & *
\end{Bmatrix}$$

Table 9 RNA16B model (Schöniger and Von Haeseler 1994) transition matrix, RNA16B is equal to a F81-like model of doublets, with all substitutions being equally likely.

$$Q = m_r \times \begin{Bmatrix}
 & AU & GU & GC & UA & UG & CG & AA & AG & AC & GA & GG & CA & CC & CU & UC & UU \\
AU & * & \pi_{GU} & 0 & 0 & 0 & 0 & \pi_{AA}\alpha_1 & \pi_{AG}\alpha_1 & \pi_{AC} & 0 & 0 & 0 & 0 & \pi_{CU}\alpha_1 & 0 & \pi_{UU}\alpha_1 \\
GU & \pi_{AU} & * & \pi_{GC} & 0 & 0 & 0 & 0 & 0 & 0 & \pi_{GA}\alpha_1 & \pi_{GG}\alpha_1 & 0 & 0 & \pi_{CU}\alpha_1 & 0 & \pi_{UU}\alpha_1 \\
GC & 0 & \pi_{GU} & * & 0 & 0 & 0 & 0 & 0 & \pi_{AC} & \pi_{GA}\alpha_1 & \pi_{GG}\alpha_1 & 0 & \pi_{CC}\alpha_1 & 0 & \pi_{UC}\alpha_1 & 0 \\
UA & 0 & 0 & 0 & * & \pi_{UG} & 0 & \pi_{AA}\alpha_1 & 0 & 0 & \pi_{GA}\alpha_1 & 0 & \pi_{CA} & 0 & 0 & \pi_{UC}\alpha_1 & \pi_{UU}\alpha_1 \\
UG & 0 & 0 & 0 & \pi_{UA} & * & \pi_{CG} & 0 & \pi_{AG}\alpha_1 & 0 & 0 & \pi_{GG} & 0 & 0 & 0 & \pi_{UC}\alpha_1 & \pi_{UU}\alpha_1 \\
CG & 0 & 0 & 0 & 0 & \pi_{UG} & * & 0 & \pi_{AG}\alpha_1 & 0 & 0 & \pi_{GG}\alpha_1 & \pi_{CA} & \pi_{CC}\alpha_1 & \pi_{CU}\alpha_1 & 0 & 0 \\
AA & \pi_{AU}\alpha_1 & 0 & 0 & \pi_{UA}\alpha_1 & 0 & 0 & * & \pi_{AG} & \pi_{AC}\alpha_1 & \pi_{GA} & 0 & \pi_{CA}\alpha_1 & 0 & 0 & 0 & 0 \\
AG & \pi_{AU}\alpha_1 & 0 & 0 & 0 & \pi_{UG}\alpha_1 & \pi_{CG} & \pi_{AA} & * & \pi_{AC}\alpha_1 & 0 & \pi_{GG} & 0 & 0 & 0 & 0 & 0 \\
AC & \pi_{AU} & 0 & \pi_{GC} & 0 & 0 & 0 & \pi_{AA}\alpha_1 & \pi_{AG}\alpha_1 & * & 0 & 0 & 0 & \pi_{CC}\alpha_1 & 0 & \pi_{UC}\alpha_1 & 0 \\
GA & 0 & \pi_{GU}\alpha_1 & \pi_{GC}\alpha_1 & \pi_{UA}\alpha_1 & 0 & 0 & \pi_{AA} & 0 & 0 & * & \pi_{GG} & \pi_{CA}\alpha_1 & 0 & 0 & 0 & 0 \\
GG & 0 & \pi_{GU}\alpha_1 & \pi_{GC}\alpha_1 & 0 & \pi_{UG} & \pi_{CG}\alpha_1 & 0 & \pi_{AG} & 0 & \pi_{GA} & * & 0 & 0 & 0 & 0 & 0 \\
CA & 0 & 0 & 0 & \pi_{UA} & 0 & \pi_{CG} & \pi_{AA}\alpha_1 & 0 & 0 & \pi_{GA}\alpha_1 & 0 & * & \pi_{CC}\alpha_1 & \pi_{CU}\alpha_1 & 0 & 0 \\
CC & 0 & 0 & \pi_{GC}\alpha_1 & 0 & 0 & \pi_{CG}\alpha_1 & 0 & 0 & \pi_{AC}\alpha_1 & 0 & 0 & \pi_{CA}\alpha_1 & * & \pi_{CU} & \pi_{UC} & 0 \\
CU & \pi_{AU}\alpha_1 & \pi_{GU}\alpha_1 & 0 & 0 & 0 & \pi_{CG}\alpha_1 & 0 & 0 & 0 & 0 & 0 & \pi_{CA}\alpha_1 & \pi_{CC} & * & 0 & \pi_{UU} \\
UC & 0 & 0 & \pi_{GC}\alpha_1 & \pi_{UA}\alpha_1 & \pi_{UG}\alpha_1 & 0 & 0 & 0 & \pi_{AC}\alpha_1 & 0 & 0 & 0 & \pi_{CC} & 0 & * & \pi_{UU} \\
UU & \pi_{AU}\alpha_1 & \pi_{GU}\alpha_1 & 0 & \pi_{UA}\alpha_1 & \pi_{UG}\alpha_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \pi_{CU} & \pi_{UC} & *
\end{Bmatrix}$$

Table 10 RNA16HKY model (modified from Schöniger and Von Haeseler 1994) transition matrix as used in this work. RNA16HKY is equal to a HKY-like model of doublets, with transitions-transversions allowed to have a different substitutions rates. This model alongside with RNA16GTR is implemented in MrBayes with parameters fixed using lset nst=x command, where "x" could be either 1,2 or 6 for F81, HKY or GTR models respectively. In this case the GTR rates are $\alpha_1=\alpha_2=\alpha_3=\alpha_5=\alpha/\beta=\kappa$-ratio and $\alpha_4=4$.

5.5.1 Model selection

In this work, several models and scenarios of analysis were considered and tested. A diagram (Figure 44) outlining all combinations of analysis is shown below.

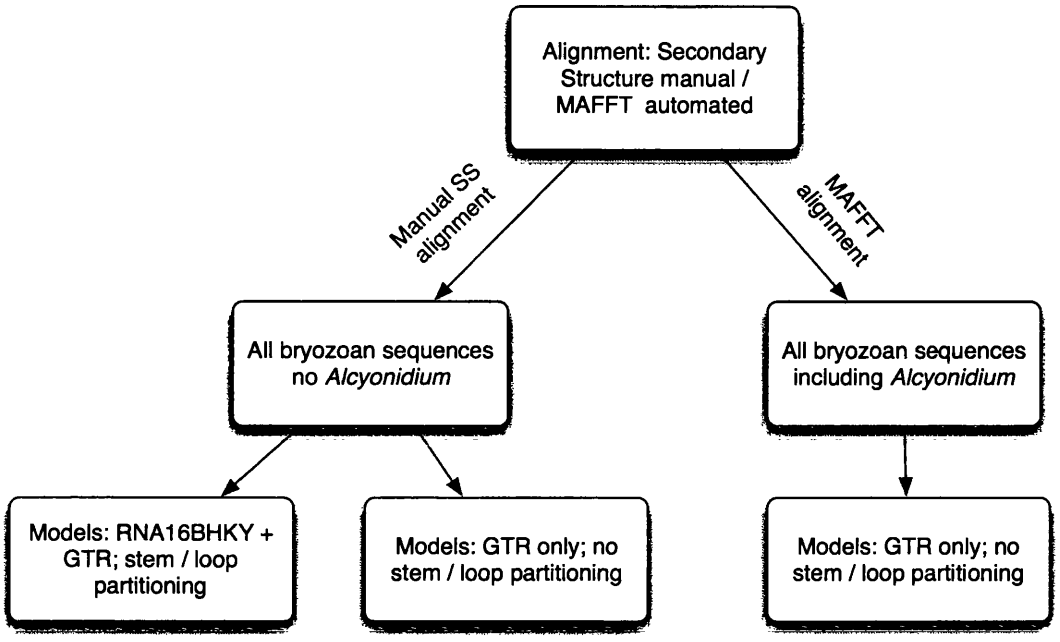


Figure 44 Diagram showing main types of analysis performed in this work.

The following different analyses were first separated based on the alignment method that was used for the creation of the dataset. Most of the bryozoan sequences were aligned using secondary structure models as a template (see previous chapter for details). However, three sequences, which belonged to the family Alcyonidiidae, namely *Alcyonidium gelatinosum*, *Alcyonidium hirsutum* and *Alcyonidium polyoum,* could not be aligned using an existing secondary structure model.

These sequences were considerably longer (2168 nucleotides on average instead of average 1797 for other sequences in general) due to several insertions as well as contents—the sequences were very difficult to align due to considerable differences in nucleotide composition.

Due to these differences the sequences of *Alcyonidium* species were excluded from the main secondary structure alignment. In order not to lose valuable data and in order to evaluate these sequences another alignment was created using the MAFFT (Katoh *et al.* 2002; Katoh *et al.* 2005; Katoh and Toh 2007) alignment package. Therefore, two principal alignment files were used in the analysis. All alignment files were prepared in NEXUS file interchange format to be used in MrBayes (including

MrBayes specific formatting). The evolutionary model selection was done using software Modeltest and MrModeltest – see below.

Further, the structure based alignment file was separated into two datasets and formatted for the use in MrBayes using Xstem utility (Telford *et al.* 2005). This utility converts a DCSE data file (which was created in the previous steps of manual alignment) into a NEXUS file format, which also includes all necessary secondary structure information suitable for the doublet model used in MrBayes. Thus the first dataset included an addition of a separate dataset block at the end of the alignment sequences which indicates two character-partitions, *loop* and *stem*, and shows the exact position of each nucleotide in the stems and loops. Also all nucleotide pairs have to be specified for the doublet model using the PAIRS command, for instance pairs 4:20, 5:19, 6:18, 7:17, 8:16, *etc*[32]. The above data partitioning allowed the performing of the phylogenetic reconstruction using separate models: one for the stem regions (RNA16BHKY+I+Γ) and one for the loop regions (GTR+I+Γ).

The second dataset, although derived from the same DCSE file (i.e. the one aligned using secondary structure) was stripped of the partition data and converted from DCSE to NEXUS format using the same utility (Xstem) to be used for MrBayes, but this time the evolutionary model was evaluated for the entire dataset. The model selection for the latter dataset was done using software script which performs a batch models evaluation in the PAUP package: Modeltest (Posada 2006) and MrModeltest (Nylander 2004). Both scripts implement two model selection methods Akaike Information Criterion (AIC) and Hierarchical Likelihood Ratio Test (hLRTs), the latter script is specifically written to evaluate only those nucleotide models implemented in MrBayes and ignores the rest.

The necessity of model selection in phylogenetics has been obvious for a long time as has the consideration of such factors as multiple substitutions per site and substitutional saturation, Felsenstein zone[33] model criticality (*i.e.* a condition when rapidly evolving taxa cause unusually long branches to attract to each other when a maximum parsimony method is employed), under and overparameterisation, all of

---

[32] For an example of the file formatted by Xstem for the use in MrBayes see Appendix C. The section specific to the secondary structure model is written in the section beginning with the *begin mrbayes* command.

[33] "Felsenstein zone" refers to the top left corner of the tree parameter simulation space which corresponds to a short internal edge and two long terminal edges of a tree (Page and Holmes 2002).

which play an important role in models and their selection (Sullivan 2005). The importance of substitution models was realised quite early—the simplest one was described by Jukes and Cantor (1969). Selection of a correct and suitable model has been long a topic of special interest of many authors. The one point of view that has been quoted by almost every author discussing the issue is that of Box (1976) "all models are wrong but some describe natural phenomena better than others"[34]. This reflects the reality of models in phylogenetics. However, many tools and methods have been proposed. One of the most popular methods of model selection (implemented both in Modeltest and MrModeltest) is LRT with likelihood score being used as a measure of model fitness:

- $\delta = 2(\ln L1 - \ln L0)$, where "*L1*" – Likelihood score of the more complex model.

This method is limited to the models which are nested, in practice all models being a special case of a GTR model. This test is implemented in the Modeltest program which uses hierarchical approach to the nested models—hence hLRTs. One of the biggest criticisms of this approach is that traversing a tree-like space of hierarchical models is done pairwise in one direction and model selection outcome can be altered and influenced by the starting model, or fail to select the best model altogether (Sullivan and Joyce 2005; Posada and Buckley 2004).

As a result some authors (e.g. Sullivan and Joyce 2005; Posada and Buckley 2004) suggest the use of alternative methods. One such is the Akaike Information Criterion (AIC) (Akaike 1974). AIC measures the amount of lost information when a specific model is used as an estimate of a real evolutionary process (Posada and Buckley 2004).

- $AIC = -2\ln L + 2k$, where "*k*" is the number of independently adjusted parameters in the model and "*L*" is the maximum value of the likelihood function.

[34] Ironically, the quote "all models are wrong, but some are useful" appears to be incorrect as nowhere in his paper Box actually says that. The article itself though is an excellent excursion into philosophy of science and scientific method.

The most important advantage of the AIC method is that it allows a simultaneous testing of independent non-nested models. This method has recently been implemented in Modeltest and MrModeltest software and thus was used in this work.

5.5.2 Results of MrModeltest and Modeltest model selection

When selection of the suitable model is completed, using Modeltest (MrModeltest is identical to Modeltest in its algorithm), the program requires the PAUP software package (Swofford 2003) to be used to calculate scores and build a NJ tree out of the data. For this work both secondary structure based and MAFFT alignments were loaded in NEXUS format into PAUP and after that Modeltest PAUP batch script was executed. After that the generated scores were evaluated using the Modeltest program. The results for both alignments are given below. For comparison hLRTs were performed as well. These gave the same model selection as the AIC method.

5.5.2.1 Results of model selection for the structure-assisted alignment

Results of the hLTRs for the dataset of structure-assisted alignment from MrModeltest are given in Table 11, Modeltest results were identical and not shown. All four independent hierarchy analyses selected the same model as depicted in the table. Results of the AIC test performed by MrModeltest are shown in Table 12 (Modeltest results are identical). In Table 12 the values of Δ AIC_c are given, which represent the difference over all presented models and are crucial when reporting AIC model selection due to AIC being on a relative scale (Posada 2004).

- Δ AIC_c = AIC_c – *min* AIC, where "min AIC" is the smallest AIC value among all candidate models.

The model selection is based on the AIC_c values—the model with the lowest AIC_c is selected. Also the Δ AIC_c allows the evaluation and consideration of more than one model for those models where Δ $AIC_c \leq 2$. This is based on the assumption that the larger the AIC difference between two models the less likely this is the best model to describe the real process of nucleotide substitution. In this case the difference of AIC ≤ 2 for a model is a proposed guideline value for models which receive substantial support (Posada and Buckley 2004).

Additionally, in all relevant tables below, Alkaike weights are given which are sometimes used for assessing the models selection uncertainty (Posada 2004). These weights are normalised approximations of the relative likelihood of the model given the data. These values are not assessed here, but as they are calculated by the software together with AIC they are given for the information only, while GTR+I+Γ is best by AIC_c and weight.

Table 11 Results of the hLRTs test for model selection using MrModeltest software. The dataset is for the structure-assisted alignment. The table gives all estimated model parameters, which may be required by some software. In this case the value of –LnL is important for model selection. The rest of the parameters are estimated by MrBayes during its run and thus given here for information only.

Model selected: GTR+I+Γ	
-lnL =	14823.7754
K (number of estimated—free—parameters) =	10
P-value =	<0.000001
Base frequencies:	
freqA =	0.2196
freqC =	0.2599
freqG =	0.2974
freqT =	0.2231
Substitution model:	
Rate matrix	
R(a) [A-C] =	0.9384
R(b) [A-G] =	1.5539
R(c) [A-T] =	1.4026
R(d) [C-G] =	1.0916
R(e) [C-T] =	3.1308
R(f) [G-T] =	1.0000
Among-site rate variation	
Proportion of invariable sites (I) =	0.3778
Variable sites (G)	
Gamma distribution shape parameter =	0.8156

Table 12 Showing results of the Akaike weights selection using MrModeltest for the structure-assisted data alignment. -lnL: negative log likelihood; K: number of estimated (free) parameters; AIC_c: Akaike values; Δ AIC_c: Akaike values differences; weight: information weight.

Model	-lnL	K	AIC_c	Δ AIC_c	Weight
GTR+I+Γ	14823.7754	10	29667.5508	0.0000	1.0000
SYM+I+Γ	14846.5254	7	29707.0508	39.5000	2.65e-09
HKY+I+Γ	14857.2910	6	29726.5820	59.0312	1.52e-13
GTR+Γ	14867.9561	9	29753.9121	86.3613	1.77e-19
K80+I+Γ	14887.5957	3	29781.1914	113.6406	2.10e-25
SYM+Γ	14890.8613	6	29793.7227	126.1719	4.00e-28
HKY+Γ	14904.3389	5	29818.6777	151.1270	1.52e-33
K80+Γ	14934.4707	2	29872.9414	205.3906	2.80e-45
F81+I+Γ	14975.9814	5	29961.9629	294.4121	0.00e+00
JC+I+Γ	15002.8955	2	30009.7910	342.2402	0.00e+00
F81+Γ	15022.0127	4	30052.0254	384.4746	0.00e+00
JC+Γ	15048.3936	1	30098.7871	431.2363	0.00e+00
GTR+I	15043.2959	9	30104.5918	437.0410	0.00e+00
SYM+I	15056.1426	6	30124.2852	456.7344	0.00e+00
HKY+I	15066.5791	5	30143.1582	475.6074	0.00e+00
K80+I	15090.7441	2	30185.4883	517.9375	0.00e+00
F81+I	15173.8408	4	30355.6816	688.1309	0.00e+00
JC+I	15196.5947	1	30395.1895	727.6387	0.00e+00
GTR	15979.7832	8	31975.5664	2308.0156	0.00e+00
SYM	15993.8037	5	31997.6074	2330.0566	0.00e+00
HKY	16042.5840	4	32093.1680	2425.6172	0.00e+00
K80	16061.8730	1	32125.7461	2458.1953	0.00e+00
F81	16144.1113	3	32294.2227	2626.6719	0.00e+00
JC	16159.3818	0	32318.7637	2651.2129	0.00e+00

5.5.2.2 Results of model selection for the non-structural alignment

Results of the hLTRs for the dataset of non-structural alignment from MrModeltest are given in Table 13. Modeltest results were identical and not shown, all four independent hierarchical searches selected the same model. Results of the AIC test performed by MrModeltest shown in Table 14 (Modeltest results were identical and not shown).

Table 13 Results of the hLRTs test for model selection using MrModeltest software. The dataset is for the non-assisted alignment of sequences, using MAFFT software.

Model selected: GTR+I+Γ
-lnL = 19651.3047
K (number of estimated—free—parameters) = 10
AIC = 39322.6094
Base frequencies:
freqA = 0.2163
freqC = 0.2661
freqG = 0.3020
freqT = 0.2156
Substitution model:
Rate matrix
R(a) [A-C] = 0.9436
R(b) [A-G] = 1.7235
R(c) [A-T] = 1.4955
R(d) [C-G] = 0.8913
R(e) [C-T] = 3.2239
R(f) [G-T] = 1.0000
Among-site rate variation
Proportion of invariable sites (I) = 0.1787
Variable sites (G)
Gamma distribution shape parameter = 0.5153

Table 14 Showing results of the Akaike weights selection using MrModeltest for non-structure specific data alignment, using MAFFT. -lnL: negative log likelihood; K: number of estimated (free) parameters; AIC_c: Akaike values; Δ AIC_c: Akaike values differences; weight: information weight.

Model	-lnL	K	AIC_c	ΔAIC_c	Weight
GTR+I+Γ	19651.3047	10	39322.6094	0.0000	1.0000
GTR+Γ	19664.9902	9	39347.9805	25.3711	3.10e-06
HKY+I+Γ	19694.1113	6	39400.2227	77.6133	1.40e-17
SYM+I+Γ	19696.9473	7	39407.8945	85.2852	3.02e-19
HKY+Γ	19708.4082	5	39426.8164	104.2070	2.35e-23
SYM+Γ	19711.0840	6	39434.1680	111.5586	5.96e-25
K80+I+Γ	19733.2441	3	39472.4883	149.8789	2.85e-33
K80+Γ	19747.9160	2	39499.8320	177.2227	3.29e-39
F81+I+Γ	19889.3867	5	39788.7734	466.1641	0.00e+00
F81+Γ	19903.8594	4	39815.7188	493.1094	0.00e+00
JC+I+Γ	19920.6934	2	39845.3867	522.7773	0.00e+00
JC+Γ	19935.2637	1	39872.5273	549.9180	0.00e+00
GTR+I	20183.5391	9	40385.0781	1062.4688	0.00e+00
HKY+I	20199.5430	5	40409.0859	1086.4766	0.00e+00
SYM+I	20207.0469	6	40426.0938	1103.4844	0.00e+00
K80+I	20229.4277	2	40462.8555	1140.2461	0.00e+00
F81+I	20367.7949	4	40743.5898	1420.9805	0.00e+00
JC+I	20393.7227	1	40789.4453	1466.8359	0.00e+00
GTR	21104.7207	8	42225.4414	2902.8320	0.00e+00
SYM	21128.5020	5	42267.0039	2944.3945	0.00e+00
HKY	21153.5176	4	42315.0352	2992.4258	0.00e+00
K80	21180.8281	1	42363.6562	3041.0469	0.00e+00
F81	21319.4160	3	42644.8320	3322.2227	0.00e+00
JC	21339.5117	0	42679.0234	3356.4141	0.00e+00

5.5.3 GTR+I+Γ model

In addition, as can be seen from the models selection above, among-site rate variation was considered and implemented. This is done using the gamma (Γ) distribution with shape parameter alpha α. Also, proportion of invariable sites (I) was estimated using the same software. Thus, the model, which was selected using MrModeltest was general time reversible with proportion of invariable sites and variable sites parameter

(GTR+I+Γ). Although, both "I" and "α" are estimated during the MrModeltest model estimation test they are actually not used in MrBayes as these parameters are estimated during the Bayesian search algorithm. These parameters are given with each tree displayed below. For the partitioned data set a loop model was selected empirically, and some other authors (Telford *et al.* 2005) suggest using a similar mode for the partitioned set, i.e. GTR+I+Γ.

5.6 Details of MrBayes analyses

MrBayes was compiled as an mpi UNIX version on an IBM UNIX Blue C cluster running 'AIX' (proprietary UNIX from IBM)[35], at the Institute of Life Sciences Swansea University. The cluster consists of 16 server-nodes each having 16 processors. As it is not possible to use MrBayes on the cluster in an interactive mode - batch files were used for each run consisting of the standard nexus file for MrBayes, which also included all necessary information for the run to be performed in non-interactive mode (see Appendix C for technical examples). Once calculations on the cluster were finished files were downloaded to a portable computer and analysed locally. All runs were limited to four days. The availability of a supercomputer cluster dramatically sped up calculation time. For instance a dataset of 33 species with a partitioned model set (*i.e* GTR+I+Γ and RNA16HKY models for loops and stems respectively) when loaded would take approximately 170 days to compute 20 million generations of four runs with four chains on a standalone Macintosh desktop computer (PowerPC G5 with 1.5Gb of RAM), exactly the same dataset was possible to calculate within 4-6 days on the Blue C cluster!

It is recommended (Ronquist and Huelsenbeck 2003) to run a minimum of four chains and two independent runs for each data set, and more independent runs starting with independent random trees generally increase the chances of "good sampling" of posterior probabilities (Ronquist and Huelsenbeck 2003). The number of chains was limited to the default setting of four – three hot chains and one cold per run. The heated chains "temperature" parameters were left at the default values (0.2) as well as priors settings. All together each analysis was therefore running four chains

[35] For details on how the software was compiled please see Appendix C. Also a modified "make" file is available upon request from the author.

for four runs thus requiring 16 individual processors, which was perfectly suitable for utilising one node on the cluster – i.e. 16 CPUs.

In order to estimate number of generatios and time required for the runs to complete, the data were tested originally using 1 million generations. Whilst the calculation is performed, MrBayes generates an output file in which each generations (as specified in the batch file) results are printed out:

```
1000 -- (-18401.318) [...15 remote chains...] -- 26:39:54 (time)
Average standard deviation of split frequencies: 0.250089
2000 -- (-17736.922) [...15 remote chains...] -- 26:39:48 (time)
Average standard deviation of split frequencies: 0.178495
```

As can be seen from the above example, results of each 1000th generation are printed out with estimated time given at the end (in here for example it 26 hours 39 minutes and 54 seconds). The estimated time is meant to show how long it remains for the process to run – this is supposed to give a rather precise estimation of time required for the run to complete (Hall 2004; Ronquist and Huelsenbeck 2003). The time is dependent on many factors such as data alignment size, model chosen, number of species etc and of course on the actual capabilities of the CPU it runs on. Unfortunately, in our case it appeared to underestimate the required time considerably. Consequently, given the initial time limit for the runs to be of maximum 96 hours the analysis was limited to a maximum 20 million generations for the "light" model (i.e. GTR+I+Γ only) and to around 14-16 million generations for the mixed model runs (i.e. GTR+I+Γ and RNA16BHLY+Γ+I).

5.7 Results for individual MrBayes runs

5.7.1 Non-structural alignment (GTR+I+Γ model)

Below are the results from a non-structural alignment for all bryozoan sequences collected in this work including those of *Alcyonidium* species and seven outgroup species (as described above). The analysis was run for 16,000,000 generations with sample frequency of every 1000th generation, thus recording 16,000 trees. The burn-in period (*i.e.* the number of generations required to attain stationarity) was determined using log likelihood plots of all values sampled during the analysis and graphical results from AWTY (Wilgenbuch *et al.* 2004) online utility. This utility allows a plot

of the cumulative posterior probabilities for taxon bipartitions (taxon splits) over the generations and visual evaluation of when the probabilities of these bipartitions stop fluctuate and become stable. First 20 bipartitions were plotted for evaluation (default value). As a result of log likelihood and bipartition plots examination the burn-in period was set as 25% of all sampled records, *i.e.* the first 25% of all sampled trees were discarded, and the remaining were used to calculate posterior probabilities.

Model parameter summary over four independent runs is given in Table 15. For each parameter its value is given as Mean, Variance, Median and its 95% confidence interval. These values are summarised over four independent runs within each analysis, summary statistics from the files produced during analysis and after the burn-in period is specified.

Table 15 Model parameter summary over all 4 runs for non-structural alignment data set: the total tree length (TL), the six reversible substitution rates (r(A<->C), r(A<->G), etc), the four stationary state frequencies (pi(A), pi(C), etc), the shape of the gamma distribution of rate variation across sites (alpha), and the proportion of invariable sites (pinvar). Additionally the lower and upper boundaries of the 95% credibility interval are given. Symbols here and in all other parameter summary tables are taken directly from MrBayes.

			95% Conf. Interval		
Parameter	Mean	Variance	Lower	Upper	Median
TL	4.955685	0.069333	4.469000	5.499000	4.945000
r(A<->C)	0.102163	0.000056	0.087866	0.117254	0.102042
r(A<->G)	0.184880	0.000103	0.165245	0.205224	0.184747
r(A<->T)	0.161049	0.000098	0.142190	0.181209	0.160808
r(C<->G)	0.097803	0.000045	0.085132	0.111244	0.097713
r(C<->T)	0.347261	0.000189	0.320683	0.374683	0.347237
r(G<->T)	0.106845	0.000054	0.092997	0.121669	0.106644
pi(A)	0.215647	0.000040	0.203507	0.228253	0.215580
pi(C)	0.266185	0.000043	0.253452	0.279177	0.266170
pi(G)	0.302030	0.000051	0.288285	0.316155	0.301961
pi(T)	0.216138	0.000036	0.204550	0.227973	0.216128
Alpha	0.513605	0.002054	0.432240	0.609963	0.510799
Pinvar	0.168987	0.000889	0.108175	0.225333	0.169712

This evaluation was done in order to assess where exactly the *Alcyonidium* sequences would fit in relation to other species and especially to other Ctenostomata species. The cladogram with the results is given on Figure 45 , the outgroup was

placed at *Neocrania anomala*, the tree is rooted via *Neocrania anomala* and *Neocrania huttoni* branch. The decision about the outgroup selection was made based on the sequences located in the European Ribosomal Database during the selection of the sequences which had RNA secondary structure already recorded, i.e. the same sequences which were used for reference alignment. Several sister taxa from Entoprocta and Brachiopoda—the latter a lophophorate—were selected. Adding several sister taxa to the outgroup is beneficial for the tree topology as it adds balance and aids in breaking possible long branch attraction (Smith 1994).

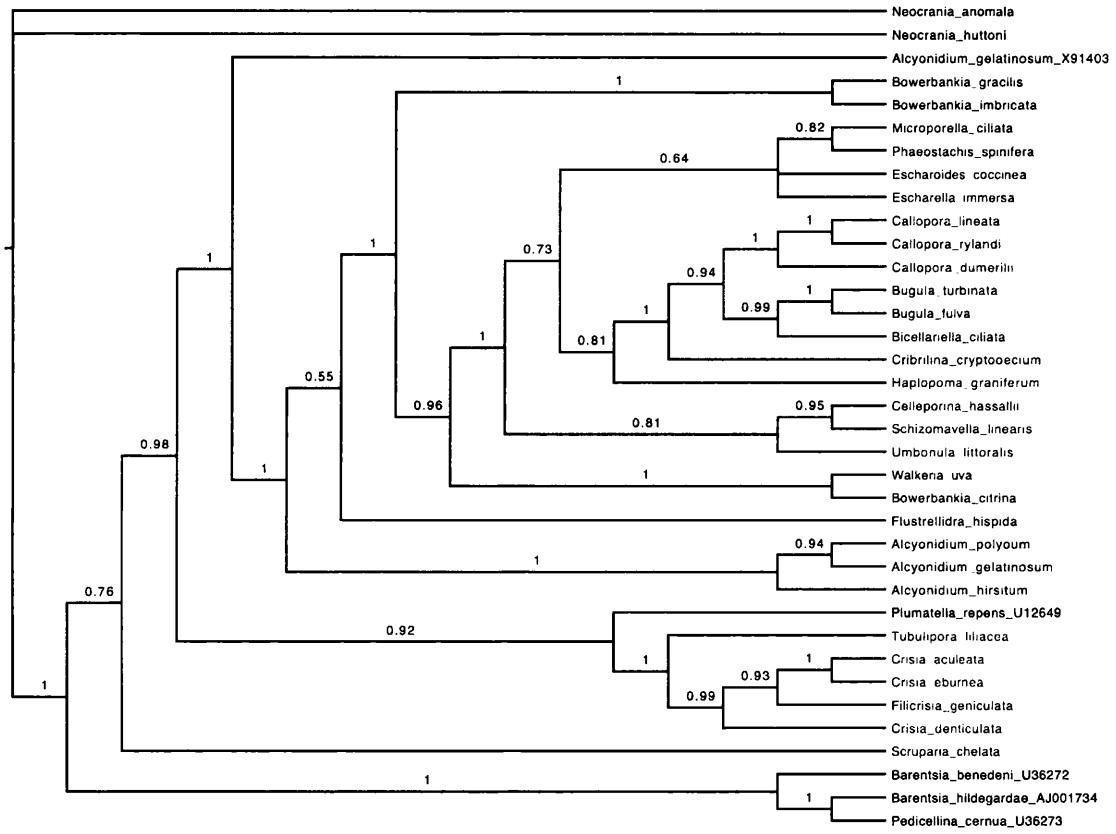


Figure 45 A Bayesian 50% majority rule consensus tree showing results of the software-aligned (MAFFT) data set. All sequences except *Bugula plumosa* are present. Node labels indicate posterior probabilities; sequences from NCBI have their accession number after the species names. All bryozoan species are coloured by their order: Cheilostomata - green; Ctenostomata – Blue; Cyclostomata – red; Phylactolaemata and the outgroup – black.

5.7.2 Structural alignment (GTR+I+Γ model)

The results given here are for the analysis of structure-aligned sequences without consideration of *Alcyonidium* species obtained in this work. All analyses were run with four independent runs each having four chains. The analysis was run for 20,000,000 generations with sample frequency every 1000th generation, thus

recording in total 20,000 trees. The burn-in period was determined using log likelihood plots and graphical results from AWTY (Wilgenbusch *et al.* 2004) which showed number of generations before 20 most variable bipartitions (taxon splits) became stable. As a result the burn-in period was set to 25% *i.e.* 5000 trees were discarded. The convergence was checked against average standard deviation of split frequencies, uncorrected potential scale reduction factor (Gelman and Rubin 1992) for all model parameters combined through four independent runs, and finally, by visually examining AWTY output plot of posterior clade probabilities as function of chain length. Table 16 shows model parameter summaries over all four runs.

Table 16 Model parameter summary for GTR only model for structure-based dataset over all 4 runs: the total tree length (TL), the six reversible substitution rates (r(A<->C), r(A<->G), etc), the four stationary state frequencies (pi(A), pi(C), etc), the shape of the gamma distribution of rate variation across sites (alpha), and the proportion of invariable sites (pinvar).

			95% Cred. Interval		
Parameter	Mean	Variance	Lower	Upper	Median
TL	2.733229	0.021532	2.464000	3.039000	2.727000
r(A<->C)	0.103030	0.000074	0.086655	0.120455	0.102772
r(A<->G)	0.161984	0.000115	0.141583	0.183456	0.161801
r(A<->T)	0.152252	0.000118	0.131647	0.174405	0.152013
r(C<->G)	0.119931	0.000075	0.103537	0.137440	0.119727
r(C<->T)	0.356861	0.000242	0.326955	0.387744	0.356688
r(G<->T)	0.105942	0.000067	0.090409	0.122452	0.105776
pi(A)	0.220336	0.000058	0.205650	0.235532	0.220218
pi(C)	0.253369	0.000057	0.238696	0.268330	0.253304
pi(G)	0.301931	0.000073	0.285364	0.318708	0.301833
pi(T)	0.224363	0.000054	0.210286	0.238967	0.224269
Alpha	0.697977	0.006803	0.554188	0.875984	0.691646
Pinvar	0.350345	0.000725	0.295087	0.400925	0.351183

Figure 46 shows a 50% consensus rule cladogram, with outgroup placed at *Neocrania anomala*. The tree is rooted via the *Neocrania anomala* and *Neocrania huttoni* clade. Because of the concerns over the credibility of *Alcyonidium gelatinosum* (accession no. X9140) sequence (Dr J. Porter, personal communication) and because of the unusual grouping of *A. gelatinosum* sequence in relation to other Ctenostomata the above analysis was repeated with *A. gelatinosum* sequence excluded from the

analysis. The parameters of the MrBayes run were the same as with the previous run and so were the methods of convergence assessment. The resulting cladogram is presented on Figure 47.

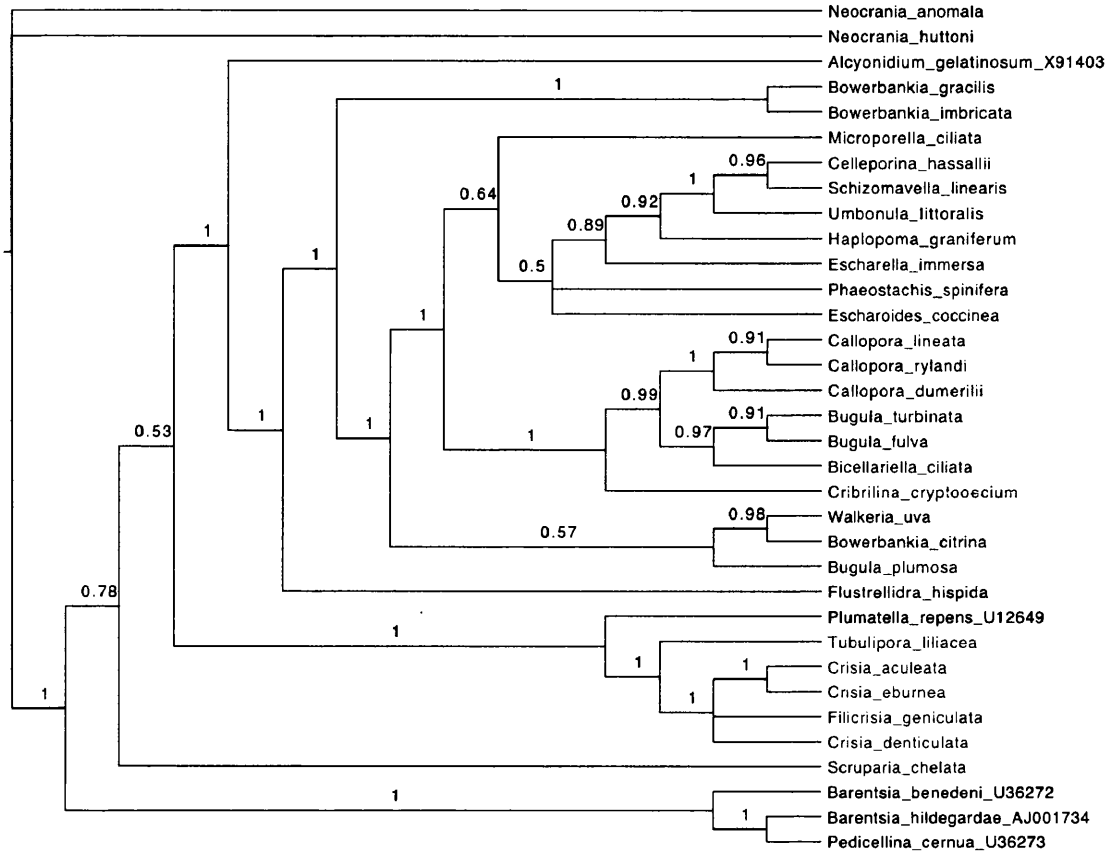


Figure 46 A Bayesian 50% majority rule consensus tree, showing results from the structure-based alignment; model GTR+I+Γ; Both *Alcyonidium gelatinosum* and *Bugula plumosa* sequences are present. Node labels indicate posterior probabilities; sequences from NCBI have their accession number after the species names. All bryozoan species are coloured by their order: Cheilostomata - green; Ctenostomata – Blue; Cyclostomata – red; Phylactolaemata and the outgroup – black. See text for details.

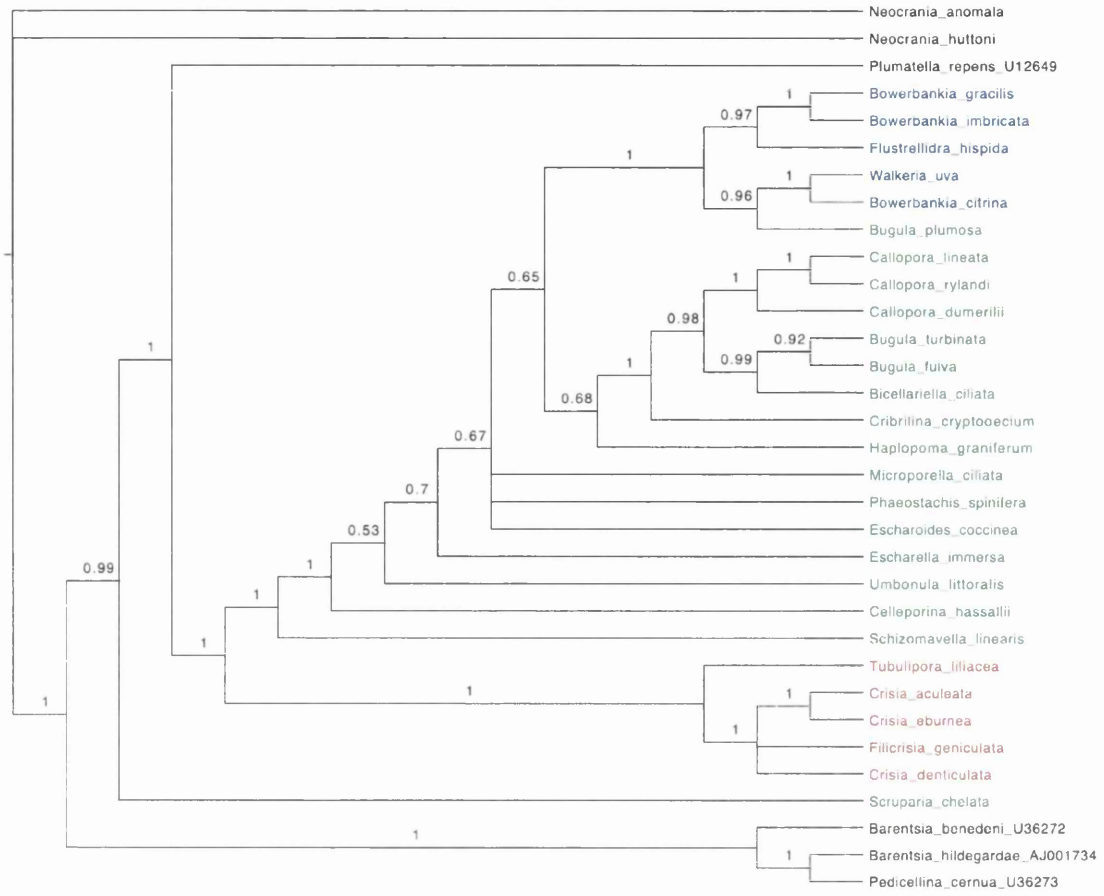


Figure 47 A Bayesian 50% majority rule consensus tree, showing results from the structure-based alignment; model GTR+I+Γ; *Bugula plumosa* sequence is present. Node labels indicate posterior probabilities; sequences from NCBI have their accession number after the species names. All bryozoan species are coloured by their order: Cheilostomata - green; Ctenostomata – Blue; Cyclostomata – red; Phylactolaemata and the outgroup – black. See text for details.

5.7.3 *Bugula plumosa* sequence

On a close examination of the trees it was discovered that the sequence of *Bugula plumosa* (obtained here), appeared to be persistently clustering with Ctenostomata (specifically with Vesicularidae), although other Bugulidae appeared to cluster with each other and with the "correct" order (Cheilostomata), see cladogram in Figure 46 for example. Because the above clustering made no taxonomic sense and because other *Bugula* sequences obtained here did not show any abnormality it was decided to exclude this sequence, *Bugula plumosa*, from further analysis. The resulting cladogram is shown in Figure 48. Also for comparison reasons the same cladogram is shown with *Alcyonidium gelatinosum* sequence present (Figure 49); this resulted in breaking of the Ctenostomata clade on the tree (see below for the detailed discussion of the results).

Several reasons could have caused this abnormal grouping but a very likely reason may be mislabelling of the sample during the sequencing process or during the computer post-processing of sequences. This conclusion was reached as the sequence *per se* appears to be a "valid" 18S rRNA sequence – *i.e.* it aligns well to the rest of the sequences obtained here and follows the secondary structure postulated for the rest of the bryozoan sequences, thus it is not likely to be a contaminant. Also other sequences of this genus obtained in this study – *Bugula turbinata* and *Bugula fulva* – cluster well with each other (see for instance Figure 48) and there is little support in taxonomic literature for non-uniformity of this well described genus (see for instance Ryland 1960). Therefore an error rather than a new taxonomic grouping is assumed. It is also possible, although less likely, that the embryo was erroneously mislabelled during the DNA extraction process (see Chapter 3).

In any case the issue with this species can only be fully resolved by sequencing another sample of *Bugula plumosa* using the oligonucleotide primers used for the other *Bugula* species in this study. Because of the above described uncertainty all trees presented further are those with *Bugula plumosa* excluded from the alignment.

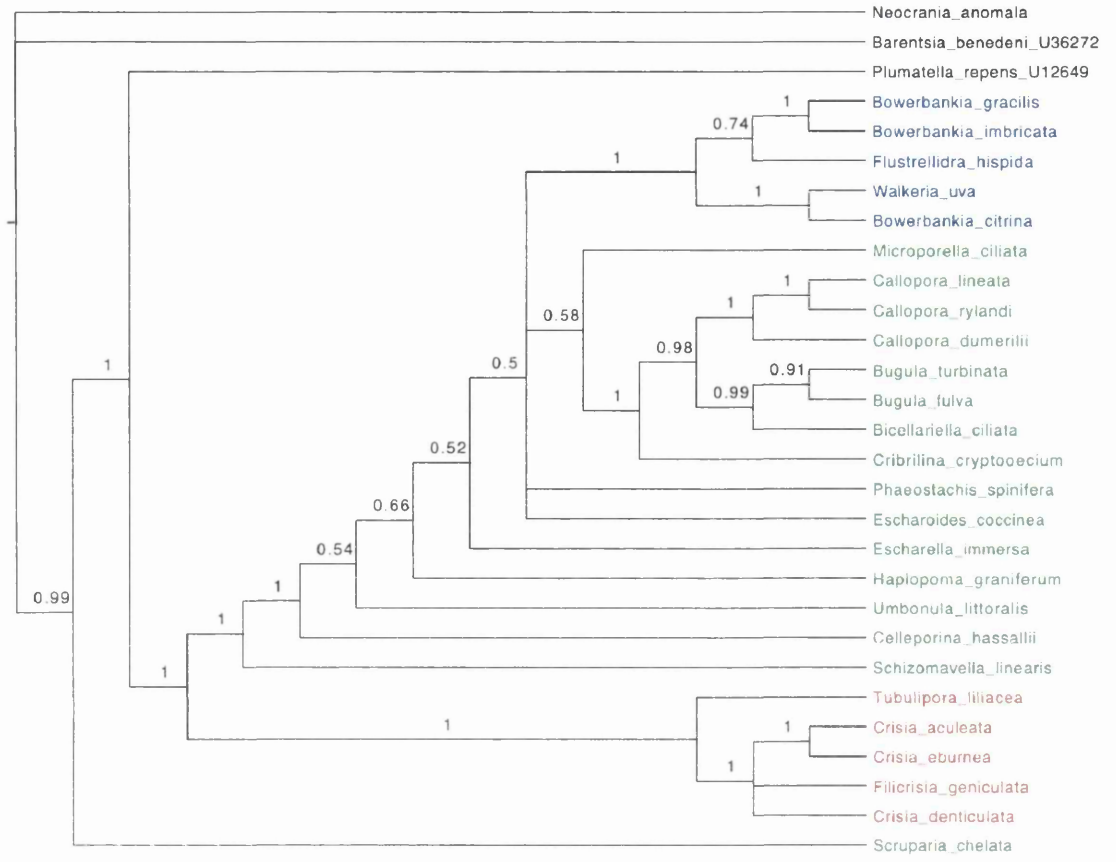


Figure 48 A Bayesian 50% majority rule consensus tree, showing results from the structure-based alignment; model GTR+I+Γ; *Bugula plumosa* and *Alcyonidium gelatinosum* sequences are excluded. Node labels indicate posterior probabilities; sequences from NCBI have their accession number after the species names. All bryozoan species are coloured by their order: Cheilostomata - green; Ctenostomata – Blue; Cyclostomata – red; Phylactolaemata and the outgroup – black. See text for details.

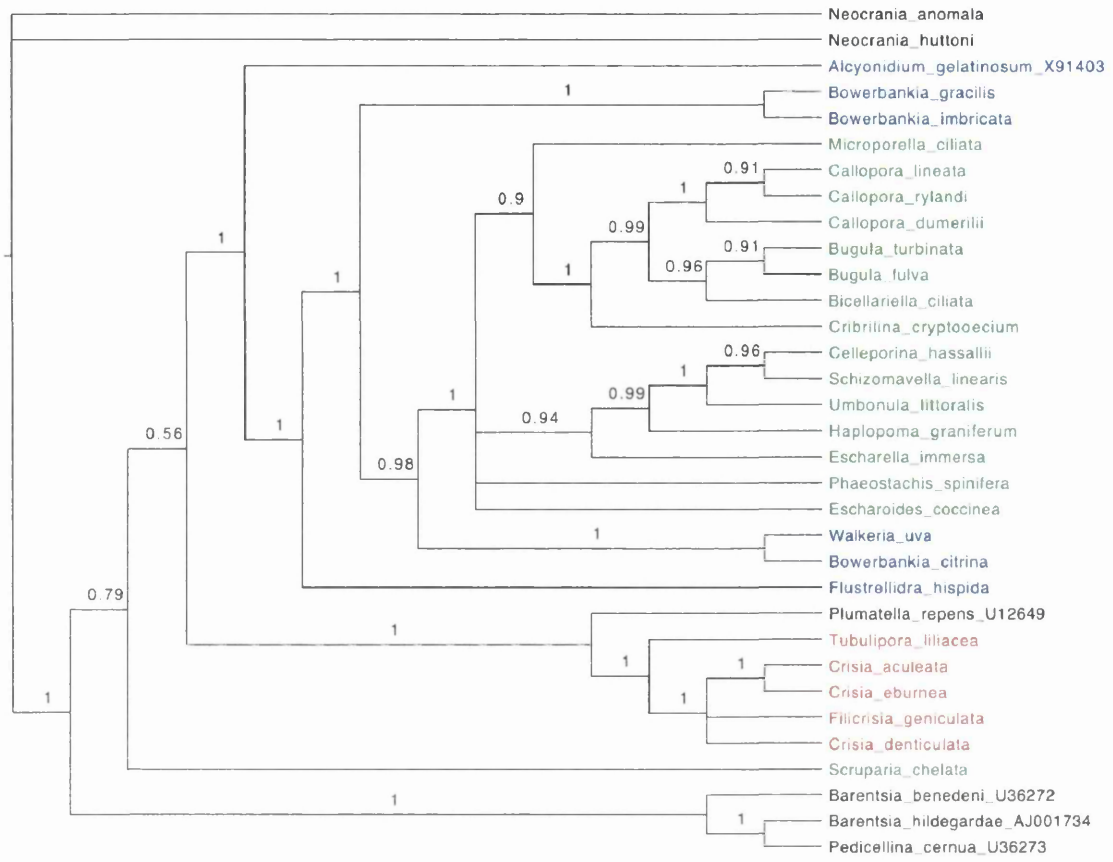


Figure 49 A Bayesian 50% majority rule consensus tree, showing results from the structure-based alignment; model GTR+I+Γ; *Alcyonidium gelatinosum* sequence is present, *Bugula plumosa* excluded. Node labels indicate posterior probabilities; sequences from NCBI have their accession number after the species names. All bryozoan species are coloured by their order: Cheilostomata - green; Ctenostomata – Blue; Cyclostomata – red; Phylactolaemata and the outgroup – black. See text for details.

5.7.4 Structure alignment (GTR+I+Γ and RNA16HKY models)

The results of the structural alignment of the partitioned data are given here. The dataset (i.e. structure-aligned sequences) was partitioned using Xstem utility and formatted for the use in MrBayes as described in Appendix C. The analysis excluded the *Bugula plumosa* sequence and *Alcyonidium gelatinosum* sequence from NCBI (see above). Further, in order to speed up the computation time only three sequences were used as outgroup, *Neocrania anomala*, *Plumatella repens* and *Barentsia benedeni*, with the same rooting as before. Thus only 29 species were used altogether in the dataset. These were the same sequences as those used in the previous alignments. The addition of the RNA16BHKY model made the analysis run much slower and thus required longer computer time – from four days required for the GTR only model it had to be extended up to 11 days. The analysis was run twice – once for

16,000,000 generations, and then it was repeated to run for 22,000,000 generations because the convergence was not reached in the first analysis.

Each analysis was done as before with four individual chains and four separate runs within each analysis. The samples were recorded each 1000th generation, and the burn-in value (16,000 samples) was determined based on the graphical output from log likelihood values and AWTY analysis of variable bipartitions stability. The value of average standard deviation of split frequencies was monitored in each analysis and in the case of 16,000,000 generations it was too high (0.056527) to be accepted, and in the consecutive analysis of 22,000,000 this value dropped to 0.024926, which is below the recommended convergence 0.05 value. The summary of the model parameters is given in Table 17.

Table 17 Model parameter summary for GTR {1} and RNA16BHKY {2} models for structure-based dataset over all 4 runs: the total tree length (TL), the six reversible substitution rates (r(A<->C), r(A<->G), etc), the stationary state frequencies for {1} and {2} models (pi(A), pi(C), etc), the shape of the gamma distribution of rate variation across sites (alpha), and the proportion of invariable sites (pinvar). PSRF: is the convergence diagnostics calculated by MrBayes during the analysis.

			95% Conf. Interval			
Parameter	Mean	Variance	Lower	Upper	Median	PSRF
TL{all}	3.789123	0.056506	3.349000	4.278000	3.780000	1.001
r(A<->C){1}	0.118858	0.000128	0.097237	0.141886	0.118654	1.000
r(A<->G){1}	0.142930	0.000158	0.119567	0.168640	0.142574	1.000
r(A<->T){1}	0.147278	0.000152	0.123793	0.172342	0.147001	1.001
r(C<->G){1}	0.132801	0.000158	0.109436	0.158252	0.132469	1.000
r(C<->T){1}	0.344124	0.000394	0.305583	0.383540	0.343908	1.000
r(G<->T){1}	0.114009	0.000131	0.092607	0.137274	0.113652	1.000
pi(A){1}	0.286500	0.000115	0.265896	0.307537	0.286390	1.000
pi(C){1}	0.215876	0.000080	0.198703	0.233767	0.215823	1.001
pi(G){1}	0.255725	0.000102	0.236184	0.275952	0.255615	1.000
pi(T){1}	0.241900	0.000089	0.223626	0.260893	0.241739	1.000
pi(AA){2}	0.007997	0.000005	0.004296	0.013114	0.007694	1.002
pi(AC){2}	0.013485	0.000006	0.009238	0.019100	0.013296	1.012
pi(AG){2}	0.011857	0.000006	0.007578	0.016721	0.011699	1.011
pi(AT){2}	0.104529	0.000120	0.083536	0.124601	0.104544	1.008
pi(CA){2}	0.015986	0.000007	0.010974	0.021714	0.015823	1.006
pi(CC){2}	0.016986	0.000007	0.012313	0.022983	0.016819	1.003
pi(CG){2}	0.257717	0.000223	0.230234	0.287040	0.257786	1.028

pi(CT){2}	0.015391	0.000006	0.011229	0.020642	0.015024	1.008
pi(GA){2}	0.008697	0.000004	0.005240	0.012820	0.008719	1.002
pi(GC){2}	0.260741	0.000214	0.233006	0.287504	0.259958	1.011
pi(GG){2}	0.025621	0.000012	0.019508	0.033464	0.025682	1.038
pi(GT){2}	0.057039	0.000036	0.045359	0.068345	0.057326	1.013
pi(TA){2}	0.123977	0.000135	0.100641	0.147557	0.124251	1.001
pi(TC){2}	0.012183	0.000005	0.008677	0.017365	0.012028	1.002
pi(TG){2}	0.055199	0.000034	0.044910	0.067146	0.055000	1.009
pi(TT){2}	0.012594	0.000008	0.008448	0.018767	0.012280	1.001
alpha{1}	0.532890	0.004902	0.414305	0.687575	0.526775	1.000
alpha{2}	0.585645	0.009528	0.437207	0.813493	0.570243	1.001
pinvar{1}	0.284400	0.001561	0.202239	0.357375	0.285807	1.000
pinvar{2}	0.207249	0.001776	0.123327	0.288157	0.207709	1.002

In addition to the model parameters in the above table also a convergence diagnostic is given – Potential Scale Reduction Factor (PSRF). This diagnostic is calculated by MrBayes and should approach 1.00 as the runs converge (Ronquist and Huelsenbeck 2003). These values can be compared to the equivalent values in other analyses (see the model summary tables). Notably, the above analysis results based on the convergence diagnostics could be considered "satisfactory', but ideally the analysis could have been run for more generations, this however was not possible due to allocated time for the use of the cluster. The resulting cladogram is presented on Figure 50.

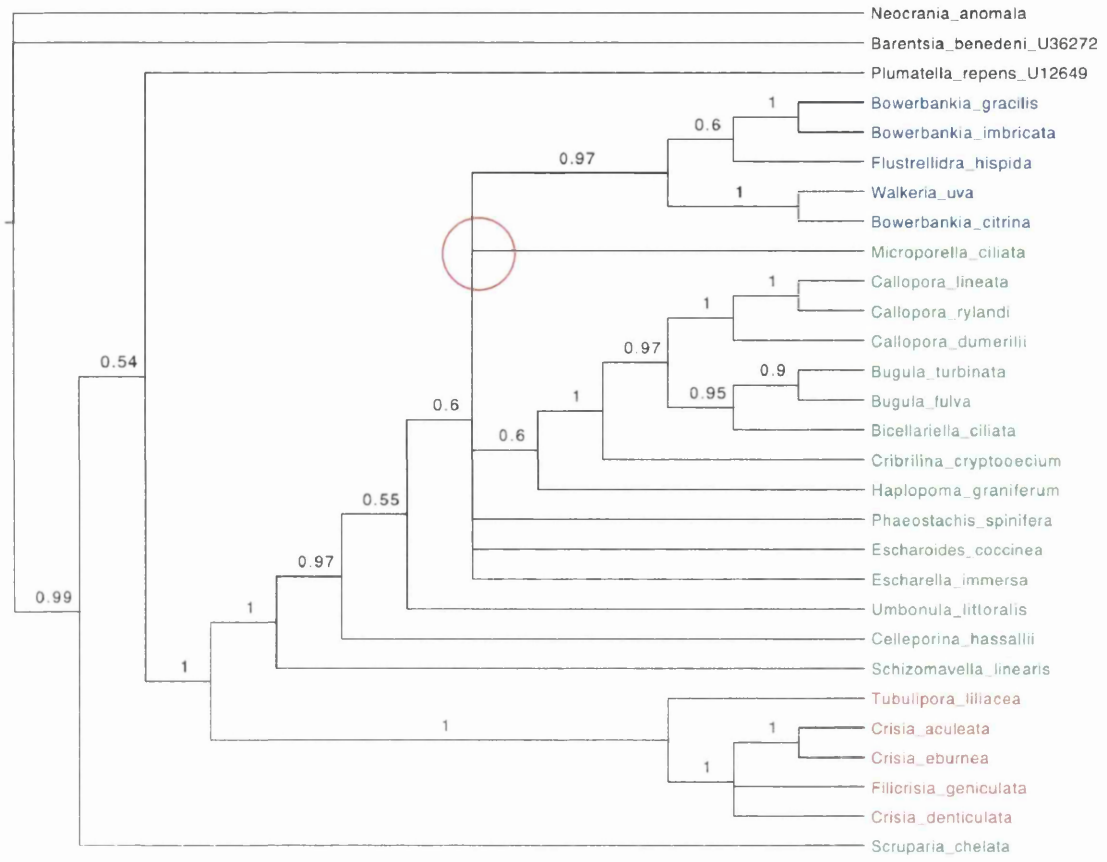


Figure 50 A MrBayes 50% majority rule consensus tree, showing results from the structure-based alignment with partitioned data set and two models GTR+I+Γ and RNA16BHKY+I+Γ. *Bugula plumosa* and *Alcyonidium gelatinosum* sequences are excluded. Node labels indicate posterior probabilities; sequences from NCBI have their accession number after the species names. All bryozoan species are coloured by their order: Cheilostomata - green; Ctenostomata – Blue; Cyclostomata – red; Phylactolaemata and the outgroup – black. Red circle indicates unresolved ctenostome-cheilostome grouping—see text for details

From the tree on Figure 50 it can be seen that *Scruparia chelata* sequences are positioned at the root of Bryozoa (including the fresh water phylactolaemate *Plumatella repens*) however in the tree generated by the shorter run (*i.e.* 16,0000,000 generations) the position of the this species is changed – the resulting cladogram is shown on Figure 51, notably the rest of the tree appears to be identical to the longer run, also the posterior probabilities of the clades are very similar.

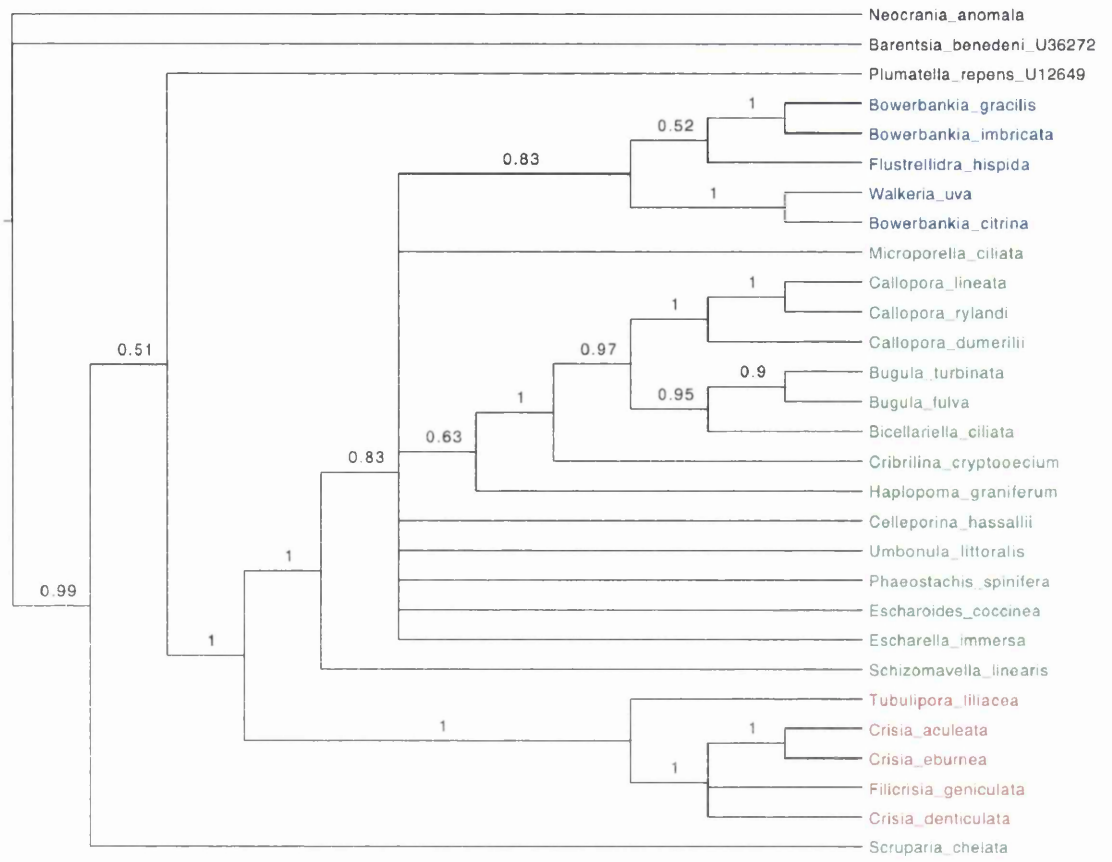


Figure 51 A MrBayes 50% majority rule consensus tree, showing results from the structure-based alignment with partitioned data set and two models GTR+I+Γ and RNA16BHKY+I+Γ. This analysis was run for 16 mln generations. *Bugula plumosa* and *Alcyonidium gelatinosum* sequences are excluded. Node labels indicate posterior probabilities; sequences from NCBI have their accession number after the species names. All bryozoan species are coloured by their order: Cheilostomata - green; Ctenostomata – Blue; Cyclostomata – red; Phylactolaemata and the outgroup – black. See text for details.

For comparison reasons the same analysis for the partitioned data was repeated using all species include in the original alignment (i.e. 33 species in total). This alignment included *Bugula plumosa*, and *Alcyonidium gelatinosum* sequences, whose validity was not certain. The results of the analysis are given here because unexpectedly, this analysis even though containing more species took less time to the convergence. The analysis was run for 22,000,000 generations, sampled at every 1000th generation and the burn-in value determined exactly as above. Based on the AWTY and log likelihood plots evaluation the burn-in value was set to 4000. The resulting cladogram is presented on Figure 52.

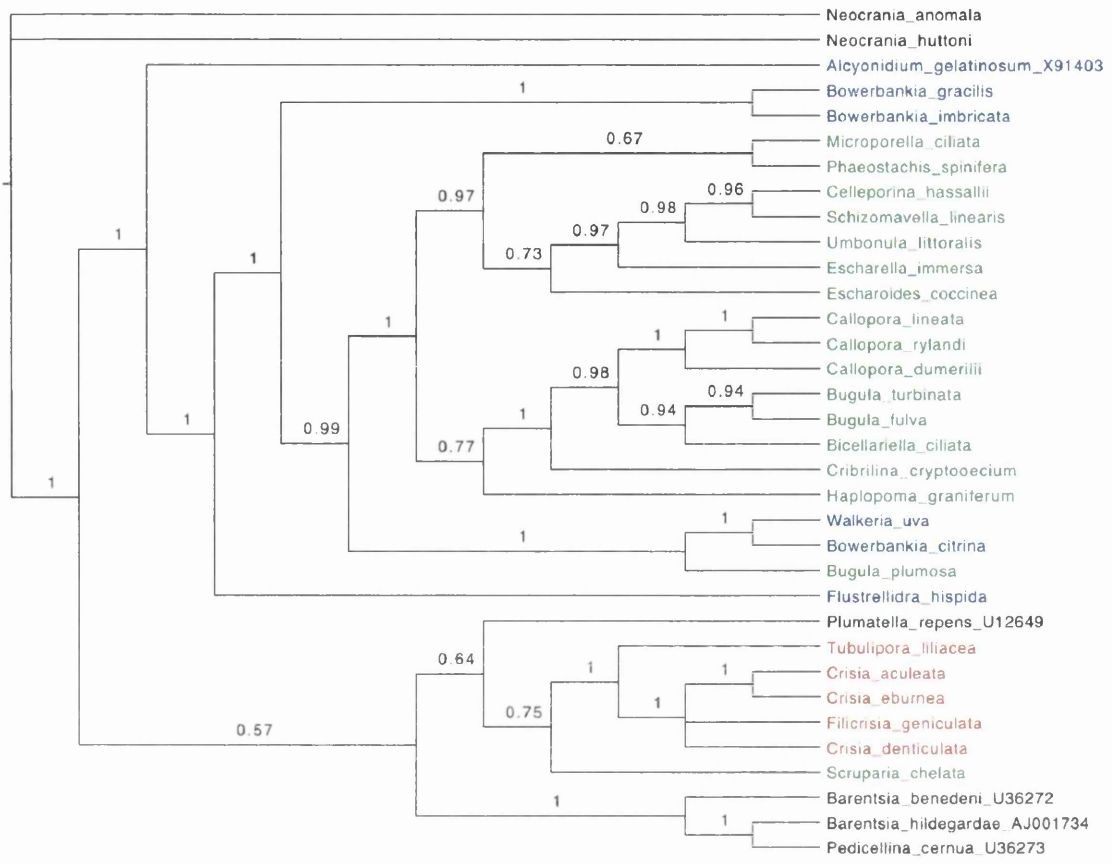


Figure 52 A MrBayes 50% majority rule consensus tree, showing results from the structure-based alignment with partitioned data set and two models GTR+I+Γ and RNA16BHKY+I+Γ. *Bugula plumosa* and *Alcyonidium gelatinosum* sequences are included. Node labels indicate posterior probabilities; sequences from NCBI have their accession number after the species names. All bryozoan species are coloured by their order: Cheilostomata - green; Ctenostomata – Blue; Cyclostomata – red; Phylactolaemata and the outgroup – black. See text for details.

5.8 Results discussions

In general, during the analysis, a total of 2042 characters were used in the data matrix with 607 unique characters for the loop partition and 258 unique characters for the stem partition. In non-partitioned dataset, where only the GTR model was used, the number of unique sites was 886, with a total character number of 1948. The results presented above showed some diversity based on the sequences included in the analysis and the method and model used.

In the outgroup there is a strong support for all three entoprocts grouping together, when present: two *Barentsia* species and *Pedicellina cernua*. Interestingly though, instead of grouping together of two *Barentsia* spp., there was an opposite situation observed in all cases with 100% support - *Barentsia hildegardae* (accession no: AJ001734) grouped with *Pedicellina cernua* (accession no: U36273), and

Barentsia benedeni (accession no: U36272) placed as a sister taxon to that clade, which is not what would be expected from two species from the same genus—obviously these three sequences need to be considered with some caution if used in any further analysis or some reconsideration has to be given to the current taxonomic status of *Barentsia* and *Pedicellina*.

The only phylactolaemate bryozoan, *Plumatella repens* (accession no. U12649) in the tree was positioned as a sister taxon to Cyclostomata in both non-partitioned and partitioned dataset (i.e. RNA-specific model did not seem to make much change). However, when *Alcyonidium gelatinosum* sequence was present in the tree, the support was low—0.56 posterior probability. When the *Alcyonidium* sequence was removed from the analysis the position of *Plumatella repens* shifted to the root of the entire bryozoan clade thus becoming a sister group to marine bryozoans (Gymnolaemata and Stenolaemata). In all the trees presented above the outgroup sequences of two brachiopods *Neocrania anomala* and *Neocrana huttoni* were used. In some cases (Figure 48, Figure 50 and Figure 51) only one brachiopod sequence – *Neocrania anomala* – was used to speed up the computation time.

In general, positions of several taxa were changed with the introduction or deletion of the *Alcyonidium* sequence. Given the uncertainty of this sequence, though, it is best to not to make many conclusions based on the changed topology of the trees when *Alcyonidium gelatinosum* is present.

5.8.1 Order Cyclostomata

Cyclostomes showed monophyly in all cases regardless of the model and the dataset used. The entire clade of Cyclostomata was positioned as the sister clade to the rest of the cheilostomes and ctenostomes (note that *Scruparia chelata* position is treated separately below). There was a very high support for this topology—1.00 posterior probability in all cases. Among the cyclostome sequences, family Crisiidae was monophyletic although genus *Crisia* was not resolved fully: *Filicrisia geniculata* sequence showed polytomy with *Crisia denticulata* with remaining two *Crisia* spp. fully resolved. *Tubulipora liliacea*, belonging to family Tubuliporidae, was at the root of the clade as a sister group.

5.8.2 Order Ctenostomata

Ctenostomes showed probably most differences based on the sequences used for the alignment. The main difference (which could be seen if comparing Figure 52 to Figure 50) is that without the *Alcyonidium gelatinosum* sequence (Figure 50) ctenostomes were a monophyletic clade positioned within the paraphyletic cheilostome assemblage. Introduction of the *Alcyonidium* sequence (Figure 52) resulted in ctenostomes becoming polyphyletic. The choice of model affected the posterior probabilities of the ctenostome clade when a partitioned dataset was used with both RNA16HKY and GTR+I+Γ models it showed slightly lower support for the clade 0.97 — rather than 1.00 for a non-partitioned dataset with GTR+I+Γ only model. However, as noted above, there was a clear convergence problem with the RNA-model analysis and if the analysis were to run any longer it is possible that the support for the clade would change. Within the ctenostomes themselves there was a slightly lower support for families: three *Bowerbankia* sequences did not form a clade (Figure 50); in the *Flustrellidra hispida* sequence grouping with two *Bowerbankia* sequences—the resulting support for this grouping was low 0.74 and 0.60 for the unpartitioned and partitioned models respectively. Consideration of the position of this group and the within group relationship definitely requires more sequences especially those belonging to Alcyonidiidae.

5.8.3 Order Cheilostomata

The situation with cheilostomes is much less certain. In general, from the analysis it is clear that they are paraphyletic, but their position in relation to the other orders is less certain. Without the *Alcyonidium gelatinosum* sequence they appear to be a paraphyletic sister group to the cyclostomes, and contain monophyletic ctenostomes within. However, the introduction of the *Alcyonidium gelatinosum* sequence breaks this assemblage and also makes ctenostomes monophyletic. As discussed above the validity of *Alcyonidium gelatinosum* sequence (accession no. X91403) from NCBI is dubious, however if we were to consider the tree that was based on the software alignment and included three chimeric *Alcyonidium* sequences obtained here they also appear to cluster within the ctenostome-cheilostome clade (see Figure 45). Therefore, assuming that at least parts of the chimeric *Alcyonidium* sequences obtained here are

correct, they show some indication on where the rest of the valid *Alcyonidium* sequences would be if they were present.

It is clear that more sequences from ctenostomes are required—especially of Alcyonidiidae—to clarify the position of cheilostomes and ctenostomes. Some of the cheilostomes on the RNA-model tree (see Figure 50) were not resolved very well and were polytomic (in the case of the tree on Figure 50 it was simply because it was a 50% consensus tree and clades with lower posterior probabilities were not individually resolved). This pattern was repeated for both non-partitioned model analysis (Figure 48) and for RNA-model partitioned dataset analysis (Figure 50). However, when the *Alcyonidium gelatinosum* sequence was added (Figure 49) an additional clade appeared, which grouped together five representatives from different families: *Schizomavella linearis* (from the family Bitectiporidae), *Celleporina hassallii* (from the family Celleporidae), *Umbonula littoralis* (from the family Umbonulinidae), *Haplopoma graniferum* (from the family Haplopomidae) and *Escharella immersa* (from the family Romancheinidae). Yet another representative of the family Romancheinidae – *Escharoides coccinea* – was not part of this clade.

5.8.4 Anasca

One interesting finding is that within cheilostomes the Anasca group (represented here by *Callopora lineata Callopora rylandi, Callopora dumerilii, Bugula turbinata, Bugula fulva* and *Bicellariella ciliata*) was monophyletic in all trees and showed very strong support for this clade—this can be seen for instance from Figure 50. The grouping support was equally strong for all model types used here, and was recovered on all trees. Malacostegoidea and Cellularioidea are thus monophyletic sister taxa of a larger anascan monophyletic clade. However, *Scruparia chelata* which is also currently placed within the anascan group was not part of this clade nor was it included in the Cheilostomata in any trees. See below a separate treatment of this species.

5.8.5 Scruparia chelata

One cheilostome sequence, which has so far been neglected in discussion here, is that of *Scruparia chelata.* This species belongs to the family Scruparidae, order Cheilostomata and until recently was the only genus of this family. However recently another species (from Antarctica) was added to this family—*Brettiopsis triplex*—

based on the similarities of the brood chamber of the brooding zooids (Gappa 1986). Scrupariidae are generally considered by some (*e.g.* P.J.Hayward, personal communication) to be the most primitive anascan cheilostome Bryozoa, with very little information available about the species and their reproductive cycle and larvae (Ryland and Hayward 1977; Zimmer and Woollacott 1977; Dr PJ Hayward personal communication). In the trees presented here *Scruparia chelata* appeared placed at the root of the entire bryozoan tree—including the phylactolaemate *Plumatella repens*—with very high posterior probabilities for both the non-partitioned dataset with the GTR-only model and for the partitioned analysis using the RNA model.

Phylogenetic findings in this work showed considerable difference from the previous molecular phylogenetic study (Dick 2000) where 16S rRNA was used to reconstruct gymnolaemate Bryozoa.

The position of cyclostomes in this study, which were monophyletic, contradicts Dick *et al.* (2000) work where they were shown to be polyphyletic. Results here show support for those assumptions proposed by Todd (2000) that cyclostomes should be a monophyletic clade. Interestingly though, there appears to be no consensus in the literature on this matter as some (Taylor and Larwood 1990) believe that this group is paraphyletic when fossil stenolaemates considered.

Contrary to the common assumptions, stenolaemates here are a sister group to gymnolaemates, whereas stenolaemates are generally considered to have been derived from the ctenostomes (Larwood and Taylor 1979; Todd 2000) and cyclostomes to be paraphyletic or even polyphyletic (Todd 2000).

This work partially supports the finding of Dick *et al.* (2000) that cheilostomes are polyphyletic, if the sequence of *Scruparia chelata* is taken into consideration. It also was in concordance with the view of Todd (2000) who believed that many cheilostome families are paraphyletic and makes further sense given the great disparity of the group based on morphology of zooids, larvae and colony in general (Gordon 2000). Here Anasca were always recovered as a monophyletic clade with good support for genera—all sequences included in the analysis belonged to the suborder Neocheilostomina.

5.8.6 Ascophora

The polyphyly of ascophorans (assuming a common ctenostome ancestor) appears to be coherent with other findings based on the differences of evolutionary models of

frontal shields (Gordon 2000) and their ontogeny and structure (Voigt 1991). Cheilostomata as a group are still shown to be paraphyletic, which defies a common concept of monophyletic Cheilostomata (Todd 2000; Gordon 2000; Taylor and Larwood 2000). However the current higher taxa grouping of Cheilostomata is based on the morphology and structure of the frontal wall only (Gordon 2000) and thus a possibility of homoplasy has to be evaluated with more molecular data as it becomes available. When more apparent resolution was shown with the addition of the *Alcyonidium gelatinosum* sequence, the grouping of several species with very high posterior probability support still did not recover any expected taxonomic grades within Ascophora. For instance, Lepraliomorpha (represented here by *Celleporina, Schizomavella* and *Phaeostachys*) were still paraphyletic—see Figure 46 and Figure 52.

Ctenostomata, which are regarded as a paraphyletic group that has arisen from a common ancestor with Cheilostomata (Taylor and Larwood 2000; Ryland 1970; Gordon 2000) here showed different results. Ctenostomes formed a polyphyletic group with cheilostomes being monophyletic within the larger ctenostomes-cheilostome assemblage (Figure 52), or a monophyletic sister group to the paraphyletic cheilostomes if *Alcyonidium gelatinosum* was removed from the tree analysis. Because of unresolved polytomy (see a circle mark on the Figure 50) between ctenostomes and cheilostomes (posterior probability below 0.5) this grouping is highly uncertain and requires further investigation.

The effect of the presence or absence of the *Alcyonidium gelatinosum* (obtained from NCBI) sequence from the analysis caused a dramatic increase of convergence time and in some cases even 22,000,000 generations was not sufficient to achieve convergence (data not shown). For instance during the evaluation of average standard deviation (ASD) which proved to be a very good guide for a quick convergence diagnostic, removal of *Alcyonidium gelatinosum* sequence from the alignment increased ASD value from 0.003634 (which in this case indicated convergence) for 20,000,000 generations to 0.03778 for the same number of generations. Further, removal of the *Bugula plumosa* sequence increased the ASD value even more to 0.0607 for the same number of generations. This dramatic change of convergence time could not be explained – normally an opposite effect would be expected when sequences are removed from the alignment *i.e.* faster calculation speeds and a shorter time required for reaching the stationarity. Although exclusion of

Bugula plumosa was necessary as this sequence was mislabelled earlier in the analysis, this sequence is valid *per se*—that is it can be treated as if sequenced from an unknown bryozoan, but still belonging to the Bryozoa. If we were to treat it that way two obvious explanations of the grouping of this sequence with ctenostomes can be given. First, and most likely, is that it indeed belongs to ctenostomes and not to cheilostomes and is thus mislabelled; alternatively, cheilostomes are polyphyletic, which is less likely solely because of this particular sequence.

In general, addition of the *Alcyonidium* sequence breaks down the monophyletic topology of Ctenostomata but has little or no effect on the other two orders Cheilostomata and Cyclostomata. It is possible therefore that if more ctenostome sequences were to be added, especially those belonging to Alcyonidiidae, some clarification of the position of Ctenostomata could be achieved. In fact the addition of more sequences to the analysis was shown to increase resolution of the trees (Poe 1998; Hillis *et al.* 2003), but also some authors suggested that adding additional characters to the same number of species could also improve resolution (Poe and Swofford 1999). Obviously given the lack of success with *Alcyonidium* sequences obtained here this question remains to be answered, but it is clear that for such a diverse group more sequences are required as well as possibly additional genes.

The validity of some sequences submitted to the NCBI is questionable, for instance during the primer design stage in this work several bryozoan sequences (already available on NCBI) were evaluated and when a simple NJ tree was built these sequences clustered not by the taxonomic group they belonged to, but instead formed aggregations based on the author who submitted the sequences (tree not shown here). The validity of these sequences was also questioned base on the method with which DNA was obtained from the specimens (Dr J.Porter, personal communication).

6 CONFOCAL MICROSCOPY OF BRYOZOAN LARVAE

6.1 Evaluation of CLSM method for larval imaging

In this chapter a method developed for describing larval morphological characters using a confocal laser microscopy (CLM) is presented. It is hoped that the use of confocal laser microscope and fluorochromes tested here and possible other ones presents a method which will allow relatively rapid evaluation of larval types of Bryozoa and their morphological characters. Although it was not possible to evaluate many larval types in this study, the method described here, could be employed to a wider survey of larval types especially for those bryozoans for which larval types are unknown.

A great part of taxonomic classification of Bryozoa is based on the structure of zooecium (or cystid) and its function. In particular, the classification of many Cheilostomata is based on the frontal wall or shield (Gordon 2000). This system is extensively used for both extant and fossil species as the calcified skeletons of Bryozoa are well preserved in fossil record of this phylum dating back to Ordovician. Although larval morphology is not widely examined, it is of a great importance both for the evolution of Bryozoa, such as the switch between planktotrophy to non-planktotrophy possibly in the Ordovician firstly, and may play an important role in the taxonomic classification of extant taxa and establishing the evolutionary traits of the larvae (Taylor and Larwood 1990, Santagata and Zimmer 2000). Thus, knowledge of larval types could answer many question of the relationship between different taxonomic groups within Bryozoa and eventually give some important addition to understanding the major steps in the evolution of the group.

Scanning electron microscopy (SEM) plays a very important role in bryozoan taxonomy and morphology. This method has seen a wide use for numerous techniques and analyses such as the zooecial external morphology, skeletal microstructure, resin casting of fossil Bryozoa and for the morphology of bryozoan larvae (Taylor 1990).

Impact of SEM on larval morphology is hard to underestimate – apart from the legendary drawings of Borrois (1877) which are still used in many studies as the reference to which bryozoan larvae are compared, the majority of other larval images are indeed taken using SEM. Examples of many works dedicated to the study of individual larvae are ample ((Reed 1977; Reed and Cloney 1982; Reed and Woollacott 1982; Reed and Woollacott 1983; Reed 1988; Reed, Ninos *et al.* 1988). However the information collected using SEM is limited to a few species and their detailed morphology. No method so far has been used which could allow relatively rapid way of assessing and identifying larval types (such as those designated by Zimmer and Woolacott system). The obvious advantage of such a method would be to perform a survey of species from many bryozoan families and identify their larval types so that some systematic information could be obtained from them.

Traditional light field microscopy in combination with histological techniques has been successfully used for the study of the morphology of bryozoan larvae (Ryland 1970; Reed 1977; Reed and Cloney 1982; Reed and Woollacott 1982; Reed 1988; Reed *et al.* 1988). This method is well tested and gives good results, however, the limitations of it is that it fails to show the surface elements of the larvae, in addition it is comparatively time consuming, especially if there is a need to review a large amount of material. Based on the data obtained from the above methods (SEM and light microscopy) a system of larval type classification and general larval morphology was proposed by Zimmer and Woollacott (1977).

Although larval morphology was covered by many separate articles (Reed 1977; Zimmer and Woolacott 1977; Reed and Cloney 1982; Reed and Woollacott 1982; Reed and Woollacott 1983; Reed 1988; Reed *et al.* 1988; Stricker *et al.* 1988a; Stricker *et al.* 1988b; Zimmer and Woollacott 1989a; Zimmer and Woollacott 1989b; Reed 1991; Zimmer and Woollacott 1993; Okano *et al.* 1996) one of the most complete treatments of the issues was given in two works (Zimmer and Woolacott 1977; Giese *et al.* 1991). While it is not the aim of this study to give a full account of the morphology of the gymnolaemate larvae, some aspects of it are crucial to the understanding of differentiation between different types. Therefore main morphological characters have to be emphasised as they play an important role in differentiating species and types of larvae. The above studies revealed a great diversity in the morphology of the Gymnolaemate larvae. As a result certain external and internal characters are used in the system reviewed below. Figure 53

gives an outlined view of the hypothetical larvae, and lists major characters, which were used by Zimmer and Woollacott (1977) in their work as described herein.

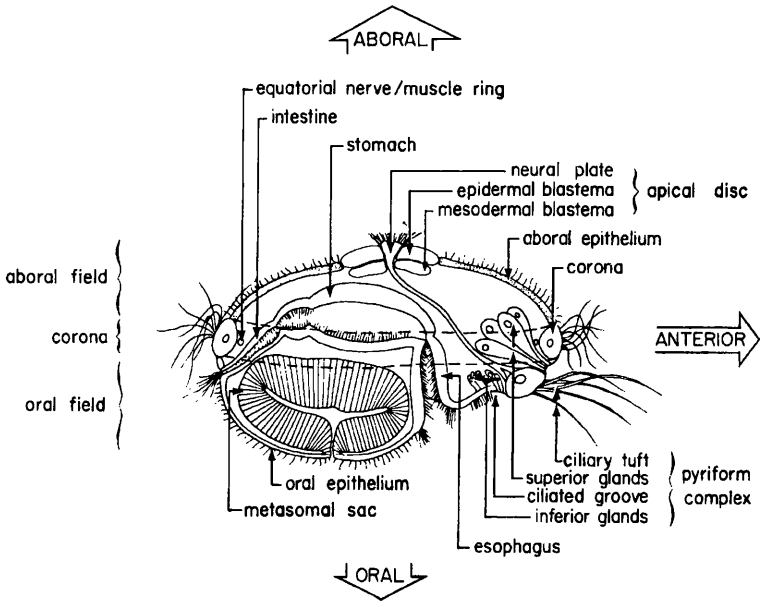


Figure 53 Generalised larva and its main morphological characters. From Zimmer and Woollacott (1977), Fig 1.

Below a brief description of the main morphological characters of the larvae are given. In addition a brief description of the larval types is given in section 6.2 .

6.1.1 Organs of the aboral field

6.1.1.1 Apical disc

This structure is found around the animal pole and could be of various sizes, from a small knob-like to a large size, occupying nearly the entire surface of the aboral field. The central zone of the apical disc is called the neural plate – which unites neuromuscular and other larval nerves with sensory systems of larvae. The peripheral zone of the apical disc is composed of epithelial cells. Epidermal blastema cells sometimes contain microvilli.

6.1.1.2 Aboral epithelium

This zone is sometimes composed of unspecialised cuboidal cells, or can have specialisation depending on the type of the larva, however in shelled larvae this zone is involved in the production of the shell. In coronate larvae (see larval types below, in Table 18) this zone carries a distinctive furrow – called pallial sinus. This

could be of various depths and is used in the system to identify different larvae. There are some indications that the cells lining the pallial sinus could have secretory nature (Zimmer and Woollacott 1977). The differences in the pallial epithelium appear to correspond to its future transformation.

6.1.2 Organs of corona

Corona is a locomotory organ. The number of cells in the corona is 32 in most species, however it could be much larger (more than 300) for example in *Bugula neretina*. For non-cyphonautes species the corona is represented by a complete band, but in cyphonautes it is localised around an inhalant aperture. In all instances corona is formed by only a single layer of cells. The degree of cilia coverage varies among species, as does its location. These characteristics could be used for differentiation of the morphotypes of larvae. The metachronal waves around the corona are responsible for the spiral movement of some larvae (Ryland 1960; Ryland 1970).

6.1.3 Organs of oral field

The oral field corresponds to the former vegetative pole of the embryo. The mouth is present in some types of larvae: cyphonautes, shelled larvae, and some lecithotrophic larvae. Depending on the development of the gut, there is a strongly invaginated vestibule present in those larvae that feed. The oral field also carries the pyriform complex, which sometimes appears to lay in the corona, due to the corona's development.

6.1.3.1 Pyriform complex

This organ is highly noticeable in most larvae, the movement of the ciliary tuft is clearly noticeable, and is composed of a bundle of ciliated cells. The complex also contains glandular inferior and superior "fields", which are cytologically identical and occupy a considerable space within the larval body. The role of the pyriform complex is unclear but a role in movement, or feeding has been suggested (Zimmer and Woollacott 1977).

6.1.3.2 Metasomal sac

This organ, also called the adhesive sac or internal sac, is an invagination of the oral epithelium and present throughout gymnolaemate larvae. The size of the metasomal sac can vary between different groups, from very small, to occupying more than half of the larval interior. It plays a major role in the settlement and metamorphosis of larva into an ancestrula as it becomes everted and the cells release cement, which enables the ancestrula to be anchored to the substratum (Reed 1977; Zimmer and Woollacott 1977; Reed and Cloney 1982; Reed and Woollacott 1983; Reed 1991).

6.1.3.3 Mouth and anus

These are present depending on the development of the larval gut and could be used as one of the important characters for identification of particular groups of larvae.

6.1.3.4 The vestibule

The entrance to the oesophagus is preceded by the vestibule in planktotrophic larvae. It is specialised as a food collecting device. It was shown to be divided into two cavities, distinguished by their function. The separation is done by the means of ciliated ridges. It is notable that coronae of cyphonautes larvae are considerably less developed, as opposed to a uniform corona of non-feeding larvae. This could be used as a differentiation character, for microscopic analysis.

6.1.3.5 The epithelium of the exposed oral field

This is the bordering epithelium between the oral field and the corona, sometimes carrying a glandular tissue. This epithelium is resorbed during ancestrula formation.

6.2 Larval types

Based on the above main features and some particularities of the larval internal morphology a system of several larval types was introduced by Zimmer & Woollacott (1977). This system was based on the principles of gross morphology and only separates seven main types of larvae. It is not attributed to any taxonomic differences, and does not allow separation at a low taxonomic level such as between species for instance. This system is important as it clearly emphasises the characters which could be investigated further for any morphological differences. In addition, it is notable that the system herein exploits both external and internal morphological

differences. Seven main types of the larvae (Table 18) were proposed by the above authors, based on an account of 45 gymnolaemate species and their corresponding larvae.

Table 18 Seven types of Bryozoan larvae as per the system proposed by Zimmer and Woollacott (1977).

Larval Type	Description	Example species
Cyphonautes Larvae	Obligatorily planktotrophic, body compressed bilaterally, lateral surfaces of the aboral epithelium produce chitinous shells. The oral field is deeply invaginated, producing a conical vestibule. Mouth and anus are present, as well as gut and a fully functional digestive system. Corona does not form a uniform ring – it is interrupted into pre- and postoral bands. Metasomal sac is small, situated between mouth and anus.	*Electra pilosa; Membranipora membranacea; Tendra repiachowi; Pyripora catenularia; Alcyonidium* spp.
Shelled Lecithotrophic Larvae	Larvae slightly compressed bilaterally have short oral-aboral and long anterior-posterior axes. The shells are rectangular. Metasomal sac is extensive. Gut is present, however, incomplete posteriorly, and not functional.	*Flustrellidra hispida; Pherusella tabulosa.*
Type O Coronate Larvae	Coronate larvae with narrow coronas that are displaced orally due to flattening or invagination of the oral field. These larvae have flattened or invaginated oral field, narrow corona is at the basal (oral) margin of the larval body. Apical disk is small knob-like, no pallial groove. Small metasomal sacs. The larvae appear to be fully differentiated only after a week; this would complicate identification of the larval type in the laboratory conditions for those larvae which have only been released from the colony.	*Tricella koreni; Alcyonidium duplex.*
Type E Coronate Larvae	Coronate larvae with narrow, equatorially positioned coronas. The position of the corona has been one of the main characters separation Type O and Type E larvae. Oral-aboral axis is short, there is lengthening along the anterior-posterior axis. The apical disc is large, pallial sinus present as a shallow furrow. Oral hemisphere is flattened interiorly in the region of the ciliated groove. In most species there is not development of the digestive system; however, a complete larval gut	*Tendra zostericola; Alcyonidium polyoum; A. variegatum; Victorella muelleri; Membraniporella nitida; Smittina pappilifera; Watersipora cucullata.*

	was reported for *Tendra zostericola.* In some species (*Alcyonidium polyoum*) the gut is transitionary and pharynx and stomach disappear during the early embryogenesis.	
Type AE Coronate Larvae	Coronate larvae with extended coronas that are aboral and equatorial AE in position.The coronal cells in this type are considerably higher, and displaced towards the aboral pole, ciliation is also found around the entire surface. Apical disc is of "modest dimension", occupies the entire aboral field, pallial furrow is not open, and appears as a cleft on the surface. Oral field is flattened or convex interiorly, and bulges posteriorly. Larvae are slightly elongated. The gut is not developed.	*Cellepora pumicosa; Catenicella cantei; Savignyella lafonti; Escharoides coccinea.*
Type AEO/ps Coronate Larvae	Coronate larvae with extended coronas that are aboral, equatorial, and oral (AEO) in position and with small pallial sinuses (ps). Corona of these larvae is expanded so much that the polar fields are limited to small circles. Apical disk of medium size, bordered by shallow pallial sinus. Epidermal and mesodermal blastemas are prominent. The subpallial aboral epithelium of these larvae appears like a minor ring. The shape of the larvae is cylinder-like with almost equal sizes in oral-aboral and anterior-posterior axes. Coronal cells cover the entire oral-aboral surface, because the corona is so extensive the pyriform organ is surrounded by the corona, rather than being at the oral margin. Metasomal sac opens near the oral pole and is large. Digestive system is lacking.	*Scupocellaria* spp.*; Bugula* spp.*; Cellaria salicornia; Cupuladria doma; Discoporella umbellata; Celleporella hyalina.*
Type AEO/PS Coronate Larvae	Coronate larvae with expanded coronas that are aboral, equatorial, and oral in position (AEO) and with exceptionally developed sinuses (PS). These larvae are elongated in oral-aboral axis, the apical disc is small, and blastemas are smaller in comparison to the apical disc. Pallial sinus extends nearly to the oral pole; metasomal sac is small, situated at the centre of the oral field. No larval gut is present.	*Amathia lendigera; Bowerbankia pustulosa*

With the advances of microscopy, new methods were successfully tested for the description of morphology of the larvae. The use of fluorochrome dyes and epi-fluorescence microscopy was successfully implemented to study the surface cells of the bryozoan larvae and thus their superficial morphology (Porter and Spencer-Jones 2000; Santagata and Zimmer 2000).

Although there is clear evidence from the literature of wide use of SEM in bryozoology, including those studies dedicated to larval morphology, and the results are exceptionally clear with detailed images, this method is limited to very detailed studies of morphology and not very suitable for a large-scale survey of larval morphological types due to its relative difficulty and large time involvement.

Santagata and Zimmer (2000) proposed a novel method of comparing the surface cells of the bryozoan larvae using fluorochrome stains which specifically target nuclear DNA (Hoechst H33342) or mitochondria (DASPEI and Mitotracker® Orange). All three stains above are cell-permeant, meaning that they are capable of penetrating cells with undamaged lipid cell membranes and therefore can be used to image viable cells. Non-permeant stains can only stain those cells that have a compromised cell membrane, such as fixed slide preparations. The uptake of the mitochondria specific stains (DASPEI, Mitotracker® Orange) is also affected by the activity of the mitochondrial membrane potential and thus by the physiological state of the organism and its cells. Therefore any disruption to the mitochondrial activity of the cell can potentially inhibit the uptake of these stains, and thus the organism in question has to be kept alive. In the case of bryozoan larvae this in practical terms means staining them shortly after their release from the colony as larval death was noted within approximately 24 hours of being kept in the artificial environment after their release from the colony.

The method described by Santagata and Zimmer (2000) allows staining of cells of the corona, neural plate, the surface cells: both transitionary and those, which will later contribute to the formation of the ancestrula. The cells were studied by means of vital, nuclear and mitochondrial stains. These stains showed the following surface elements in several bryozoan species – neural plate, ciliated ray cells, coronal cells, vibratile plume, border cells with ciliary tuffs, sensory cells of eyespots, oral ciliated cells, and some other cells with less fluorescence. This method was subsequently used (Porter and Spencer-Jones, unpublished personal communication) to study larval morphology of *Alcyonidium* and *Bowerbankia*

species. The above method allows relatively simple and quick way of examination of the surface larval components, which could help during the identification of larval types.

6.3 Epi-fluorescence microscopy method

Its relative simplicity and the considerable amount of morphological information which can be acquired about the larvae using an epi-fluorescent microscopy method made it an obvious choice for the larval imaging in the present study. The method was tested using the same fluorochromes as described in the Santagata and Zimmer (2000) method. During their test they found that DASPEI gave similar results to Mitotracker® Orange, however Mitotracker® Orange was found to be more specific to the mitochondria in its binding and gave in general better resolution. As a result only two stains were selected to be tested in this work: Mitotracker® Orange CMTMRos (Molecular Probes M7510) and Bisbenzimide Hoechst 33342 (Molecular Probes H1399). The protocols of stain preparation and concentration were based on the Molecular Probes recommendation and that of the Santagata and Zimmer (2000) method and the details of the resulting protocol are given below.

6.3.1 Mitotracker Orange fluorochrome

Mitotracker® Orange (Figure 54) CMTMRos (Molecular Probes M7510) is a fixable cell-permeant derivative of tetramethlrosamine, which binds specifically to mitochondria, with the binding site of this stain to be co-located next with antibody for subunit I of cytochrome oxidase. The molecular weight of this compound ($C_{24}H_{24}Cl_2M_2O_2$) is 427.37. Its excitation and emission spectral maxima are 554 nm and 576 nm respectively.

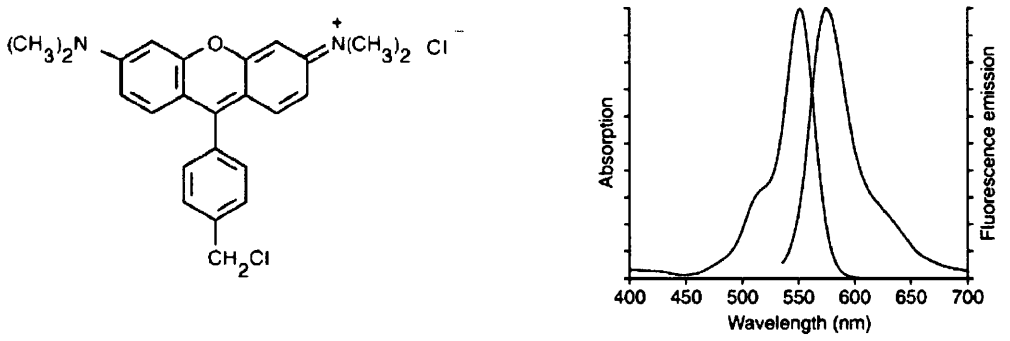


Figure 54 Mitotracker® Orange. On the left its chemical structure is shown. On the right its absorption and emission spectra (left and right peaks respectively). Images reproduced from the Molecular Probes online database.

This fluorochrome is soluble in dimethyl sulfoxide (DMSO), which was used as the main solvent for the stain. In addition magnesium sea water (MSW) was used as part of the preparation of the stain. MSW is a part of the anaesthetic medium for the larvae (see below a separate section on the issues related to the larval sedation). Mitotracker® Orange is supplied in vials containing 50 μg of the lyophilised solid stain ready for reconstitution when required. It was stored at -20°C in the supplied vials and only one vial at a time was diluted and prepared as required use as the diluted stain has much shorter shelf life compared to its lyophilised form. The working solution of the fluorochrome was prepared as following.

- 50 μg (one vial) of Mitotracker® Orange as supplied.
- 585 μl of 100% DMSO (spectrophotometric grade, Sigma), melt on water bath at 42°C thus making 200 μM solution of diluted fluorochrome.
- 150 μl of this solution is mixed with 100 ml of MSW (two parts of sterile sea water and one part of 7.5% $MgCl_2$)

The above dilutions were performed in dark conditions as the fluorochrome stains are light sensitive. Once dilutions were prepared as per the above recipe a working solution of approximately 300 nM was ready to be used. This was aliquoted onto several 1.5 ml Eppendorff tubes and wrapped in foil (to minimise exposure to light) and frozen at -20°C until required.

6.3.2 Hoechst 33342

Bisbenzimide Hoechst 33342 (Figure 55) (Molelcular Probes, H1399) is cell-permeant bisbenzimidazole derivative which binds specifically to the minor groove

of DNA with AT selectivity. Molecular weight of this compound ($C_{25}H_{37}Cl_3N_6O_6$) is 623.36. Its excitation and emission spectral maxima are 350 nm and 461 nm respectively.

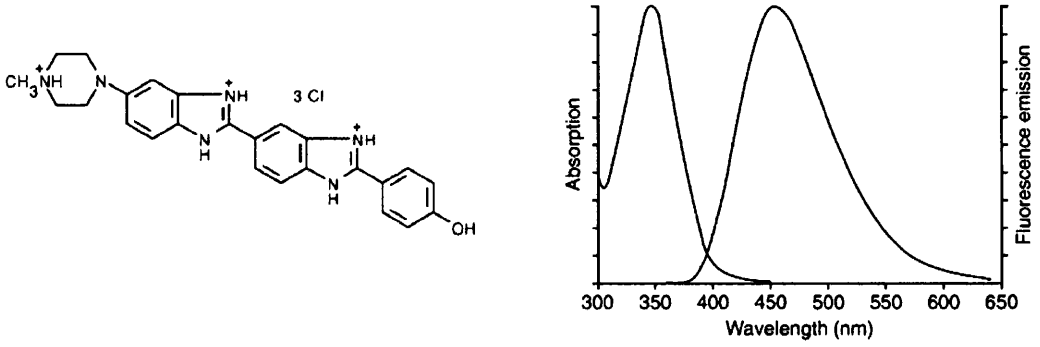


Figure 55 Hoechst 33342 fluorochrome. On the left its chemical structure is shown. On the right its absorption and emission spectra. Images reproduced from the Molecular Probes online database.

This stain is both water and DMSO soluble and in this case was directly dissolved in water using the following steps to achieve 100 μg/ml working concentrations.

- NaCl 475 mM solution in deionised H_2O
- KCl 25 mM solution in deionised H_2O
- MSW (two parts of sterile sea water and one part of 7.5% $MgCl_2$)
- Hoechst 33342 stock solid to bring it to 100 μg/ml concentration

This was done directly prior to the staining or a stock of aliquots was prepared using 2 mg of Hoechst 33342 and 20 ml of the NaCl, KCl and MSW solution as per concentrations above. The stock was aliquoted into 1.5 ml Eppendorff tubes wrapped in foil (to minimise exposure to light) and frozen at -20°C until required.

6.3.3 Larval extraction and staining

Live extraction of larvae from the colony was based on the known positive phototaxis of the larvae just after their release. Once the colony with larvae were identified (see Chapter 2 for description of sample collection) they were placed together with the substrate they were found on (in many cases a stone) in the light insulated tank built specifically for this purpose and located in the constant temperature room[36]. This tank had constant seawater and air supply and colonies

[36] Temperature in the CT room was maintained equal to that of the ambient seawater temperature at the time.

could be kept in it for several days. The larval release is usually triggered by the light in nature (presumably light which reaches the colony in the morning). Once ready for the extraction of larvae the colonies were transferred into a small transparent container placed underneath a stereo light microscope (Olympus, SZ60) and "cold" light[37] from a fibre optics illuminator source was positioned at one side of the container. That way the container could be observed under the stereomicroscope and once larvae released they would swim and congregate towards the light source (on one side of the container) at which point they could be easily collected using a 1 ml Gilson micropipette. Larvae were released usually within 20 minutes from the exposure to the light.

Once collected, larvae were anaesthetised using MSW (one part of 7.5% $MgCl_2$ and 2 parts of sea water). This solution usually worked well with the larvae and within 30 minutes they would cease moving, with only some cilia movement noticeable.

Once anaesthetised larvae were transferred to watch glasses and stained using method specific to the stain. For the Mitotracker® Orange, once stained for approximately 20-30 minutes using previously prepared working solution (see above), larvae were washed twice in the MSW to clear away the fluorochrome, then placed on the slide for imaging. The larvae were mounted on a microscope slide with the cover glass placed on four small pads made out of plasticine (Blu-Tack). For the Hoechst 33342 once working solution was prepared it was added directly to the watch glass containing MSW in the proportion of 1 part of stain to 10 parts of the MSW (this gave an approximately 10 μg/ml working dye concentration) and larvae were stained for 30 minutes. No washing was required after the staining was done.

Images were taken using a Nikon Eclipse E600 epi-fluorescence microscope, the operational principle of which is shown on Figure 56. For Mitotracker® Orange a green fluorescent filter was used (block G-2A, Excitation filter wavelength 510-560 nm, dichromatic mirror 565 nm, barrier filter 590 nm). For Hoechst 33342 an ultraviolet filter was used (block UV-2B, excitation filter 330-380 nm, dichromatic mirror 400 nm, barrier filter 435 nm).

[37] It is important to use fibre-optics in this case to minimise heat shock to the colony as the ambient temperature in the laboratory is always above that of the sea temperature.

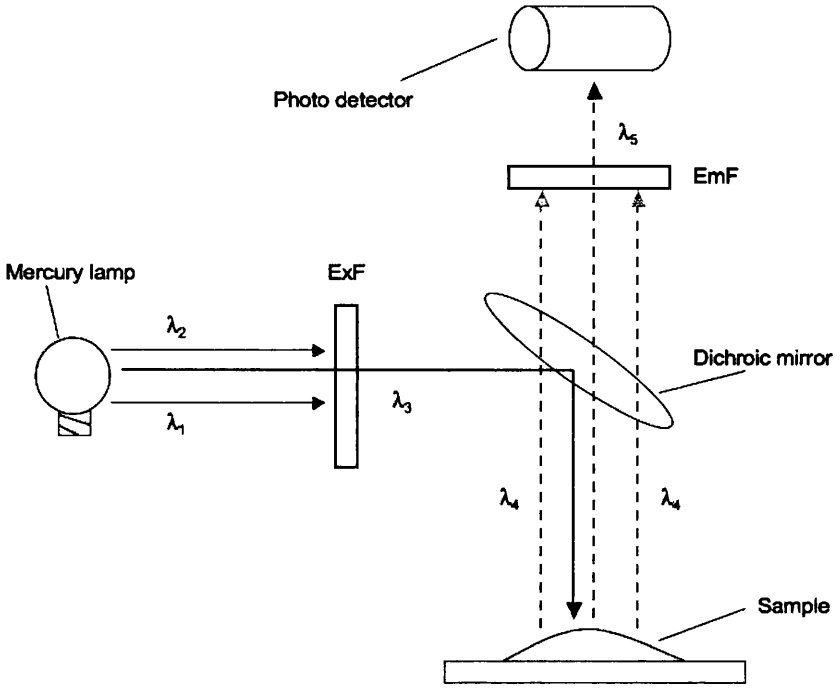


Figure 56 Principal schematics of the epi-fluorescence microscope. Mercury lamp emits wide spectra of light (λ_1 λ_2 λ_3), the desired excitation spectrum (λ_3) is selected by the use of excitation filter (ExF). The light is then directed to the dichroic mirror which separates emitted from the mercury lamp spectrum (λ_3) and the scattered emitted light of the sample. Light is reflected from sample with the florescence spectra (λ_4 λ_5). The desired emission spectrum of the fluorochrome (λ_5) is filtered by the emission filter (EmF) and collected by photo equipment (usually a digital still camera). Image adopted from Haugland (2002).

Whilst the initial results with the above method were successful, some problems emerged linked to the method in use. Firstly, there was a problem with mounting the larvae on the slides. Once the larva is stained and placed on the slide glass with the cover glass, it was no longer possible to change its position as the larva became damaged. Also a common drawback of the epi-fluorescence was obvious – the images suffered from lack of sharpness both because of the common limitations of the epi-fluorescence method linked to the out-of-focus glare and consequently lack of resolution (Amos and White 2003). As with any light microscopy there were general difficulties associated with observing a relatively large three-dimensional object under a microscope (*i.e.* out of focus problems). Also images suffered with some autofluorescence which added to the uncontrolled glare in the image. Similar problems were observed by other authors who worked with bryozoan larvae using this method (Santagata and Zimmer 2000, Porter Spencer-Jones 2000). The above problems are generally reported as one of the major

disadvantages of epi-fluorescence microscopy (Amos *et al.* 1987, Amos *et al.* 2003, Claxton *et al.* [no date]).

6.3.4 Some results of epi-fluorescence method

Shown below are several images taken using epi-fluorescence microscope. Whenever possible interpretation of morphological characters are given and larval types allocated according to the system of Zimmer and Woollacott.

Escharella immersa larva lateral view is shown in Figure 57. Here many organs of the hypothetical larvae can be identified. The larvae is most likely to be Type AE larvae.

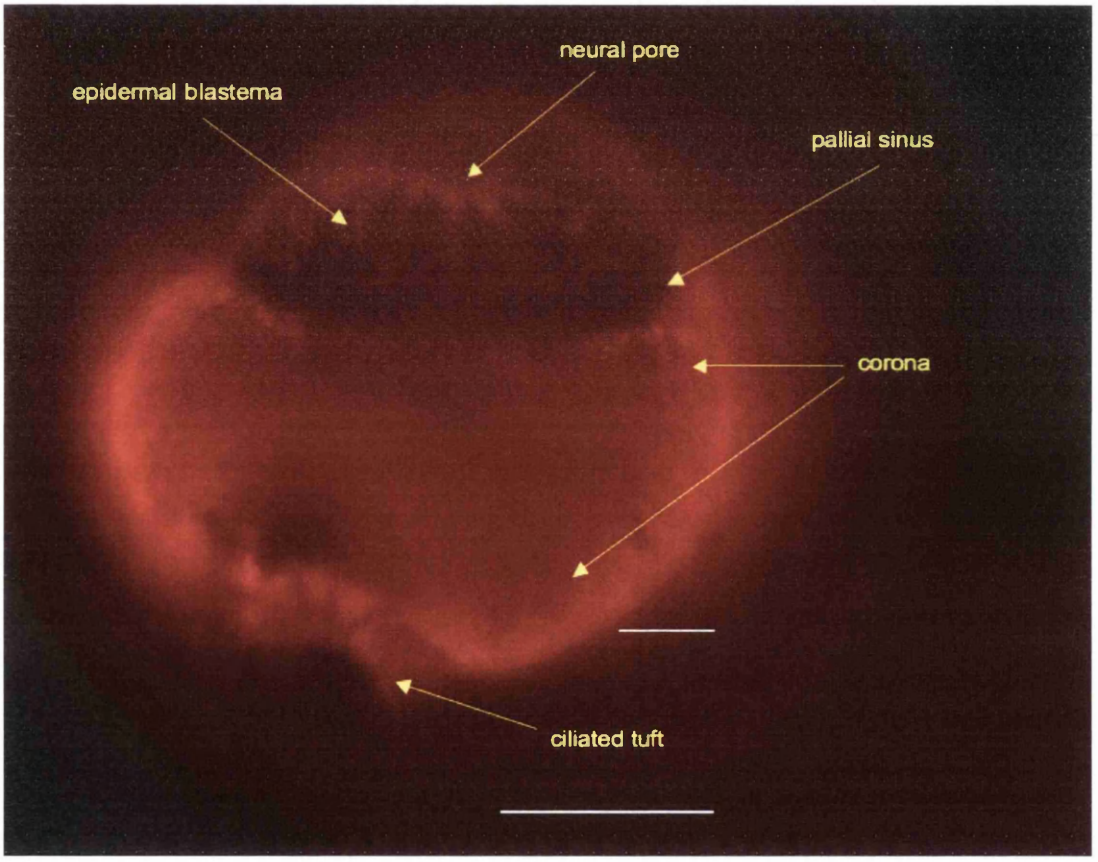


Figure 57 Epi-fluorescence image of the lateral side of *Escharella immersa* larva, stained with Mitotracker orange. Hazy glow around the larvae is from the ciliated cells of corona. Scale bar is 100 μm.

On the image above an out of focus glow artefact is clearly visible. Although a series of images were taken with this specimen the glow caused by the fluorochrome prevents successful sharpening of the image and makes many morphological characters of the larva very difficult to identify.

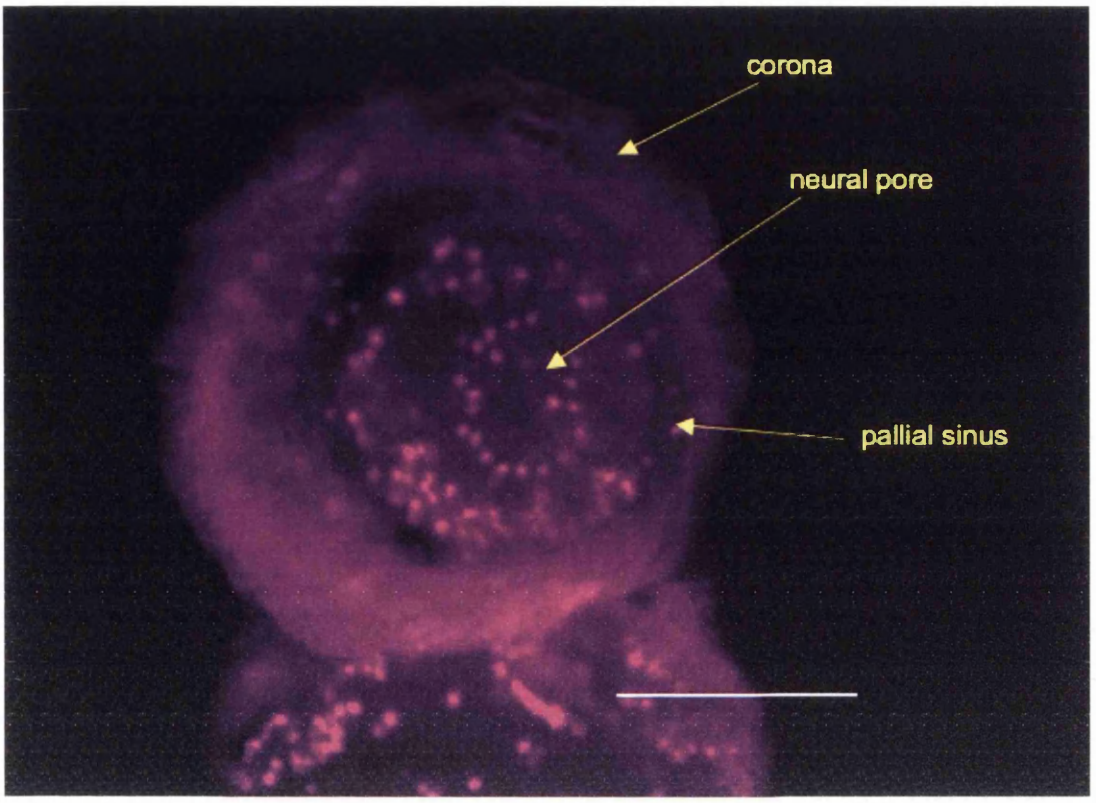


Figure 58 Epi-flourescence image of the aboral pole of *Phaeostachys spinifera*, stained with Hoechst 33342. Scale bar is 100 μm.

Phaeostachys spinifera larvae were imaged using both Hoechst 33342 (Figure 58) and Mitotracker® Orange (Figure 59). On both images, organs of the aboral pole can be seen. The haze from the corona is mostly noticeable on the image stained with the mitochondrial stain. This larvae is type AE or type AEO/ps as per the Zimmer and Woollacott system.

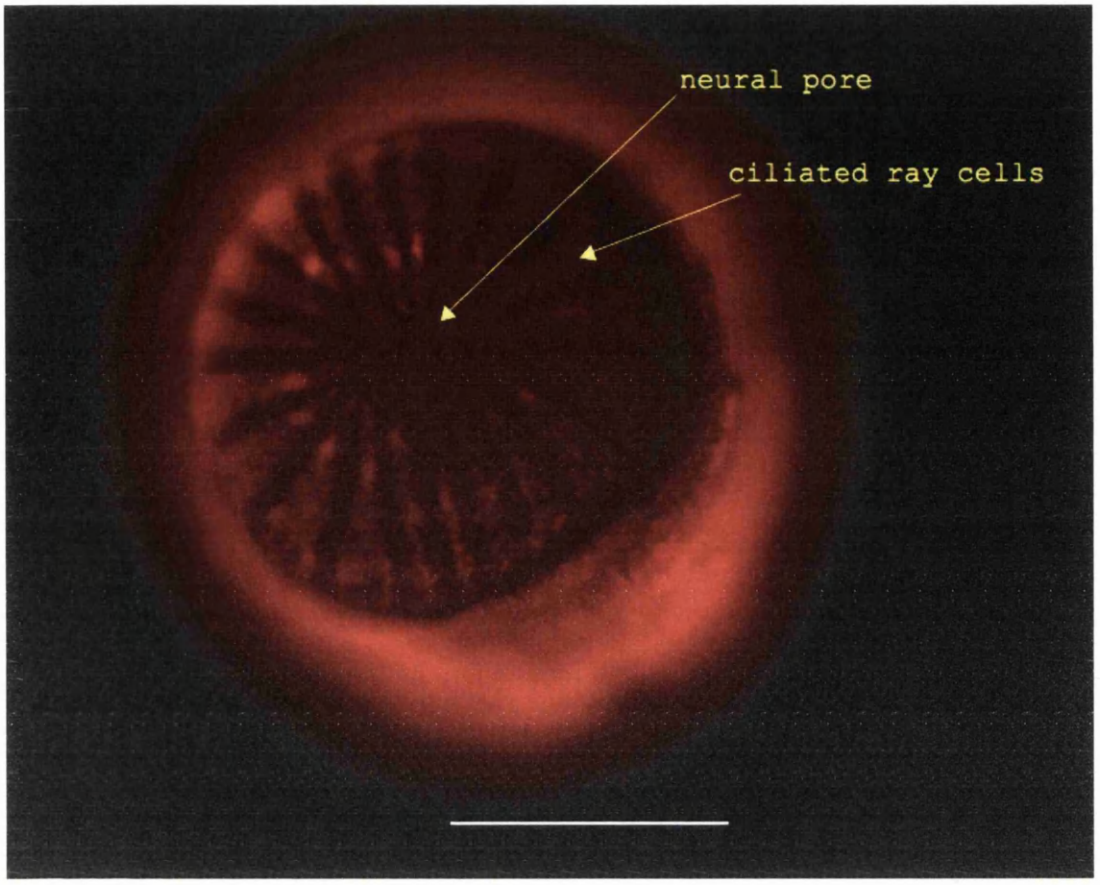


Figure 59 Epi-fluorescence image of the aboral pole of *Phaeostachys spinifera*, stained with Mitotracker® Orange. Scale bar is 100 μm.

From the images above the problems associated with the epi-fluorescence method mentioned above can be clearly seen – especially out of focus glare. The inability to move larvae or re-position them once they have been mounted on the slide and a cover glass was placed on top has affected the relative position of larvae on the images.

Due to advantages of the confocal microscopy over epi-fluorescence optics in general (see section below for discussion) and the availability of the confocal laser microscope (CLM) for research at Swansea University, it was decided after a few attempts with the epi-fluorescence microscope to halt its use and concentrate on the development of a new method which would allow to use CLM system.

6.4 Confocal microscopy

The main limitation of the above method is in out-of-focus parts of the specimen, which give rise to a uniform glow that is clearly noticeable in the images presented above. This glow prevents the fine details of the specimen to be seen. Epi-

fluorescence microscopy produces somewhat satisfactory results and its use is justified, although the limitations of this method are clear (Amos and White 2003).

The development of confocal laser microscope and its adoption in biology took a long time. The concept of the microscope was developed in 1955 by Marvin Minsky (Minsky 1988) but only was wider accepted in the biological research when a working prototype was developed in Cambridge in 1986 specifically for biological samples and rapidly took over many fields in biology and was widely adapted by the late 1990s (Amos and White 2003, Claxton *et al.* [no date]). The word confocal refers to illumination confined to a diffraction limited spot on the specimen whose plane is confocal (that is having the same foci, or conjugate) to the pinhole aperture plane. This provides among other things a possibility of optical sectioning of the specimen (Amos and White 2003).

The most important advantages of using a confocal microscope are complete elimination of the glow artefact produced in the epi-fluorescence method due to the fact that most of the out-of-focus fluorescence is filtered out by the pinhole aperture confocal to the objective focal plane. The method also allows filtering out autofluorescense by means of spectral unmixing and finally allows a 3D reconstruction of the specimen via the z-stack[38]. Two major types of confocal microscopes exist: laser scanning confocal microscopy (LSCM) and spinning disk confocal microscopy (SDCM). The difference is that in the former method the laser scans the entire specimen line by line using one spot, whereas SDCM uses a spread beam of laser and spinning disk (so-called Nipkow disk) which has holes in it and thus allows create several simultaneous spots on the specimen. In this study a LSCM was used, in particular a Carl Zeiss LSM 510 META microscope.

Confocal microscopy is similar to the epi-fluorescence microscopy in that a fluorochrome dye is irradiated with light of a certain wavelength λ_{ex} causing electrons in this fluorochrome to be raised to higher energy levels then when they drop back to they original energy levels they emit light (photons) of a lower wavelength λ_{em} thus $\lambda_{em} > \lambda_{ex}$ (Figure 60).

[38] The microscope optical axis is parallel to the specimen plane z (vertical), as opposed to the x and y, which are perpendicular lateral dimensions of the optical and specimen plane. Z-stack is called so because several images or optical slices in z plane can be stacked (assembled) together to reconstruct a 3D image of the specimen.

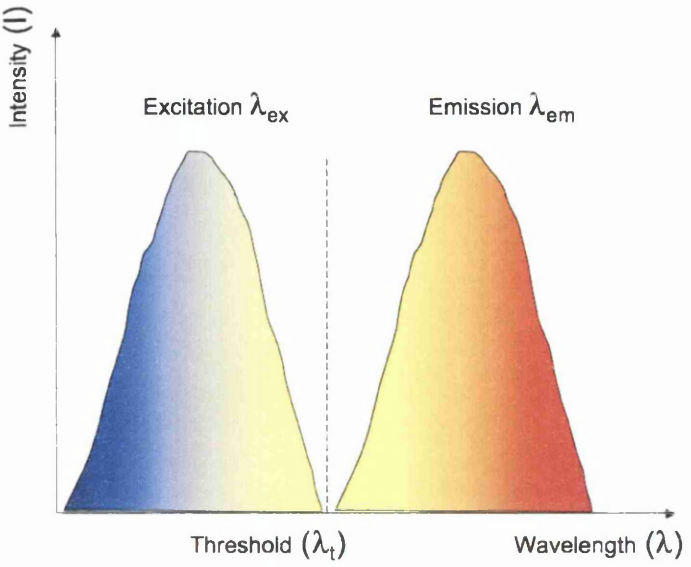


Figure 60 Principle of fluorescence. Graph shows excitation and emission intensity (I) as function of wavelength (λ). The threshold (λ_t) wavelength refers to the separation wavelength of the dichroic beam splitter (or mirror), thus separating excitation and emission spectra.

The mercury lamp which is used in epifluorescence microscope is replaced in LSCM by lasers with a fixed wavelength which are responsible for the excitation of the fluorochrome. Fixed wavelength (*i.e.* being monochromatic) and light intensity of a laser is better suited to the excitation process of the fluorochrome than conventional light source because of the loss of illumination in exciting the fluorescence (Rochow and Tucker 1994). The advantages of the laser are also in the pin-source coherent illumination of the specimen, further enhanced by the presence of the confocal pinhole. Spectral channels unmixing (META detector in Carl Zeiss LSM 510 microscope) allows separation or optical grating of the emission spectra into 32 channels, thus enabling very precisely filtering out of only the required spectrum in the image and giving a spectral signature to each acquired pixel.

In the confocal microscope (Figure 61) the laser light is directed towards the specimen via the dichroic beam splitter (or mirror), which has a 80:20 transmission reflection coefficient. Then it is focused via the objective on the sample. The focal plane of the specimen could be precisely regulated thus allowing a series of images at different focal planes to be taken for later assembly in a vertical stack (z-stack). This assembly allows a 3D reconstruction of the object. Emission from the sample is focused back through the lens; it passes again through the dichroic mirror and continues towards the emission filter that further separates excitation and emission spectra then passing it further into the confocal pinhole. The pinhole cuts off all out-

of-focus light and passes only parfocal with the excitation point image to the photomultiplier tubes (electronic light detector similar to that of the digital camera charge coupled device). It is important to note about the photomultiplier tubes is that they do not detect any colour that is they are "colour blind" and only generate an electron when a photon presence is detected. Thus any colouring of the image is done during the post processing of the captured image and can be freely changed by the investigator.

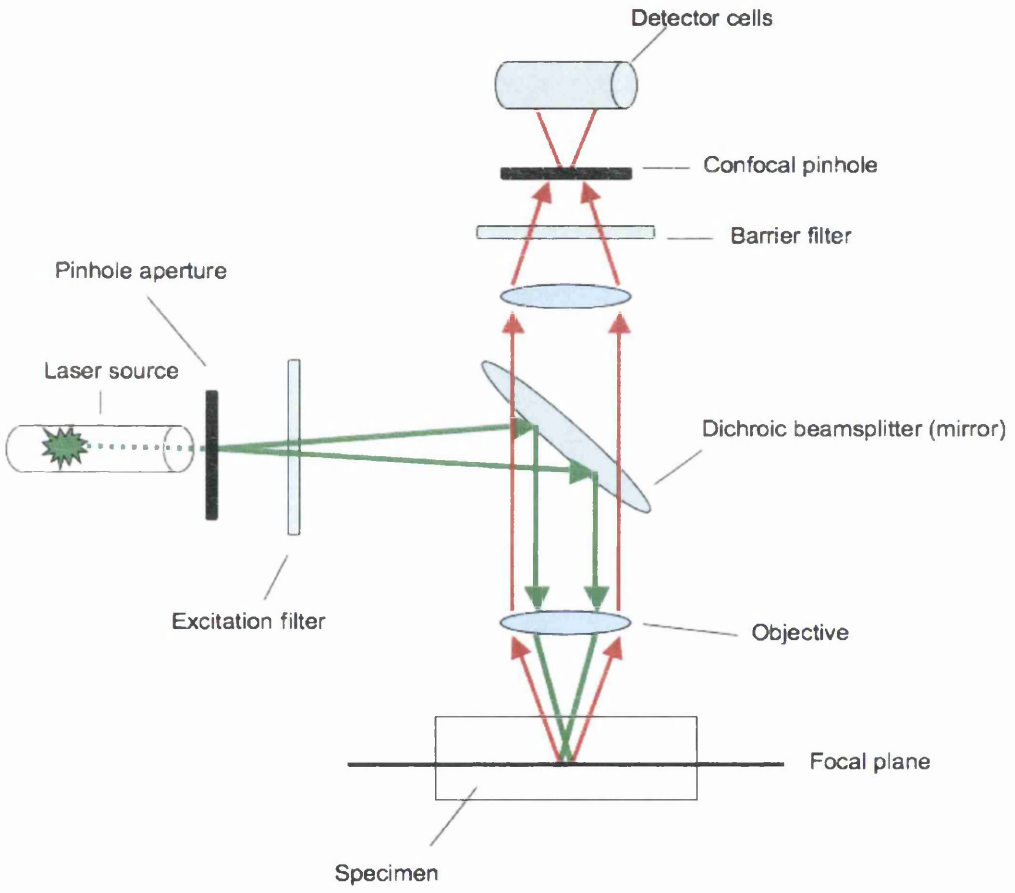


Figure 61 Light beam path in the CLSM. The light from the laser before it reaches the specimen is coloured green, after it was reflected it is coloured in red. Image adopted from the Carl Zeiss LSM 510 META manual.

The entire process of confocal laser microscopy is controlled via an integrated computer system which allows full manipulation of the specimen, microscope, its components and finally a post processing of the acquired images (including 3D reconstruction when necessary). The theoretic resolution of the CLSM is determined by the pinhole size (the smaller the pinhole the higher the resolution) and the numerical aperture of the lens (similar to that of the light microscope resolving power). In practical terms the number of pixels in the final image as it is

captured during the laser scanning also affects the final image resolution and therefore can be treated as one of the factors of resolving power. In general though, the resolution of the confocal image apart from obvious theoretical limits is also affected by the contrast and the thinness of the specimen and is never as high as that of the scanning electron microscope (Claxton *et al.*[no date], Amos and White 2003).

6.4.1 CSLM method evaluation

In this study once some images were acquired using an epi-fluorescence microscope (Nikon Eclipse E600) and clear disadvantages were noted a development of the method suitable for the imaging using a CLSM commenced. Because of full compatibility of the fluorochrome stains used in the epi-fluorescence method and their apparent suitability for the cause, the new method was based on the same fluorochrome (Mitotracker® Orange), which were tested before. Although Hoechst 33342 also gave good results with epi-fluorescence microscope and was suitable for confocal microscopy its use was hindered by the lack of the laser line suitable for its excitation[39] (its excitation spectrum was 350 nm). This would require an Argon UV laser (351/364 nm) which was not available at the time at Swansea University.

6.4.1.1 Larval extraction and staining

Larval extraction was performed in the same manner as for the epi-flourescence imaging method (see section above). Once extracted, live larvae were stained using Mitotracker® Orange fluorochrome as per the method described above. Once stained the larvae were sedated in order to completely immobilise them. This step had to be amended from the epi-fluorescence technique as it was found that larvae once exposed to the laser in the confocal microscope became active regardless of the time for which they were sedated. Larger larvae were affected more by this problem *i.e.* they became active quicker and responded less to the sedation method.

Although it is possible to image the larva using a confocal microscope if they are moving, the resulting image can not be used for the 3D reconstruction. One of the greatest advantages of using a confocal microscope is the possibility to use z-stack to electronically rebuild a 3D image of the entire organism similar to a

[39] The following laser lines were available for LSM 510 META: 405 nm, 458 nm, 477 nm, 488 nm, 514 nm, 543 nm and 633 nm.

hologram. Unfortunately for this to work the series of images have to be taken with x and y plain (*i.e.* perpendicular lateral dimensions of the specimen) to be completely fixed. Then once the specimen is fully immobilised several "slices" of images taken in z-axis could be assembled together. Further, once a 3D image is reconstructed it can be digitally flattened to include all important features observed in an individual slice of the z-stack thus giving a much fuller and sharper two dimensional image of the organism than it could be possible for instance with an epi-fluorescence microscope. Because of the problems with larval movement encountered here a large amount of time was spent attempting to optimise the technique of staining-sedation-imaging workflow.

6.4.1.2 Larval sedation media

Several different media were tried to immobilise larvae. A narcotisation method should be relatively rapid (20-30 minutes) and provide full immobilisation of the larvae but must not kill them as this would have a detrimental effect on the staining process as the fluorochromes used here require live material and would be better absorbed into the cells and adequately bind to the target organelles in the cells. As the method originally described by Santagata and Zimmer (2000) appeared to work worse with the CLS microscope (possibly due to the higher intensity of the light from the laser) several other methods were evaluated.

Many methods recommended for the sedation of marine invertebrates are either too time consuming (requiring several hours to take effect) or if sufficiently rapid they kill the target organism (Smaldon and Lee 1979). One of the methods tested was a modified method of using benzamine hydrochloride (eucaine) in 0.1% solution added to seawater (Smaldon and Lee 1979). Eucaine was not possible to acquire due to the legal restrictions[40] and little availability and therefore its functional relative benzocaine was tested as it was successfully used for sedating aquatic organisms (Prof. D.O.F.Skibinski, personal communication). Benzocaine was used in a several concentrations 0.1%, 0.5%, 1% 1.5% and 2%. However it had no apparent effect on the larvae. Higher concentrations simply killed larvae.

[40] Eucaine (benzamine hydrochloride) is an artificial substitute for cocaine as a local anaesthetic and is not available though usual biochemical suppliers.

Another method evaluated here which is often used for marine invertebrate sedation is carbon dioxide CO_2 enriched water (Smaldon Lee 1979, Ross and Ross 1999, P.J. Hayward, personal communication). This is done simply by using a soda-siphon with filtered sea water as a medium. Carbon dioxide narcotisation worked well on larger larvae (such as those of *Flustrellidra hispida*) but had no effect on the smaller[41] ones (such as *Escharoides coccinea*). In addition bubbles of CO_2 formed in the medium whilst imaging was done and interfered with the process of image taking.

Finally, a magnesium sea water (MSW) sedation method using 7.5% w/v $MgCl_2$ $6H_2O$ diluted with an equal volume of sea water (Ross and Ross 1999) was used. This method differs from the one used by the Santagata and Zimmer (2000) by the proportion of water to magnesium chloride volumes (they recommend using 2:1 sea water to magnesium chloride solution, instead of 1:1). Modification of the concentrations of $MgCl_2$ $6H_2O$ in the MSW from 7.5% to 20% (7.5%, 10%, 15%, 20%) was also attempted. However, although it had a faster immobilising effect, it also had a detrimental effect on the imaging – higher concentrations of $MgCl_2$ $6H_2O$ caused less stain binding to the organelles or even killed larvae. Therefore, the original method suggested by Ross and Ross (1999) was finally accepted for all types of larvae.

Once immobilised and stained, larvae were viewed under the microscope. Carl Zeiss LSM 510 META is an inverted microscope, which means that the objective lens is located under the object. This dramatically simplifies preparation of the slides with the sample. Instead of placing stained larvae on a slide glass and covering them with a cover glass positioned on the wax (in order not to damage the larvae) the stained larvae were placed into a German 8 chambered coverglass (LabTech® II no: 155409) designed specifically for live cell imaging on an inverted microscope (Figure 62A). This coverglass system essentially inverts the positions of the cover glass and slide glass and makes the latter redundant (although a plastic cover is supplied to prevent evaporation of the medium). This chamber glass (Figure 62B) allowed placing several larvae into individual chambers.

[41] Size gradations described here are entirely subjective and were assessed and tested on an individual basis, however as a rule of thumb larvae smaller than 0.5 mm were much more difficult to sedate with CO_2.

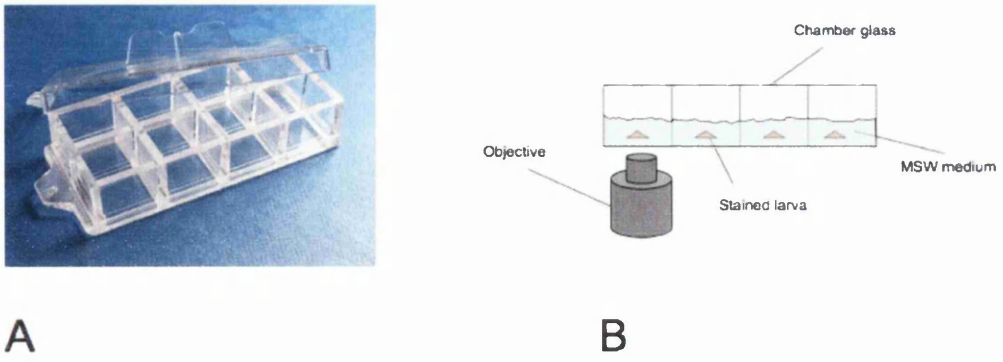


Figure 62 Chambered coverglass (LabTech® II) with 8 chambers. A) - Photograph showing the chamberglass with a lid opened. B) – Schematics showing how the larvae were positioned in relation to the objective lens of an inverted microscope.

By doing so two goals were achieved: firstly the slide preparations did not require any cover glass on top and thus no damaged was done to the larvae by the cover glass; secondly larvae could easily be moved and positioned on the required side once under the microscope by means of a simple preparation needle or a very fine paint brush (size 000). When larvae were stained and positioned on the right side in the chambers of the cover glass the chamber glass was placed on a focusing stage. All functions of the LSM 510 META microscope can be controlled from the guiding computer software. The microscope has an automatic seek function which makes it possible to search for a subject under low magnification, this requires short exposures to the laser and may bleach the sample and weaken the fluorochrome. Therefore manual stage adjustment was used under low magnification (10x objective) with a standard light source prior to switching to the use of the lasers of the microscope. Once the specimen was found, imaging was done first at a lower possible magnification to assess the image contrast and larva position, then a higher magnification was used (allowing filling of the image frame with the larva). Each image was taken in series in the z axis (vertical axis parallel to the sample plane). Because of the photo-bleaching effect a balance between time of exposure and image size had to be achieved – higher image resolution (in pixels) required longer exposure to the laser beam and thus bleached the sample quicker. On average, approximately 20 minutes of working time per sample were enough to bleach the fluorochrome so that no more imaging was possible. Photo-bleaching is a common

problem in confocal microscopy (Amos *et al.* 1987, Longin *et al.* 1993) and is one of the major drawbacks of using live material as anti-fading agents are usually toxic or require special fixing media (such as buffered glycerol) and can only be applied to a fixed non-live material such as slides (Login *et al.* 1993).

A green helium neon 543 nm laser was used for the imaging with the main dichroic beam splitter for 488/543 nm, with an optional secondary dichroic beam splitter set to 545 nm. Manufacturer instructions were followed for the operation of the LSM 510 META microscope. Channel settings (such as Pinhole size, Detector gain, Amplifier Offset) were adjusted on a per sample basis as well as the laser transmission power. The latter was determined in many cases by the condition of the stained material (such as time after the larvae were stained). Once the desired specimen was located and preliminary images taken, the lower and higher focusing planes were assigned and the automated series of z-stack images were acquired for the 3D reconstruction. In all cases, the images were taken using 8 Bit depth with a frame size of 1024x1045 pixels resolution (this provided optimal time acquisition and bleaching vs. quality balance).

6.4.1.3 Results of CLSM imaging method

The main problem encountered with method, which consequently affected its optimisation, was with the lack of material. Most suitable species were those which produced many larvae and whose release was relatively easy to monitor such as *Flustrellidra hispida* or *Alcyonidium* spp. or *Bowerbankia* spp. and whose larvae when released would be abundant. Unfortunately, some species (such as those of *Bowerbankia* spp. for instance) have a short reproductive period and thus offer a limited supply of live material. Problems encountered with larvae sedation caused even further delays with the method development and consequently many larvae were "wasted" during the optimisation method. This problem can be overcome by developing a method based on a readily available model organism. Such an attempt was made using *Artemia salina* nauplii, which appeared to have similar agility and size to many bryozoan larvae. Unfortunately, the response to the sedation using the methods tried for the bryozoan larvae was inadequate – nauplii did not appear to react well to the MSW in the concentrations used for the Bryozoa (*i.e.* 7.5% of $MgCl_2$). As a result their use was discontinued.

Below several images taken using the LSC microscope method are presented. For each species of larva its larval type is identified according to the system of Zimmer and Woollacott (1977) whenever possible. The advantages of the method as described above are clear from the images. For instance *Phaeostachys spinifera* imaged with epi-fluorescence method (Figure 59) and using CLSM method (Figure 67 and Figure 68) gives completely different resolution to similar larvae. Images of *Flustrellidra hispida* (Figure 65 and Figure 66) were particularly interesting as they gave a hologram-like 3D image, which resembled those of the SEM images. Many features of the larvae can be clearly seen such as shells of the larva. On Figure 66 an image of *Flustrellidra hispida* larva is given from the oral pole. This image although giving a good resolution at the posterior side of the larva is very fuzzy at the anterior part of the larva. This was caused by a sudden movement of the larvae during the imaging and resulted in the distortion of the x-y plane and as a result, misalignment of the z-stack images during the 3D assembly. It is a very good example of the problems associated with method when the movement of larvae cannot be controlled. The remaining images (Figure 63, Figure 64, Figure 67 and Figure 68) give some indication of the method capabilities. None of the images were manipulated and are here represented as they were taken using the microscope. The only adjustments were made to the contrast of some images (in order to show less stained organs) and the channel colouring of the images was changed (yellow, white, red) to vary the relative perception of contrast of some organs.

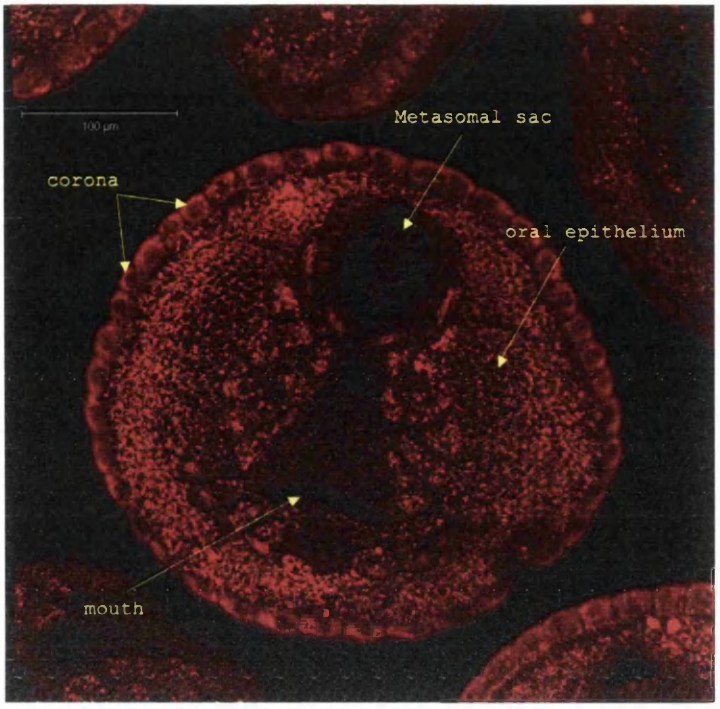


Figure 63 CLSM image of the oral pole view of *Alcyonidium hirsitum* type E larva, stained with Mitotracker® Orange. Image is pseudocoloured. Several other larvae can be seen around. Scale bar is 100 µm.

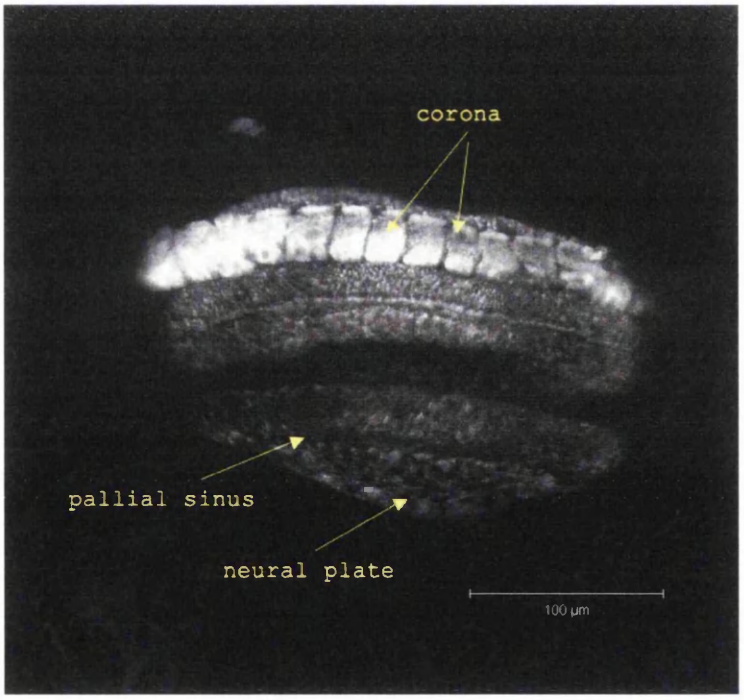


Figure 64 CLSM lateral view image of *Alcyonidium hirsitum* type E larva, stained with Mitotracker® Orange. Scale bar is 100 µm.

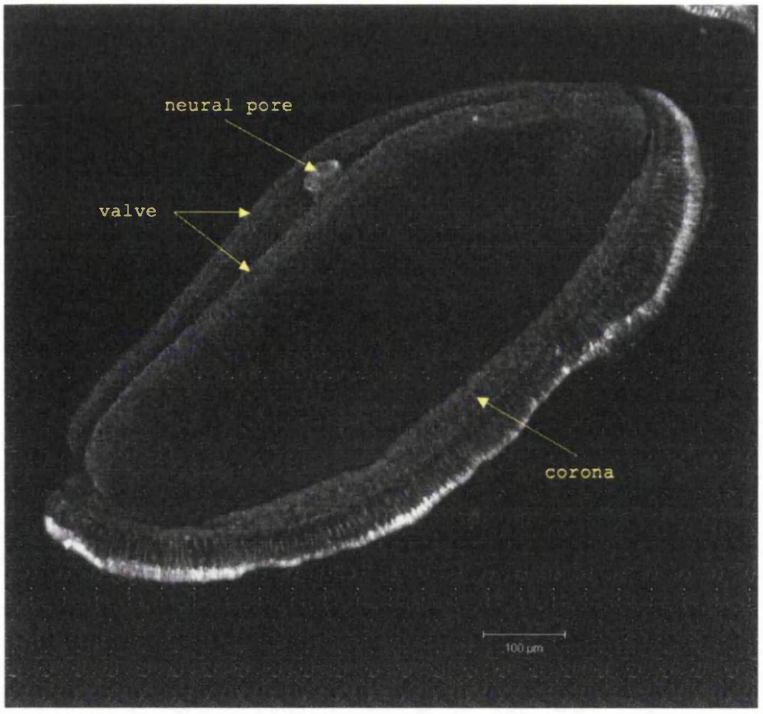


Figure 65 CLSM lateral-aboral view image of *Flustrellidra hispida* shelled lecithotrophic larva, stained with Mitotracker® Orange. Scale bar is 100 µm.

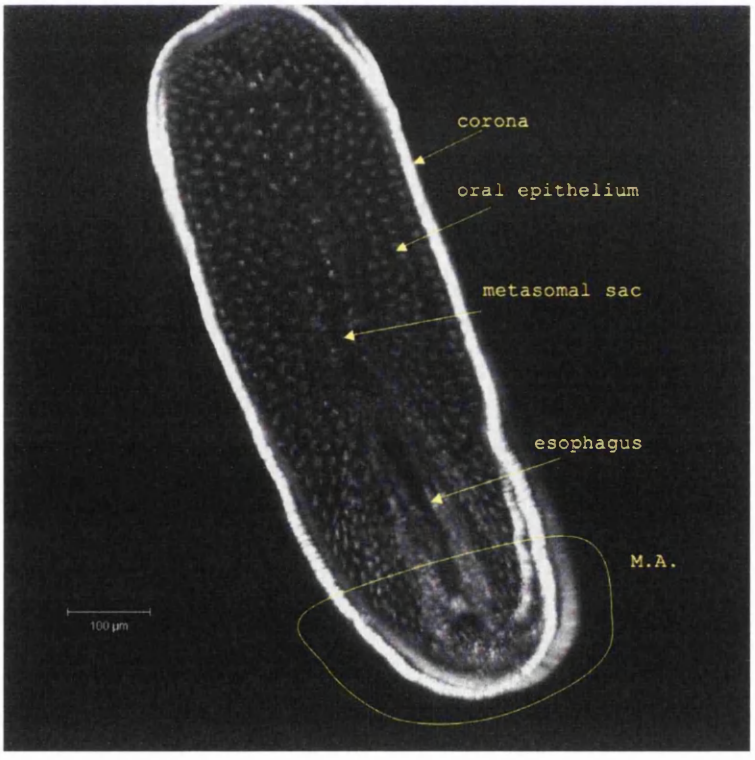


Figure 66 CLSM image of the lateral pole view of *Flustrellidra hispida* shelled lecithotrophic larva, stained with Mitotracker® Orange. Misaligned z-stack planes (M.A.) are marked by yellow line. This was caused by sudden movement of the larva during imaging. Scale bar 100 µm.

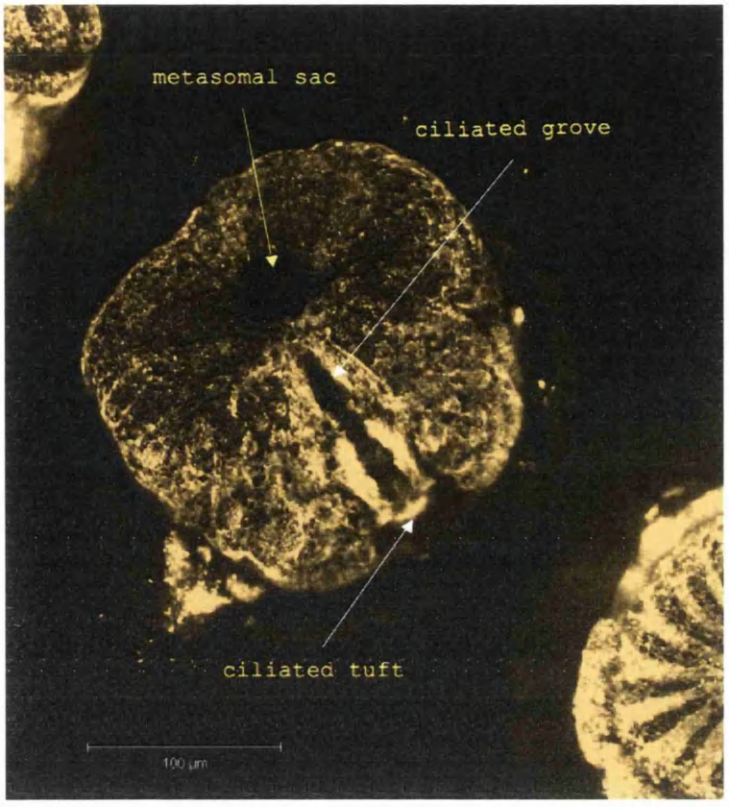


Figure 67 CLSM image view of the oral pole of the *Phaeostachys spinifera* Type E larva, stained with Mitotracker® Orange. Image is pseudocoloured. Scale bar is 100 µm.

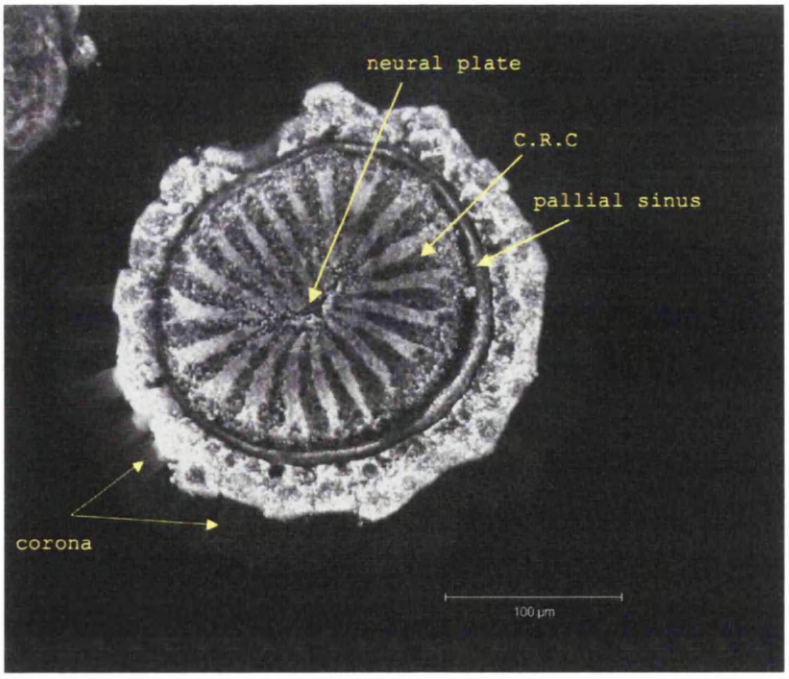


Figure 68 CLSM image view of the aboral pole of *Phaeostachys spinifera* type E larva, stained with Mitotracker® Orange. Scale bar is 100 µm.

The CLSM method in combination with a mitochondrial stain showed good results for the surface imaging of bryozoan larvae and demonstrated that it could be used for quick assessment of major larval morphological characters and identification of larval types. However, it is view of the author that more fluorochrome stains should potentially be evaluated as they may give more detailed surface structure of the larva. It is believed by the author that specific attention should be given to the BODIPY® 505/515 stain (Molecular Probes). This stain belongs to a group of membrane fluorochromes which are fluorescent analogues of phospholipids capable of incorporating themselves into cell membranes. These are relatively new stains that have shown great results for cytoplasmic staining with some model organisms such as zebrafish embryos (Cooper *et al.* 2005). Their other advantage is a very low photobleaching rate, which means that the stain can be used for much prolonged time, and thus offer better imaging of the bryozoan larvae.

The method presented above can clearly aid in identification of larval types based on the system of Zimmer and Woollacott (1977). However not in all cases presented here could larval type identification be performed with 100% certainty. This was mainly due to the fact that more images would be required which would show a particular larva from different angles. The 3D reconstruction technique is limited in bryozoan larvae by their size. The larger the object under the microscope, the more difficult it is to perform a complete scan of all focal planes, therefore requiring several independent images taken from different larvae positioned at different angles. Thus for a complete larval identification a larva has to be imaged from at least oral and aboral poles, and preferably a lateral image has to be taken as well. This would enable a very precise collection of morphological characters required for the identification of larval type based on the Zimmer and Woollacott (1977) system. This is especially true for those larvae whose morphology is somewhat similar, for instance those larvae belonging to Type E or Type AE. In some cases, such as *Flustrellidra hispida* (Figure 65), the type of the larva (*i.e.* shelled lecithotrophic) can be clearly identified because of the distinctive characters of this type.

7 GENERAL DISCUSSION

7.1 Primer design

One of the aims of this study was design and presentation of a working set of primers suitable for a broad range of bryozoan species. Ideally, so-called universal primers were sought which would enable rapid generation of 18S rRNA sequences from the material. Although universal primers were described before (Hillis and Dixon 1991; Halanych *et al.* 1995; Halanych *et al.* 1998), they did not appear to give good results in this study. Their use, application and universality is based on the fact that they anneal to very conserved regions of the 18S rRNA gene. These areas, described in Chapter 5, are those which most commonly correspond to the stems of the 18S rRNA and thus evolve very slowly and undergo a very low substitution rate. The use of these primers of course can speed up data accumulation, but also can introduce additional problems as these universal primers can potentially anneal to the DNA of foreign organisms (*i.e.* contaminants) due to their universality. This is especially critical for marine invertebrates where cross contamination is rather common. Contaminants were reported for two bryozoan sequences (*Lichenopora* sp. and *Membranipora* sp.) by Waeschenbach (2003) and in this study some sequences, especially those of Hao *et al.* (2005), were questioned due to the method with which the DNA was obtained and their unusual clustering. In addition,the published *Alcyonidium gelatinosum* sequence validity was questioned by others (Dr J. Porter, personal communication).

The methods involved in DNA extraction are important and were specifically discussed in Chapter 3 and a more reliable method which is less prone to contamination based on the DNA extraction from embryos was used (Porter *et al.* 2001). Although this method minimises possible cross contamination it has its own drawbacks. Firstly it is dependent on the reproductive cycles of Bryozoa and thus DNA material can only be obtained from those species which are found to be reproducing, therefore limiting the number of species available to the investigator. In this study 55 species were observed in total, however only 42 were collected in the reproducing stage. In addition, the method is limiting because of the sometimes low amount of DNA which can be obtained, as sometimes embryos are lost during the extraction and more than one is required for a guaranteed effective DNA

extraction. Obviously, the absolute compatibility of primers is required for the DNA obtained to be useful. In some cases here the primers did not work well with a given species and required optimisation which eventually led to the loss of DNA material during the process of primer optimisation.

The design of oligonucleotide primers in this work took a very long time because of the lack of published bryozoan sequences available for the design of the first Bryozoa-specific primers.

The primers designed in this study can be separated into those which work with most species, and thus can be used with a broad range of bryozoan species, and primers which were designed specifically for *Alcyonidium* species. When the design of the bryozoan-wide primers was done, the concept of universality was applied in the sense that the annealing sites were picked up in such a way so that they were conserved among as many bryozoan 18S sequences as possible. Given the wide universality of the main set of primers designed here when tested with over 20 species of Bryozoa, it is likely that these primers should work with other bryozoan species and might be useful for future studies which extend this work.

Although there was no need for designing Ctenostomata-specific primers, *Alcyonidium* species did not work with the main sets of primers. Therefore *Alcyonidium*-specific primers had to be designed. This was complicated by the actual lack of *Alcyonidium* sequences available upon which the new primers could be built and apparent differences between *Alcyonidium* and other Ctenostomata sequences. For instance, standard bryozoan primers worked very well with *Flustrellidra hispida* (a ctenostome species), and yet did not work with *Alcyonidium* (also a ctenostome).

The sequences obtained here for *Alcyonidium hirsutum*, *A. gelatinosum* and *A. polyoum* were found to be longer than expected (average length was 2168 bp, see Chapter 4 for details). These sequences did not align well with sequences of other bryozoan species. Detailed discussion of the possible problems and causes of the anomalies of *Alcyonidium* species sequences are presented in Chapter 4. Their close affinity to the *Uncultured metazoan* (accession no. AY172989) sequence from GenBank and their failure to align to the secondary structure of 18S rRNA from Bryozoa and other taxa indicate a possible abnormality of the *Alcyonidium* sequences. Based on the above, further work is urgently required for this genus,

especially given that the validity of the only other sequence of *Alcyonidium* deposited in NCBI GenBank is questionable.

In all cases in this work, DNA extraction for *Alcyonidium* species was done using the method described in Chapter 3. However, in future work, given the lobose nature of colonies of some *Alcyonidium* species (such as *Alcyonidium diaphanum*) and their size, it should be possible to evaluate another method of DNA extraction, namely from the colony tissue, should further attempts based on the larval DNA extraction fail. Alternatively, a method of combining several larvae and using commercial DNA extraction kits could be used to increase DNA recovery (such as the method described in Waeschenbach *et al.* 2006). This way at least one sequence of *Alcyonidium* can be obtained after which it can be used to build further genus-specific primers.

Sequence contamination has to be taken very seriously, not only in ongoing studies but also in considering sequences which have previously been deposited in GenBank or other databases. A recent study (Ashelford *et al.* 2005), found 5% of errors with more than 60% of these being chimeras out of the sample of sequences obtained from the 16S rRNA database. The tool developed by the same authors, MALLARD (Ashelford *et al.* 2006), allows an evaluation of 16S sequences and identification of possible suspect errors in the examined sequences and eventually chimeras.

In this work, the bryozoan 16S sequences submitted to the NCBI Genbank by Dick *et al.* (2000) were examined. Two, *Scrupocellaria varians* (accession no. AF156291) and *Electra pilosa* (accession no. AF161176), were found to be anomalous. Unfortunately, the above software currently does not support identification of 18S rRNA sequence anomalities. However, based on the above findings and findings of other authors, it is necessary to exclude those sequences which have been found to be contaminants and further evaluate the suspect ones.

As an aid for further research work a separate database of proven valid sequences (listing corresponding species and their accession numbers) perhaps can be established on the website of the International Bryozoology Association to simplify further sequence tracking.

7.2 Secondary structure alignment

The alignment of sequences of 18S rRNA using secondary structure considerations was discussed in detail in Chapter 4, specifically the importance of using such an alignment as opposed to an automated computer-assisted alignment. Because of the highly variable substitution rates in some regions of 18S rRNA throughout the gene and their affect on the results of phylogenetic reconstruction (Abouheif *et al.* 1998; Xia *et al.* 2003), secondary structure has to be considered during alignment.

In this work alignment was done using already established secondary structure data from the ERRD (see Chapter 4). The sequences which were as closely related to bryozoans as possible were selected. The alignment (which took a large amount of time because of lack of automated methods) was done in two stages: first by computer assisted alignment (ClustalX) of the conserved motifs, followed by a manual evaluation of each of the individual rRNA helix loop segments.

Currently no software exists which performs an automated secondary structure alignment of 18S rRNA and most of the process has to be performed "by hand" by the investigator. This process is highly time consuming, tedious and prone to errors. In the current study it was aggravated by the lack of a published secondary structure model of bryozoan 18S rRNA. Because of the problems encountered with *Alcyonidium* sequences these were not included in the secondary structure alignment. Originally the possibility of large insertions into the rRNA sequences of *Alcyonidium* was suspected. Similar insertions were reported in the hypervariable regions of some insects (Kwon *et al.* 1991; Gillespie *et al.* 2005). However, apart from length differences, the sequences of *Alcyonidium* species did not align well at all to any tested metazoan 18S rRNA sequences throughout most of their lengths. The only similarity to other sequences was noted for several very conserved regions corresponding to the stems of the helices.

Because of the lack of published secondary structures of 18S rRNA for Bryozoa, once the alignment of the sequences was performed the sequence which covered all regions of the 18S gene (helices 1 to 50) was used to reconstruct a bryozoan rRNA secondary structure model. This structure model of *Bugula turbinata* is presented for future reference should other 18S sequences become available in future studies. The drawing of the secondary structure was done using RNAViz software (*de Rijik* et al. 2003), now unfortunately not further developed

and unsupported. In addition, assistance from the software script supplied by M. Telford (personal communication) was used.

It is hoped by the author that with the availability of the 18S rRNA secondary structure model of *Bugula turbinata,* more studies can benefit from it by using it as a reference when other bryozoan sequences are aligned. It is also hoped that from a purely pragmatic point of view this secondary structure will enable other researchers to speed up the tedious manual process of secondary structure alignment.

Currently there are several projects dealing with rRNA folding algorithms being developed. The most well known, which has been in existence for several years and which was used in this work, is the algorithm implemented in the Mfold program (Zuker and Steiger 1981). The algorithm of this program is based on the minimum free energy foldings. The same algorithm is implemented in the RNAfold package (Hofacker 2003). This widely used algorithm has been criticised because of its shortcomings, namely its inability to correctly fold sequences larger than 400 bp, complete disregard of tertiary interactions (*i.e.* pseudoknots) and general drawbacks of the method in comparison with the comparative approach of the conserved elements of the several RNA sequences (Reeder *et al.* 2006). Several other algorithms and software packages are being developed which show some improvement in their ability to overcome the drawbacks of Mfold and which may be able to assist in secondary structure folding of new bryozoan sequences. Using these new packages should improve the reliability of the automated method and speed up the process of secondary structure alignment, which is currently done by hand. Unfortunately, these new packages are currently limited in some way and mostly suitable only for shorter RNAs sequences. Some of these packages are discussed in Chapter 5 and include: RNAshapes (Steffen *et al.* 2006), which allows a quicker selection of the optimal structures based on the abstract shape analysis of the folded structure; pknotsRG (Reeder *et al.* 2004) which allows detection of simple pseudoknot structures based on two helices; and RNAforester (Höchsmann *et al.* 2003) which allows multiple sequences alignment based on the local similarities of the RNA structures.

An ARB project environment (Ludwig *et al.* 2004), mentioned in Chapter 5, was originally considered for the secondary structure alignment. Despite its lack of support the software was compiled, installed and configured for use in this study.

However, it is only suitable for comparison and alignment of rRNA sequences against already existing predefined secondary structure alignments. These, however, only exist for 16S rRNA genes for this software, so as a result the software was not used.

7.3 Phylogenetic analysis

One of the main aims of this project was reconstruction of phylogeny of the three bryozoan orders Ctenostomata, Cheilostomata and Cyclostomata based on 18S rRNA, as well as a more detailed investigation of the relationship between families of this complex group.

The study depended on the availability of 18S rRNA bryozoan sequences, not many of which were available in the public databases, and thus part of the study was obtaining these sequences. In total 26 valid sequences were obtained from three orders of Bryozoa and this allowed evaluation of the relationship between the orders, as well as critically review of some of the previous findings of other molecular studies, for instance Dick *et al.* (2000).

Based on the collected data, a full recovery of the Cyclostomata group was obtained; it was observed as a monophyletic clade on all trees using mixed RNA-specific and GTR models. In addition a possible monophyletic clade of Ctenostomata was observed. Unfortunately, because of the problems encountered with the *Alcyonidium* species sequences, the question remains open on the position of Ctenostomata in this study. Also the relationship of Ctenostomata with Cheilostomata was inconclusive based on the trees obtained here. The order Cheilostomata was found to be either paraphyletic, monophyletic or polyphyletic depending on the inclusion or exclusion of the *Alcyonidium gelatinosum* sequence (accession no X91403) from GenBank and the sequence obtained here for *Scruparia chelata*. The inclusion of the *Alcyonidium gelatinosum* sequence (the validity of which has been questioned) in the dataset resulted in breaking the ctenostome clade and making it polyphyletic but leaving a monophyletic cheilostome clade within ctenostomes. Conversely, the removal of this sequence resulted in paraphyletic cheilostomes, which contained monophyletic ctenostomes. These contrasting results once more emphasise the necessity of obtaining valid *Alcyonidium* sequences and possibly more sequences from Ctenostomata.

Anasca is no longer recognised as a separate suborder of Cheilostomata and instead is split into four suborders (as reported in Chapters 1 and 5). However, analysis of 18S rRNA showed a complete recovery of Anasca, regardless of the inclusion or exclusion of the *Alcyonidium gelatinosum* sequence. The above finding of course has to be treated as tentative as it is based on a limited number of sequences. Representatives of only three anascan families (or two according to the unpublished D.P.Gordon Cheilostomata Treatise classification) were used—six species belonging to three genera. *Scruparia chelata* which was formerly included in the suborder Anasca but is now separated in the suborder Scrupariina[42], is not considered here as part of recovered Anasca clade due to its wide separation on all trees from the other three "anascan" suborders. For any further studies it would be very interesting to obtain as many different representatives from other former anascan families as possible to test for their monophyletic grouping as observed in this work.

Several trees were evaluated in the study; they differed in sequence alignment method, evolutionary model and inclusion or exclusion of particular taxa. After review two trees were considered to be the most favoured—these trees are shown as Figure 50 and 52 in Chapter 5. The only difference between them is the exclusion and inclusion respectively of the *Alcyonidium gelatinosum* and *Bugula plumosa* sequences. Despite this, the difference in tree topology was considerable (as described above). Because of the uncertainty with the sequence of *Bugula plumosa* and the suspect *Alcyonidium gelatinosum* sequence (see Chapter 5 for detailed discussion), the tree shown in Figure 52 was considered less reliable and thus most of the emphasis below is given to the tree displayed in Figure 50 (redrawn below – Figure 69).

[42] As per D.P.Gordon working Treatise (personal communication). However, despite the new classification proposed in the Treatise of D.P. Gordon, Anasca is still often used and includes suborders Malacostega, Inovicellata, Scrupariina and one infraorder from suborder Neocheilostomina.

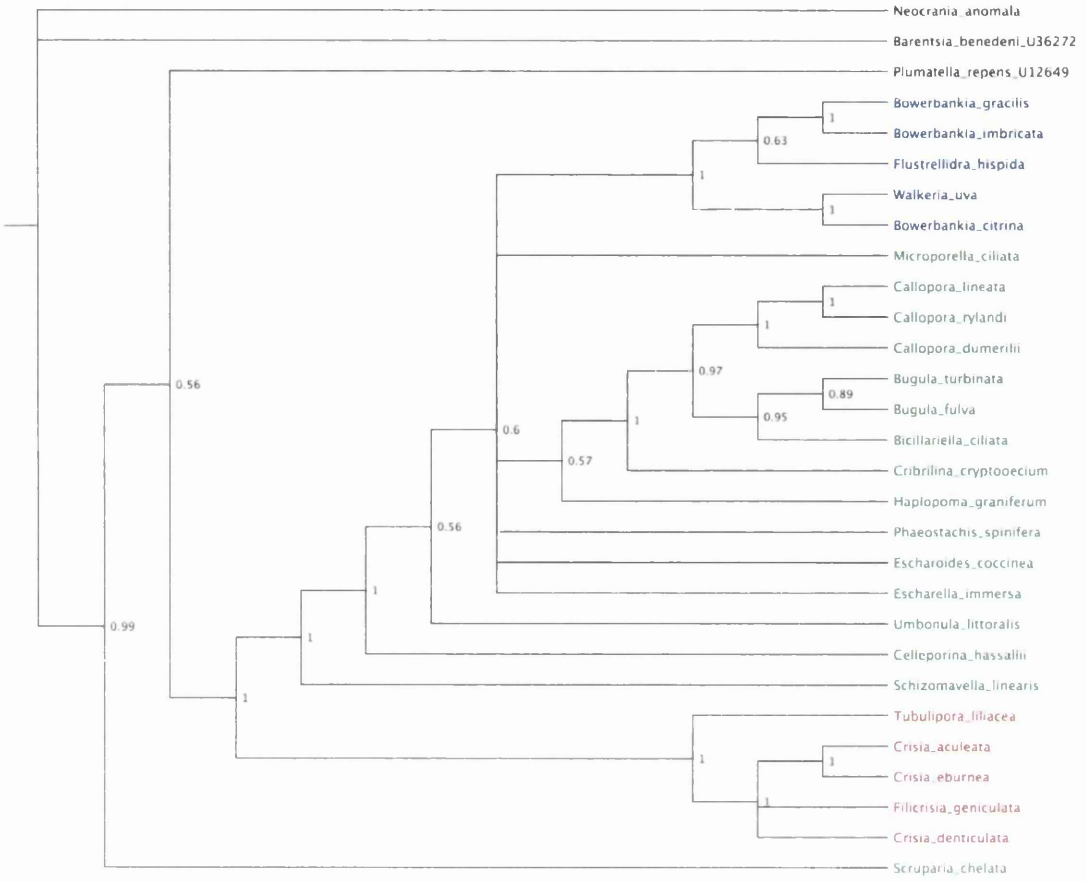


Figure 69 The most favoured bryozoan tree. Bayesian 50% majority rule consensus tree based on the RNA16BHKY+I+Γ and GTR+I+Γ models split between two partitions. Cheilostomata – green; Ctenostomata – blue; Cyclostomata – red; the outgroup and Phylactolaemata – black. Node labels indicate posterior probabilities.

This tree showed monophyletic cyclostomes with a sister clade comprising cheilostomes (excepting *Scruparia chelata*) plus ctenostomes. The above relationships between stenolaemates (cyclostomes) and gymnolaemates (ctenostomes + cheilostomes) contrasts with the notion that ctenostomes are ancestral to stenolaemates (Larwood and Taylor 1979). In particular, ctenostomes are believed to be paraphyletic and have cyclostomes as well as cheilostomes nested within them (Todd 2000). Notable is the low posterior probability (below 0.5; not indicated on the tree) and thus the resulting polytomy between Ctenostomata and Cheilostomata. However, regardless of the polytomy the support for the sister clades (Cyclostomata and Cheilo-Ctenostomata) is very high and would leave ctenostomes within the cheilostomes regardless of how this polytomy would be resolved.

Further, relationships within the cheilostomes are in contradiction with what is commonly believed. A complete recovery of Anasca within the paraphyletic

Ascophora is in contradiction with the paradigm of the anascan cheilostomes giving rise to the more complicated and advanced ascophorans (Gordon 2000; Ryland 1970). Evidence from morphological and palaeontological data that ascophorans are nested within paraphyletic cribrimorphs (Gordon 2000) is in direct contradiction to the findings here. Nesting of the monophyletic Anasca within Ascophora is hard to explain, and definitely requires more sequences from 18S rRNA and possibly other genes to test. As noted above, the introduction of the *Alcyonidium* sequence into this tree broke the paraphyletic cheilostomes topology, but did not change the relationship between anascans and ascophorans.

7.3.1 Stratigraphic congruence

For the purpose of assessing the stratigraphic consistency between the phylogenetic tree and the palaeontological record, three indices which are commonly used for that purpose were calculated for the most favoured tree (Figure 69). These were Stratigraphic Consistency Index (SCI) (Huelsenbeck 1994), Relative Congruence Index (RCI) (Benton and Storr 1994) and, its derivative, Gap Extension Ratio (GER) (Wills 1998). All three indices were calculated using software – Ghosts2 (Willis 1998). The RCI and GER indices are almost identical in their calculation, however they give slightly different values. The significance levels for all three indices were calculated using a permutation test proposed by Huelsenbeck (1994). The RCI and GER indices were calculated in addition to the SCI as it was shown that RCI values match stratigraphic data better for molecular trees, whereas SCI is more suitable for morphological trees (Benton 1998).

Stratigraphic data were taken from Taylor (1993) for Bryozoa, Smirnova (1997) for Brachiopoda, and Todd and Taylor (1992) for Entoprocta. Because of the lack of genus level information, the stratigraphic ranges were assigned based on the family level—fossil records for most families represented in the tree above are documented in the literature. For instance, all *Callopora* species were assigned the stratigraphic age based on the family they belong to (*i.e.* Calloporidae), in this case ALB to Recent. For those families with no fossil record the range was assigned as Recent to Recent.

The results of the tests are given in Table 19 below. The indices were calculated for the most congruent tree displayed above as well for the tree that had *Alcyonidium gelatinosum* and *Bugula plumosa* sequences added to it. This was done

as there was a considerable difference in the topography of Ctenostomata when these sequences were added.

Table 19 Results of stratigraphic congruence tests for two trees based on RCI, GER and SCI indices. The RCT and GER significance results are identical. Significance values are given in percentage and those significant are marked with an asterisk. Tree numbers correspond to those trees shown in Figures 50 and 52.

Tree #	RCI	GER	RCI & GER Significance	SCI	SCI Significance
52	64.90	0.95	0.1*	0.56	0.3*
50	56.46	0.94	0.1*	0.44	1.5*

The results for the randomisation test shown in Table 19 are for 1000 permutations. The significance values that are below 5% are considered to be significant (all values in this case), and essentially indicate that the fit of the cladogram to the stratigraphic data is better than that which would be observed by chance (Huelsenbeck 1994; Wills 1998). Because of the identical calculation of the RCI and GER indices, their significance values are identical and reported together.

The results for the main tree (marked as tree 50 in the table) are lower than expected and inconclusive. The SCI index results obtained here were in line with other published data for the molecular trees, *i.e.* between 0.4 and 0.6 (Benton 1998; Wills, personal communication). However, the SCI metric has to be taken critically though as it was shown to much better suited for the morphological data derived trees than for molecular trees (Benton 1998). The RCI values were slightly lower than observed for the published data, ca 80 (Benton 1998), however, they are still considered to be very good (M.A. Wills, personal communication).

The RCI metrics were shown to much better suited for molecular data congruence to the stratigraphic data in the reviewed literature (Benton 1998). As GER index is identical in its calculation to the RCI index, but simply expressed in a different way (Finarely and Clyde 2002), it is logical to expect that this index performs equally well to the RCI index for molecular data. Values of GER close to 1.0 are possible and the index has an advantage over the RCI in that it can be used to compare different trees. Given the lower than expected values for the SCI and RCI indices obtained here for tree #50, it is clear that there is less agreement with the stratigraphical fossil data than would be desired. The stratigraphic congruency of this tree may change if the topology, (*i.e.* sequence composition) were to change

and of course some change may be expected if more genus-level stratigraphy can be added.

The results of the tree marked as #52 in the table (also see Figure 52), which had sequences of *Alcyonidium gelatinosum* and *Bugula plumosa* added, had higher value for all three indices and, as GER index can be use to compare different trees, this index value was slightly higher for tree #52 than for tree #50. This indicated that tree #52 is in better agreement with the stratigraphical record. It is clear that the alteration of the topology, which was created by the introduction of the *A. gelatinosum* sequence to the tree, creates a tree more congruent with the stratigraphic record. This was caused by the alteration of the topology of the Cheilo-Ctenostomata group. Although the tree with *A. gelatinosum* and *B. plumosa* sequences cannot be treated as the most favoured, because of the uncertainty of these two sequences, it is clear from the stratigraphic congruency indices that addition of more Alcyonidiidae sequences may improve the tree fit to the stratigraphical data by changing the topography and thus the relationship between Cheilostomata, Ctenostomata and, possibly, within cheilostomes as well.

7.3.2 Scruparia chelata position

As discussed in Chapter 5 the position of *Scruparia chelata* was unexpected. In all but one case, *Scruparia chelata* was placed at the root of the bryozoan tree including the Phylactolaemata species. However, in one case, when the *Alcyonidium gelatinosum* sequence was added to the tree, the *Scruparia chelata* sequence appeared as a sister branch to the Cyclostomata clade. As in previous cases given the uncertain status of the *Alcyonidium gelatinosum* sequence, the result is inconclusive and further investigation is required. Due to the difficulty of obtaining reproducing colonies of this species in South Wales further samples can perhaps be sought in other locations. This species is distributed around Western Europe and the Mediterranean (Hayward and Ryland 1998). In addition, another representative of this genus – *Scruparia ambigua* – may be used.

When adding more sequences to the phylogenetic analysis, a balance has to be reached in relation to taxon sampling. The balance has to be between the number of taxa used for the analysis and the length of the sequences (Hillis *et al.* 2003; Poe 1998). The issue of taxon sampling is still debated in the literature (Poe 1998; Poe and Swofford 1999; Hillis *et al.* 2003); however, there appears to be a consensus

that adding more taxa to the tree decreases the phylogenetic error. These observations may well explain the disagreement which was observed when for instance *Alcyonidium gelatinosum* sequence was added to the trees in this study.

One of the observations in this study during the computer analysis of the sequence data was an issue related to the computer hardware and thus the speed of calculation. The complexity of the data set with the addition of the RNA-specific evolution model considerably increased the computation time and placed a special emphasis on the necessity to use up-to-date hardware. Whilst a supercomputer cluster (similar to the one used in this study) may not be available to every study, the use of outdated hardware or that which is not optimised for the computational tasks is equally inappropriate. Hall (2005, p 73) states, thinking of desktop PCs, that "if [phylogenetic analysis] takes longer than about 14 hours, I will probably choose another method". Thus the calculation time is fitted into the capabilities of an average desktop computer or even a portable computer. Whilst this approach is definitely convenient and time saving, it does not necessarily allow the desired results. For instance, in this study the introduction of the RNA-specific model into the partitioned dataset shifted the convergence generation time from an average of 12 million generations to over 20 million. It was also shown (D.Swofford 2006, unpublished) when evaluating several previously published studies which used MrBayes for their phylogenetic analysis that convergence was not reached because of the insufficient time the analysis was run. This of course in most cases was because of the computational power available to the researchers rather than their personal beliefs. Other researchers have emphasised the necessity of using powerful, up to date hardware in phylogenetics (Sanderson and Shaffer 2002).

The version of MrBayes which was used in this study – 3.2 – currently does not support multithreading[43], however the next release of MrBayes (version 4) is going to add the ability of splitting each chain calculation between different CPUs and thus multithreading the calculation process and ability to speed up considerably the calculation time (F.Ronquist, personal communication). In addition it is planned to include new evolutionary models as well as evaluation of the model space on the go during the analysis. This has many positive implications for further studies and

[43] Multithreading in computing is a type of parallel execution of a process on the computer where the same calculation process can be split between different processors. This functionality has to be supported on the software level.

has to be actively used. For instance in this study, if multithreading had been available it would have been possible to spread the calculation of an individual chain of Bayesian analysis between several cluster nodes and considerably shorten the computation time.

7.3.3 Possibility of using other genes

A relatively new direction in molecular phylogenetics has recently emerged which could be promising for investigation of molecular relationships between larger taxonomic groups, in particular orders of Bryozoa. Instead of using a particular gene sequence, unique genomic rearrangements relating to gene order are investigated (Boore and Brown 1998). One of the greatest advantages of this method is that does not suffer from homoplasy sometimes encountered by other methods (Adoutte *et al.* 2000). However, this method has a certain drawback in that it requires that the genes in question be present in all groups of organisms investigated and this may be difficult to find. One recent study, which investigated the possibility of using Hox gene clusters for inferring metazoan phylogeny (Halanych and Passamaneck 2001), emphasised the advantages of this method but also pointed out that this method requires much higher technological involvement and time.

Recently the mtDNA genome was shown to give some promise in relation to gene order (Boore and Brown 1998). This method of course requires sequencing a complete mitochondrial genome, with relative arrangements of genes recorded. Currently there are 40 species of Lophotrochozoa in which the complete mtDNA genome has been published (Valles and Boore 2006). Recently, the complete mitochondrial genome of *Flustrellidra hispida* was published (Waeschenbach *et al.* 2006), adding a first bryozoan to the list of completely sequenced taxa.

7.4 Confocal microscopy and larval morphology

In this study, an evaluation of a microscopy method was performed to test for a relatively resource efficient and quick method of bryozoan larvae imaging.

The method of confocal laser microscopy tested here showed good results which sometimes enabled a more detailed larval morphological character evaluation compared to light microscopy methods. The main purpose of the method was to assess larval types based on the system proposed and described by Zimmer and Woollacott (1977).

During the method evaluation, some drawbacks were observed. These are linked to that fact that larvae have to be stained and visualised alive. This has two implications. Firstly it is almost impossible to preserve larvae for later imaging. In other words material collection and larval extraction has to be done immediately prior to microscopic imaging. This has, in turn, two associated problems. Firstly, a laboratory equipped with facilities for imaging (i.e. CLS microscope) has to be readily available when larval collection is done, and secondly it is not possible (or highly resource demanding) to transport larvae from a remote site. For instance it would be not practical to do field work in the Mediterranean sea and larval imaging in a UK laboratory. A second implication of the live larvae imaging is their movement – larvae have to be 100% immobilised prior to imaging (see Chapter 6 for examples). Despite the above drawbacks the method offers a quick away of larval examination in suitable conditions and requires much less time and preparation than SEM.

In a recently published work (Santagata 2008) exploring evolutionary relationships and significance of the ciliary fields and musculature of bryozoan larvae, a great diversity of structures within the studied taxa was observed. The method employed by Santagata (2008) was based on light microscopy, SEM and confocal laser imaging. However, unlike the confocal method described in this work, Santagata (2008) employed fixed staining for a detailed examination of larval musculature. In total seventeen species were evaluated by Santagata and one of the most interesting implications is that four new larval types were described in addition to those designated by Zimmer and Woollacott (1978). The new larval types are one for Cyclostomata (cycloform) and three for Ctenostomata (nolelliform, sundanelliform and aeverrilliform). This adds very interesting information to the larval type diversity discussed in Chapter 6 and expands on the larval type system introduced by Zimmer and Woollacott (1978). Obviously, the four new types of larvae have to be considered when larval identification is performed in future. However, the identification of the new larval types by Santagata (2008) was based not only on the shape and position of the surface elements, such as corona or apical disk for instance, but also on the actin staining with Phalloidin, *i.e.* internal musculature. This may hinder or limit the possibility of larval type identification if the cell-permeant stains cannot be acquired for fibrous actin. The finding of these four new types raises questions about larval type

diversity in general and how many new types of larvae can be found should a large scale study be performed.

One of the problems associated with fluorescence microscopy and discussed in Chapter 6 was rather rapid deterioration of the fluorochrome and inability to use anti-fading agents due to their apparent toxicity to the live larvae. Recently introduced methodology – quantum dots (artificiality created nanometre-size semiconductors) – is promising to overcome this problem. Quantum dots being semiconductors can be adjusted to emit any required spectrum, emission of which is triggered by photon stimulation (Claxton *et al.*[no date]). Unlike ordinary fluorochromes quantum dots appear to have a very long photostability and thus could potentially solve the problem of photo bleaching which was encountered in this study.

The method tested in this study can work with any cell-permeant fluorochrome. Based on this method an evaluation of larval types from a wide array of species can be performed and the gathered information can be used for the future taxonomic study of Bryozoa where larval morphology can be taken into account. A phylogenetic analysis can be performed using larval morphological and molecular characters combined in one dataset and thus giving a tree which will take into account both morphological and molecular information. The results of these findings can be evaluated in respect to the currently accepted systematics of Bryozoa based on adult morphological characters.

APPENDIX A

Here 18S rDNA sequences of all species obtained in this study are given in FASTA format. Sequences of the *Alcyonidium* species are given at the end for the record only as their validity is uncertain. The sequence of *Bugula plumosa* was found during the phylogenetic reconstruction to present problems (see Chapter 5) and therefore cannot be positively identified as a *Bugula* species and shall not be submitted to the GenBank. All valid sequences can be supplied in a digital format upon request from the author.

>*Bicellariella ciliata*

CGAAGGTTGCGGCCAGATAGCCATGCATGTCTAAGTGCAAGCCGCGTATGCGGCGAGA
CTGCGGACGGCTCATTAAATCGGTTATGAATCCACTGGGGCCAGACTCACCCGTGGATA
ACGTGCGGTAACTCCGGTGCTAATACATGCAACAAGGCCTTGACCCCGTCCTCGGGCGG
GGGAAGGGCGCACTTATTAGGCGAAAACCAACGGCCGGCCTCGGCCGGCCTTGGTGGA
CGACACCCGAGTAATTGCCGCCGATCGCACGGCCTCAGCGCCGGCGACGCCTGCACCGA
GTTTCTGATCTATCATGCTGTCGACGGTTGGTGCTATGCCAACCGTGGCGTTGACGGATA
ACAGAGAATCTGGGTTCGATTCTGGAGAGGCCGCATGAGAAACGGCGACCACTTCCAA
GGAAGGCAGCAGGCACGCAAATTACCCACTTCGGACACCGAGAGGTAGTGACGAACAA
TACCGATGCGGCGCACTTACGTGTCGCCGTAATCGGAATGAGTACACTTCAAATCCTTT
AACGAGGATCCACTGGAGGGCAAGCCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCA
GCAGCGTATATTTATATTGCTGCGGTTAAAAAGCTCGTAGTTGAAGCGTCAGACGCGGG
AGGGCGGACGGCCGCACCGGTCGCTCTCGCCCGACCGAGCCAGGCGCGGAGGGCGCGC
GTCGCTTGCACTTCACCGTGTGGGCGCGCCGCCGCTGCACGTTCACCTTGAAGAAATTG
AACCGCTTAGAGGGGGCGAGCAGCTTGAACAGCTCAGCATGGTATGATGGAACATGGG
CTTGTACTCATTTTGTTGGTTAGAGAGTCGGCGAGCCAATGATTAACAGGGACTGCCGG
GGGCATTAGTACTCGGACGGGAGAGGTGAAATTCAAGGATCGTCCGAAGACTTCCTACT
GCGAAAGCATTTGCCAAGAATGTTTTCATTAATCAAGAACGAAAGTCAGAGGTGCGAA
GGCGATCAGATACCGCCGTAGTTCTGACCGTAAACGCTGCCAACCGGCAATTGGGCGCA
CTTAGCAATAAGTTGCCGCCAGCAGTTTCTTCTGCCGGGGAAACCAGAGTCACTGGGTT
CCGGGGGGAGTATGGTTGCAAAGCTGAAACTTAAAGGAATTGACGGAAGGGCACCACC
AGAAGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAACCTCACCCGGCCCGAAC
ACTGTTATGACAGACAGGTTGAGAGCTCTTTCTCGATTCAGTGGTTGGTGGTGCATGGC
CGTTCTTAGTTCGTGGAGTGATCTGTCTGGTTAATTCCGCTAACGAACGAGACTCTCGCC
TGCTAAATAGACGGCGCCGAGCTTCGGCTGACGGCGACCGCTCGCTTCTTAGAGGGACA
ACCGGCTTTTAGCCGGTTGAAGCGGAGAGCAATAACAGGTCAGTGATGCCCTCAGATGT
TCGGGGCCGCACGCGCGCTACACTGTTTGCATCAGCGTGTTTCTCCCGCGCCGATCGGC
GCGGGCAACCCGTTGAACCGCAAACGTGCTAGGGATCGGAGATTGCAATTGTTCTCCGT
CAACGAGGAATTCCTTGTACTGGCGAGTCATCAGCTCGCGGGGAATCTGTCCCTGCCCT

TTGTACACACCGCCCGTCGCTACTACCGATTGAGTGGTTTAGTGAGGCTCACGGACTGC
GAGCTGTGCCGGCGGCGGCTTCGGCCGTCGTGGCGCAGCGAGCGGAAAGTGAGACGAA
CTTGATCACTCTAGAAGTAAAAGCGACCTGCCAGGGAGTAGTAGTGACCTGGCCGTAGG
AT

>*Bowerbankia citrina*

TGCTTAGTCAGCCGCGCACGCGGCGAGACTGCGGACGGCTCATTAAATCGGTTACGACT
CCGCTGGGGCCAGACTCCTACGTGGATAACGTGCGGTAACTCCGGTGCTAATACATGCA
ACCAGGCTCCGACCGCGTCTTCGGGCGCGGGAAGGGCGCACTTATTAGGCGAAAACCA
ATCGCCGGCCTCCGGGTCGGCGTTGGCGGACGACACCCGAGTAATTGCCGCCGATCGCA
CGGCCTCGAGCCGGCGACGCTTCCATCGAGTTTCTGATCTATCATGCTGACGACGGTTG
GCGCTATGCCAACCGTGGCGTTGACGGATAACAGAGAATCTGGGTTCGATTCTGGAGAG
GCCGCATGAGAAACGGCGACCACTTCCAAGGAAGGCAGCAGGCACGCAAATTACCCAC
TTCGGACACCGAGAGGTAGTGACGAACAATACCGATGCGGCGCGCTCACGCGTCGCCG
TAATCGGAATGAGTACACTTCAAATCCTTTAACGAGGATCCACTGGAGGGCAAGCCTGG
TGCCAGCAGCCGCGGTAATTCCAGCTCCAGCAGCGTATATTTATATTGCTGCGTTTAAA
AAGCTCGTAGTTGAAGCGTCAGACGCGGGAGGGCGGACGGCCGCACTGGTCGCTCTCG
CCCGACCGAGCCAGGCGCAGAGGGCGTCCGTCGCTTGCACTTCGCCGTGTGAGCGCGGC
GCCGCTGCACGTTCACCTTGAGGAAATTGAACCGCTCAGAGGGGGCGAGCAGCTTGCAC
AGCTCAGCATGGTATGATGGAACACGGGCTCGTACTCGTTTTGTTGGTTTTAGAGTCGG
CGAGCCAATGATTAAGAGGGACTGCCGGGGGCATTCGTACTCGGACGGGAGAGGTGAA
ATTCAAGGATCGTCCGAAGACGCCCTACTGCGAAAGCATTTGCCAAGAATGTTTTCATT
AATCAAGAACGAAAGTCAGAGGTGCGAAGGCGATCAGATACCGCCGTAGTTCTGACCG
TAAACGATGCCGACTGGCAATTGGGCGCACTTCTGTAGAAGTTGCTGCCAGCAGCGCGT
CCCGGGAAACCAAAGTCATTGGGTTCCGGGGGGAGTATGGTTGCAAAGCTGAAACTTA
AAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAAC
ACGGGAAACCTCACCCGGCCCGAACACTGTTATGACAGACAGGTTGAGAGCTCTTTCTC
GATTCAGTGGTTGGTGGTGCATGGCCGTTCTTAGTTCGTGGAGCGATCTGTCTGGTTAAT
TCCGATAACGAACGAGACTCTCGCCTGCTAAATAGACGGCGCCGACGTACGCGCGGCG
ACCGCGCAGCTTCTTAGAGGGACAAGCGGCGTTTAGCCGCGTGAAGCCGAGAGCAATA
ACAGGTCAGTGATGCCCTCAGATGTTCGGGGCCGCACGCGCGCTACACTGTTTGCATCA
GCGTGTCTCTCCCTCGCCGGCGGGCGCGGGCAACCCGTTGAAACGCAAACGTGCTAGGG
ATCGGAGATTGCAATTGTTCTCCGTCAACGAGGAATTCCTTGTACTTGCGAGTCATCAGC
TCGCGGGGAATCTGTCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGATTGAGTG
GTTTAGTGAGGCTCGCGGACGGCGCGCGGCAACCGTCGGGTTCGCCCGTCGCGTCGCGC
GCCGGAAGCGAGACGAACCTGATCACTTCTAGGANGTTAAAAGTCGTACGTGTTTAAAA
AAAAAAA

>*Bowerbankia* gracilis

ACCTAGTCCTTGTCTGCAAAGATTAAGCCATGCACGTCTAATGTACAAGCCGCTAAGAC
GNGCGAGACTGCGGACGNGCTCATTAAATCGGTTACGACTCCGCCGGGGACAGACAAC
CCTAGTGGATAACGTGCGGTAACTCCGGTGCTAATACATGCAACGAGGCTCCGACTCGG
CGCGTTTCGGCGCGTCGGGGAAGGGCGCACTTATTAGGCGAAAACCAGTCGGGCGGCC
GTTCGCGGTCGCCCGCCGGTGGACGACACCCGAGTAATTGCCGCCGATCGCACGGCCTC
GAGCCGGCGACGCCATCGACGAGTTTCCGATCTATCATGCTGACGACGGTTGGCGCTGT
GCCAACCGTGGCGTTGACGGATAACAGAGAATCTGGGTTCGATTCTGGAGAGGCCGCCT
GAGAAACGGCGACCACTTCTAAGGAAGGCAGCAGGCACGCAAATTACCCACTTCGGAC
ACCGAGAGGTAGTGACGAACAATACCGATGCGGCGCGCTCACGCGTCGCCGTAATCGG
AATGAGTACACTTCAAATCCTTTAACGAGGATCCACTGGAGGGCAAGCCTGGTGCCAGC
AGCCGCGGTAATTCCAGCTCCAGCAGCGTATATTTATATTGCTGCGTTTAAAAAGCTCGT
AGTTGAAGCGTCAGACACGGAGAGGCGGACGGCCGCACTGGTCGCGCTCGTCTCTCCG
GGCCAGGCGCGGAGGGCGTCCGTCGCTTGCACTTTACCGTGTGAGCGCGGCGCCGCCGC
TCGTTCACCTTGAGGAAATTGAACCGCTCAAAGGAGGCGAGCAGCTCGAACAGCTCAG
CATGGTATGATGCAAGACGGGCTCGTACTCGTTTTGCTGGTTTTAGAGTCGGCGAGCCA
ATGATTAATAGGGACTGCCGGGGGCATTCGTACTCGGACGGGAGAGGTGAAATTCAAG
GATCGTCCGAAGACGCCCGACTGCGAAAGCATTTGCCAAGAATGTTTTCATTAATCAAG
AACGAAAGTCAGAGGTGCGAAGGCGATCAGATACCGCCGTAGTTCTGACCGTAAACTA
TGCCGACCGGCAATTGGGCGCACTTCTGCAGAAGTTGCTGCCAGCAGCATTGCCCGGGA
AACCAGAGTCATTGGGTTCCGGGGGGAGTATGGTTGCAAAGCTGAAACTTAAAGGAAT
TGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAA
AACTCACCCGGCCCGAACACTGTTATGACAGACAGGTCGAGAGCCCTTTCTCGATTCGG
TGGTTGGTGGTGCATGGCCGTTCTTAGTTCGTGGAGCGATTTGTCTGGTTAATTCCGATA
ACGAACGAGACTCTCGCCTGCTAAATAGACGGCGCCACGCGTGCGTGGCAGCCGCTCA
GTTGCTTCTTAGAGGGACAAGCGGCGTTTAGCCGCGGGAAGCGGAGAGCAATAACAGG
TCAGTGATGCCCTCAGATGTTCGGGGCCGCACGCGCGCTACACTGTCCGCATCAGCGTG
TCTGTCCCACGCCGGCCGGCGCGGGCAACCCGTTGAACCGCGGACGTGCTAGGGCTCGG
AGATTGCAATTCTTCTCCGTCAACGAGGAATTCCTTGTACTCGTGGGTCATCAGCTCGCG
GGGAATCCGTCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGATTGAGTGGTTTA
GTGAGGCCCGCGGACCGCGGGCGGCATCCGGCGGGAGACCGTCGCGCCGCCTGCTCGG
GAAGCGGTACGNNACANTGATCACTTCCTAGGTAAGTAACAGCTCGTAACGAGTGATCT

> *Bowerbankia imbricata*

GGNACGTTGTCGCTTGGTCGCGGAGATGAAGCACATGCACGGCTAAGCTACAAGCCGCT
AAGACGGCGAGACTGCGGACGGCTCATTAAATCGGTTACGAACTCCGCCGGGGACAGA
ACACCCTAGTGGATAACGTGCGGTAACTCCGGTGCTAATACATGCAACGAGGCTCCGAC
TCGGCGACGTGTTCGGCGCGTCGGGGAAGGGCGCACTTATTAGGCGAAAACCAGTCGG
GCGGCCGTTCGCGGTCGCCCGCCGGTGGACGACACCCGAGTAATTGCCGCCGATCGCAC
GGCCTCGAGCCGGCGACGCCATCGACGAGTTTCCGATCTATCATGCTGACGACGGTTGG
CGCTGTGCCAACCGTGGCGTTGACGGATAACAGAGAATCTGGGTTCGATTCTGGAGAGG

CCGCCTGAGAAACGGCGACCACTTCTAAGGAAGGCAGCAGGCACGCAAATTACCCACT
TCGGACACCGAGAGGTAGTGACGAACAATACCGATGCGGCGCGCTCACGCGTCGCCGT
AATCGGAATGAGTACACTTCAAATCCTTTAACGAGGATCCACTGGAGGGCAAGCCTGGT
GCCAGCAGCCGCGGTAATTCCAGCTCCAGCAGCGTATATTTATATTGCTGCGTTTAAAA
AGCTCGTAGTTGAAGCGTCAGACACGGAGGGGCGGACGGCCGCACTGGTCGCGCTCGT
CCCTCCGGGCCAGGCGCGGAGGGCGTCCGTCGCTTGCACTTCACCGTGTGAGCGCGGCG
CCGCCGCTCGTTCACCTTGAGGAAATTGAACCGCTCAAAGGAGGCGAGCAGCTCGAAC
AGCTCAGCATGGTATGATGCAAGACGGGCTCGTACTCGTTTTGCTGGTTTTAGAGTCGG
CGAGCCAATGATTAATAGGGACTGCCGGGGGCATTCGTACTCGGACGGGAGAGGTGAA
ATTCAAGGATCGTCCGAAGACGCCCGACTGCGAAAGCATTTGCCAAGAATGTTTTCATT
AATCAAGAACGAAAGTCAGAGGTGCGAAGGCGATCAGATACCGCCGTAGTTCTGACCG
TAAACTATGCCGACCGGCAATTGGGCGCACTTCTGCAGAAGTTGCCGCCAGCAGCATTG
CCCGGGAAACCAGAGTCATTGGGTTCCGGGGGGAGTATGGTTGCAAAGCTGAAACTTA
AAGGAATTGACGGAAGGGCACCCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAA
CACGGGAAAACTCACCCGGCCCGAACACTGTTATGACAGACAGGTCGAGAGCCCTTTCT
CGATTCGGTGGATTGGTGGTGCATGGCCGTTCTTAGTTCGTGGAGCGATTTGTCTGGTTA
ATTCCGATAACGAACGAGACTCTCGCCTGCTAAATAGACGGCGCCGCGCGTGCGCGGCT
GCCGCTCAGTTGCTTCTTAGAGGGACAAGCGGCGTTTAGCCGCGGGAAGCGGAGAGCA
ATAACAGGTCAGTGATGCCCTCAGATGTTCGGGGCCGCACGCGCGCTACACTGTCCGCA
TCAGCGTGTCTGTCCCGCGCCGGCCGGCGTGGGCAACCCGTTGAACCGCGGACGTGCTA
GGGCTCGGAGATTGCAATTCTTCTCCGTCAACGAGGAATTCCTTGTACTCGTGGGTCATC
AGCTCGCGGGGAATCCGTCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGATTG
AGTGGTTTAGTGAGGCCCGCGGACCGCGGGCGGCATCCGGCGGGAGACCGTCGCGCCG
CCTGCTCGGGAAGCGGTACGAACATTGATCACTTCTAGGAAGTAAAAGTCGTAAGCAA
GATTNTT

>*Bugula fulva*

CATTTCCCCATGTTTGGTGCCCCGCACGTCTAANTANANGNCCGCGTANCGCCGGCGAG
ACTGCGGACGGCTCATTAAATCGGTTATGAATCCACTGGGGCCAGACTCACCCGTGGAT
AACGTGCGGTAACTCCGGTGCTAATACATGCAACAAGGCTCCGACCCTGCCTTCGGGCG
GGGGAAGGGCGCACTTATTAGGCGAAAACCAATGGCCCGTTTCGGCGGGCGTTGGTGG
ACGACACCCGAGTAATTGCCGCCGATCGCACGGCCACAGCGCCGGCGACGCCTGCACT
GAGTTTCTGATCTATCATGCTGTCGACGGTTGGTGCTATGCCAACCGTGGCGTTGACGG
ATAACAGAGAATCTGGGTTCGATTCTGGAGAGGCCGCATGAGAAACGGCGACCACTTCT
AAGGAAGGCAGCAGGCACGCAAATTACCCACTTCGGACACCGAGAGGTAGTGACGAAC
AATACCGATGCGGCGCACTTACGTGTCGCCGTAATCGGAATGAGTACACTTCAAATCCT
TTAACGAGGATCCACTGGAGGGCAAGCCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCC
AGCAGCGTATATTTATATTGCTGCGTTTAAAAAGCTCGTAGTTGAAGCGTCAGACGCGG
GAGGGCGGACGGCCGCACTGGTCGCTCTCGCCCGACCGCGCCAGGCGCAGAGGGCGCG
CGTCGCTTGCACTTCGCCGTGTGAGCGCGCCGCCGCTGCACGTTCACCTTGAAGAAATT

GAACCGCTTAGAGGGGGCGAGCAGCTTGAACAGCTCAGCATGGTATGATGGAACATGG
GCTCGTACTCATTTTGTTGGTTAGAGAGTCGGCGAGCCAATGATTAACAGGGACTGCCG
GGGGCATTAGTACTCGGACGGGAGAGGTGAAATTCAAGGATCGTCCGAAGACTTCCTA
CTGCGAAAGCATTTGCCAAGAATGTTTTCATTAATCAAGAACGAAAGTCAGAGGTGCGA
AGGCGATCAGATACCGCCGTAGTTCTGACCGTAAACTATGCCGCCTGGCAATTGGGCGC
ACTTCTGTCGAAGTTGCCGCCAGCAGTATTGCCCGGGAAACCAAAGTCATTGGGTTCCG
GGGGGAGTATGGTTGCAAAGCTGAAACTTAAAGGAATTGACGGAAGGGCACCACCAGG
AGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAACCTCACCCGGCCCGAACACTG
TTATGACAGACAGGTTGAGAGCTCTTTCTCGATTCAGTGGTTGGTGGTGCATGGCCGTTC
TTAGTTCGTGGAGCGATCTGTCTGGTTAATTCCGATAACGAACGAGATTCTTGCCTGCTA
AATAGACGGCGCCGACTCGTGGCGGCGACCGCTAGCTTCTTAGAGGGACAAGCGGCTTT
TAGCCGCGTGAAGCTGAGAGCAATAACAGGTCAGTGATGCCCTCAGATGTTCGGGGCC
GCACGCGCGCTACACTGTTTGCATCAGCGTGTTTCTCCCGCGCCGATCGGCGTGGGCAA
CCCGTTGAACCGCAAACGTGCTAGGGATCGGAGATTGCAATTGTTCTCCGTGAACGAGG
AATTCCTTGTACTGGCGAGTCATCAGCTCGCGGGGAATCTGTCCCTGCCCTTTGTACACA
CCGCCCGTCGCTACTACCGATTGAGTGGTTTAGTGAGGCTCACGGACTGCGAGCTGCGC
CAGTCGGCCCCGGCCGTCACGGCGCAGCGAGCGGAAAGTGAGACGAACTTGATCACTT
CTAGGAAGTAAAAGTCGTAACAAGGTTTTCGTAGGTGAACCTCGGGAAGGATCACACCT
AACGAAAAACCTTGTTACGACTTTTATTTCCTAGAAGTGATCAAGTTCGTCTCACTTTCG
CTCGCTGCGCCTGACGTGGGGCCAAAGGGCAGCTGCATTCGGAACTCATAAACGCTCAA
TCGTATAGCAAGGGGGTGGTGTTACAAGGGCAGGGACGATTGCGACTAAGTGACCTCA
GGCGAGTCGTCAGAACTGTGCACTCACGTTTGGTTCACGGTTGCCATTGCGCGGAGAAA
CATAATCACAGATAAAGCGTCGCCGAATCGAAGCGACTGAGTATGAATGTCCTCATGCG
CGACC

>*Bugula plumosa*[44]

TTGTCTTCAAAAAGAGAAAGCCATGCATGTCTAAGTACANNCCGCGCACGCGGCGAGA
CTGCGGACGGCTCATTAAATCGGTTACGACTCCGCTGGGGCCAGACTCCTACGTGGATA
ACGTGCGGTAACTCCGGTGCTAATACATGCAACCAGGCTCCGACCGCGTCTTCGGGCGC
GGGAAGGGCGCACTTATTAGGCGAAAACCAATCGCCGGCCTCCGGGTCGGCGTTGGCG
GACGACACCCGAGTAATTGCCGCCGATCGCACGGCCTCGAGCCGGCGACGCTTCCATCG
AGTTTCTGATCTATCATGCTGACGACGGTTGGCGCTATGCCAACCGTGGCGTTGACGGA
TAACAGAGAATCTGGGTTCGATTCTGGAGAGGCCGCATGAGAAACGGCGACCACTTCC
AAGGAAGGCAGCAGGCACGCAAATTACCCACTTCGGACACCGAGAGGTAGTGACGAAC
AATACCGATGCGGCGCGCTCACGCGTCGCCGTAATCGGAATGAGTACACTTCAAATCCT
TTAACGAGGATCCACTGGAGGGCAAGCCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCN
AGCAGCGTATATTTATATTGCTGCGTTTAAAAAGCTCGTAGTTGAAGCGTCAGACGCGG

[44] This sequence was mislabelled and therefore cannot be properly identified and should not be used – please see special notes about this sequence in the Phylogenetic reconstruction chapter.

GAGGGCGGACGGCCGCACTGGTCGCTCTCGCCCGACCGAGCCAGGCGCAGAGGGCGNN
CGTCGCTTGCACTTCGCCGTGTGAGCGCGCCGCCGCTGCACGTTCACCTTGAAGAAATT
GAACCGCTTAGAGGGGGCGAGCAGCTTGAACAGCTCAGCATGGTATGATGGAACATGG
GCTCGTACTCATTTTGTTGGTTAGAGAGTCGGCGAGCCAATGATTAACAGGGACTGCCG
GGGGCATTAGTACTCGGACGGGAGAGGTGAAATTCAAGGATCGTCCGAAGACTCCCTA
CTGCGAAAGCATTTGCCAAGAATGTTTTCATTAATCAAGAACGAAAGTCAGAGGTGCGA
AGGCGATCAGATACCGCCGTAGTTCTGACCGTAAACGTTGCCGACTGGCAATTGGGCGC
ACTTCTGTCGAAGTTGCCGCCAGCAGTATTGCCCGGGAAACCAAAGTCATTGGGTTCCG
GGGGGAGTATGGTTGCAAAGCTGAAACTTAAAGGAATTGACGGAAGGGCACCACCAGG
AGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAACCTCACCCGGCCCGAACACTG
TTATGACAGACAGGTTGAGAGCTCTTTCTCGATTCAGTGGTTGGTGGTGCATGGCCGTTC
TTAGTTCGTGGAGCGATCTGTCTGGTTAATTCCGATAACGAACGAGACTCTCGCCTGCTA
AATAGACGGCGCCGACGTACGCGCGGCGACCGCGCAGCTTCTTAGAGGGACAAGCGGC
GTTTAGCCGCGTGAAGCCGAGAGCAATAACAGGTCAGTGATGCCCTCAGATGTTCGGGG
CCGCACGCGCGCTACACTGTTTGCATCAGCGTGTCTCTCCCTCGCCGGCGGGCGCGGGC
AACCCGTTGAAACGCAAACGTGCTAGGGATCGGAGATTGCAATTGTTCTCCGTCAACGA
GGAATTCCTTGTACTTGCGAGTCATCAGCTCGCGGGGAATCTGTCCCTGCCCTTTGTACA
CACCGCCCGTCGCTACTACCGATTGAGTGGTTTAGTGAGGCTCGCGGACGGCGCGCGGC
AACCGTCGGGTTCGCCCGTCGCGTCGCGCGCCGGAAGCGAGACGAACTTGATCACTTCT
AGNAAGTAANNNNNNNANNANNTNNCNTAGGTGACCTGCGGAAAGATTCA

>*Bugula turbinata*

TCTGGTTGAGGTCCTGCCAGTAATCATNTGNNGTCNNCANNGNTNAAAGCCATGCATGT
CTAAGTACAAGCCGCGTACGCGGCGAGACTGCGGACGGCTCATTAAATCGGTTATGAAT
CCACTGGGGCCAGACTCACCCGTGGATAACGTGCGGTAACTCCGGTGCTAATACATGCA
ACAAGGCTCCGACCCTGCCTTCGGGCGGGGGAAGGGCGCACTTATTAGGCGAAAACCA
ATGGCTCGCTTCGGCGGGCGTTGGTGGACGACACCCGAGTAATTGCCGCCGATCGCACG
GCCACAGCGCCGGCGACGCCTGCACTGAGTTTCTGATCTATCATGCTGTCGACGGTTGG
TGCTATGCCAACCGTGGCGTTAACGGATAACAGAGAATCTGGGTTCGATTCTGGAGAGG
CCGCATGAGAAACGGCGACCACTTCTAAGGAAGGCAGCAGGCGCGCAAATTACCCACT
TCGGACACCGAGAGGTAGTGACGAACAATACCGATGCGGCGCACTTACGTGTCGCCGT
AATCGGAATGAGTACACTTCAAATCCTTTAACGAGGATCAACTGGAGGGCAAGCCTGGT
GCCAGCAGCCGCGGTAATTCCAGCTCCAGCAGCGTATATTTATATTGCTGCGTTTAAAA
AGCTCGTAGTTGAAGCGTCAGACGCGGGAGGGCGGACGGCCGCACTGGTCGCTCTCGC
CCGACCGCGCCAGGCGCAGAGGGCGCGCGTCGCTTGCACTTTACCGTGTGAGCGCGCCG
CCGCTGCACGTTCACCTTGAAGAAATTGAACCGCTTAGAGGGGGCGAGCAGCTTGAACA
GCTCAGCATGGTATGATGGAACATGGGCTCGTACTCATTTTGTTGGTTAGAGAGTCGGC
GAGCCAATGATTAACAGGGACTGCCGGGGGCATTAGTACTCGGACGGGAGAGGTGAAA
TTCAAGGATCGTCCGAAGACTCCCTACTGCGAAAGCATTTGCCAAGAATGTTTTCATTA
ATCAAGAACGAAAGTCAGAGGTGCGAAGGCGATCAGATACCGCCGTAGTTCTGACCGT

AAACGATGCCGACTGGCAATTGGGCGCACTTCTGTCGAAGTTGCCGCCAGCAGTATTGC
CCGGGAAACCAAAGTCATTGGGTTCCGGGGGGAGTATGGTTGCAAAGCTGAAACTTAA
AGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAACA
CGGGAAACCTCACCCGGCCCGAACACTGTTATGACAGACAGGTTGAGAGCTCTTTCTCG
ATTCAGTGGTTGGTGGTGCATGGCCGTTCTTAGTTCGTGGAGCGATCTGTCTGGTTAATT
CCGATAACGAACGAGATTCTTGCCTGCTAAATAGACGGCGCCGACTTTGTACGGCGACC
GCTAGCTTCTTAGAGGGACAAGCGGCTTTTAGCCGCGTGAAGCTAAGAGCAATAACAG
GTCAGTGATGCCCTCAGATGTTCGGGGCCGCACGCGCGCTACACTGTTTGCATCAGCGT
GTTTCTCCTGCGCCGATCGGCGCGGGCAACCCGTTGAACCGCAAACGTGCTAGGGATCG
GAGATTGCAATTGTTCTCCGTGAACGAGGAATTCCTTGTACTGGCGAGTCATCAGCTCG
CGGGGAATCTGTCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGATTGAGTGGTT
TAGTGAGGCTCACGGACTGCGAGCTGCGCCAGTCGGCCTCGGCCGTCACNGGCGCAGC
GAGCGGAANAGTGAGACGNAACTTGATCACTTCTAGGAAGTAATACGTCGTAAGGCCA
AGTTTTCGTACT

>*Callopora dumerilii*

ACCCGAGCGAGTGGTCAGATAGCATGCATGTCTAAGTACAAGCCGCGTATGCGGCGAG
ACTGCGGACGGCTCATTAAATCGGTTATGACTCCACTGGGGCCAGACTAACCCGTGGAT
AACGTGCGGTAACTCCGGTGCTAATACATGCAACAAGGCTCCGACCCTGTCTTCGGGCG
GGGGAAGGGCGCACTTATTAGGCGAAAACCAATGGCTCGCTTCGGCGGGCGTTGGTGG
ACGACACCCGAGTAATTGCCGCCGATCGCACGGCCACAGCGCCGGCGACGCCTGCACT
GAGTTTCTGATCTATCATGCTGTCGACGGTTGGTGCTATGCCAACCGTGGCGTTAACGG
ATAACAGAGAATCTGGGTTCGATTCTGGAGAGGCCGCATGAGAAACGGCGACCACTTC
CAAGGAAGGCAGCAGGCACGCAAATTACCCACTTCGGACACCGAGAGGTAGTGACGAA
CAATACCGATGCGGCGCACTCATGTGTCGCCGTAATCGGAATGAGTACACTTCAAATCC
TTTAACGAGGATCCACTGGAGGGCAAGCCTGGTGCCAGCAGCCGCGGTAATTCCAGCTC
CAGCAGCGTATATTTATATTGCTGCGTTTAAAAAGCTCGTAGTTGAAGCGTCAGACGCG
GGAGGGCGGACGGCCGCACTGGTCGCTCTCGCCCGACCGAGCCAGGCGCAGAGGGCGC
GCGTCGCTTGCACTTCACCGTGTGAGCGCGCCGCCGCTGCACGTTCACCTTGAAGAAAT
TGAACCGCTTAGAGGGGGCGAGCAGCTTGAACAGCTCAGCATGGTATGATGGAACATG
GGCTCGTACTCATTTTGTTGGTTAGAGAGTCGGCGAGCCAATGATTAACAGGGACTGCC
GGGGGCATTAGTACTCGGACGGGAGAGGTGAAATTCAAGGATCGTCCGAAGACTCCCT
ACTGCGAAAGCATTTGCCAAGAATGTTTTCATTAATCAAGAACGAAAGTCAGAGGTGCG
AAGGCGATCAGATACCGCCGTAGTTCTGACCGTAAACGTTGCCGACTGGCAATTGGGCG
CACTTCTGTCGAAGTTGCCGCCAGCAGTATTGCCCGGGAAACCAAAGTCATTGGGTTCC
GGGGGGAGTATGGTTGCAAAGCTGAAACTTAAAGGAATTGACGGAAGGGCACCACCAG
GAGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAACCTCACCCGGCCCGAACACT
GTTATGACAGACAGGTTGAGAGCTCTTTCTCGATTCAGTGGTTGGTGGTGCATGGCCGTT
CTTAGTTCGTGGAGCGATCTGTCTGGTTAATTCCGATAACGAACGAGATTCTTGCCTGCT
AAATAGACGGCGTCGACTGTGTGCGACGGCCGTTCGCTTCTTAGAGGGACAAGCGGCTT

TTAGCCGCGTGAAGCTGAGAGCAATAACAGGTCAGTGATGCCCTCAGATGTTCGGGGCC
GCACGCGCGCTACACTGTTTGCATCAGCGTGTTTCTCCCGCGCCGGCCGGCGCGGGCAA
CCCGTTGAACCGCAAACGTGCTAGGGATCGGAGATTGCAATTGTTCTCCGTGAACGAGG
AATTCCTTGTACTGGCGAGTCATCAGCTCGCGGGGAATCTGTCCCTGCCCTTTGTACACA
CCGCCCGTCGCTACTACCGATTGAGTGGTTTAGTGAGGCTCACGGACCGCGAGCGGCAC
CCGGCGGCCTCGGCCGTCAAGGCGTCGCGAGTGGGAAGTGAGACGAACTTGATCACTTC
TAGGAAGTAAAAGTCGTAACAAGGTTTTCGAGGTGACTCGGGGAAAGAGGGACCTACC
AAAAAACCTTGTTACGACTTTACTTCCTAGAAGTGATCAGTTCGTCTCATTTCCCTCGCG
ACGCCTGACGCGAGCGCGGTGCGCTGGGGTCGTGAGCTCATAACACTCATCGTAGTAGC
GACGGGGGTGTGTACAGGGCAGGACGAATGCGAGTATGACTCAGTAAGTCTCGAAGTA
AATGACTATCTACGTTGCGTTCGGTGGCCGCCGCGGAAAAACATAATCACAGGTAGCGG
TGCAATTAGGACTCGAGTATTGTCTCGGCATAACGCGTCACAAGGCATCAGCATACTTT
AGCGGGACGCTTTTTGAACAGGTTTA

>*Callopora lineata*

ANATCTAATGTACCAAACCGCATATGCGGCGAGACTGCGGACGGCTCATTAAATCGGTT
ATGACTCCACTGGGGCCAGACTAACCCGTGGATAACGTGCGGTAACTCCGGTGCTAATA
CATGCAACAAGGCTCCGACCCTGTCTTCGGGCGGGGGAAGGGCGCACTTATTAGGCGA
AAACCAATGGCTCGCTTCGGCGGGCGTTGGTGGACGACACCCGAGTAATTGCCGCCGAT
CGCACGGCCACAGCGCCGGCGACGCCTGCACTGAGTTTCTGATCTATCATGCTGTCGAC
GGTTGGTGCTATGCCAACCGTGGCGTTAACGGATAACAGAGAATCTGGGTTCGATTCTG
GAGAGGCCGCATGAGAAACGGCGACCACTTCCAAGGAAGGCAGCAGGCACGCAAATTA
CCCACTTCGGACACCCGGAGAGGTAGTGACGAACAATACCGATGCGGCGCACTCATGT
GTCGCCGTAATCGGAATGAGTACACTTCAAATCCTTTAACGAGGATCCACTGGAGGGCA
AGCCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAGCAGCGTATATTTATATTGCTGC
GTTTAAAAAGCTCGTAGTTGAAGCGTCAGACGCGGGAGGGCGGACGGCCGCACTGGTC
GCTCTCGCCCGACCGAGCCAGGCGCAGAGGGCGCGCGTCGCTTGCACTTCACCGTGTGA
GCGCGCCGCCGCTGCACGTTCACCTTGAAGAAATTGAACCGCTTAGAGGGGGCGAGCA
GCTTGAACAGCTCAGCATGGTATGATGGAACATGGGCTCGTACTCATTTTGTTGGTTAG
AGAGTCGGCGAGCCAATGATTAACAGGGACTGCCGGGGGCATTAGTACTCGGACGGGA
GAGGTGAAATTCAATGCGGATCGTCCGAAGACTCCCTACTGCGAAAGCATTTGCCAAGA
ATGTTTTCATTAATCAAGAACGAAAGTCAGAGGTGCGAAGGCGATCAGATACCGCCGTA
GTTCTGACCGTAAACGTTGCCGACTGGCAATTGGGCGCACTTCTGTCGAAGTTGCCGCC
AGCAGTATTGCCCGGGAAACCAAAGTCATTGGGTTCCGGGGGGAGTATGGTTGCAAAG
CTGAAACTTAAAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAAT
TTGACTCAACACGGGAAACCTCACCCGGCCCGAACACTGTTATGACAGACAGGTTGAGA
GCTCTTTCTCGATTCAGTGGTTGGTGGTGCATGGCCGTTCTTAGTTCGTGGAGCGATCTG
TCTGGTTAATTCCGATAACGAACGAGATTCTTGCCTGCTAAATAGACGGCGTCGACTGT
GTGCGGCGGCCGTTCGCTTCTTAGAGGGACAAGCGGCTTTTAGCCGCGTGAAGCTGAGA
GCAATAACAGGTCAGTGATGCCCTCAGATGTTCGGGGCCGCACGCGCGCTACACTGTTT

GCATCAGCGTGTTTCTCCCGCGCCGGCCGGCGCGGGCAACCCGTTGAACCGCAAACGTG
CTAGGGATCGGAGATTGCAATTGTTCTCCGTGAACGAGGAATTCCTTGTACTGGCGGTC
ATCAGCTCGCGGGGAATCTGTCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGAT
TGAGTGGTTTAGTGAGGCTCACGGACCGCGAGCGGCACCCGGCGGCTTCGGCCGTCAAC
GGCGTCGCGAGTGGGAAGTGAGACGAACCTGATCACTCTTAGGAAGTACAGTCGTAAG
GGGTGAAAATATTCAGNTTCNTCCCTTCCCTCCCACCCGTTGACGGCNAACCCCGGNTC
GCTCCGGTCGTACCTCCTACCCTTTCGNTATTAGCAACGGCGGTGGTTNAAGGGGCAGG
GNACGGACGCATTATCTCTCAG

>*Callopora rylandi*

GNNACNNTNGNCCTNGNNCCGAGATGAAGANCTGCTGTCTAAGTNTNNAGCCGCGTAT
GCGGCGAGACTGCGGACGGNTNATTAAATCGGTTATGACTCCACTGGGGCCAGACTAA
CCCGTGGATAACGTGCGGTAACTCCGGTGCTAATACATGCAACAAGGCTCCGACCCTGT
CTTCGGGCGGGGGAAGGGCGCACTTATTAGGCGAAAACCAATGGCTCGCTTCGGCGGG
CGTTGGTGGACGACACCCGAGTAATTGCCGCCGATCGCACGGCCACAGNCGCCGGCGA
CGCCTGCACTGAGTTTCTGATCTATCATGCTGTCGACGGTTGGTGCTATGCCAACCGTGG
CGTTAACGGATAACAGAGAATCTGGGTTCGATTCTGGAGAGGCCGCATGAGAAACGGC
NACCACTTCCAAGGAAGGCAGCAGGCACGCAAATTACCCACTTCGGACACCGAGAGGT
AGTGACGAACAATACCGATGCGGCGCACTCATGTGTCGCCGTAATCGGAATGAGTACAC
TTCAAATCCTTTAACGAGGATCCACTGGAGGGCAAGCCTGGTGCCAGCAGCCGCGGTAA
TTCCAGCTCCAGCAGCGTATATTTATATTGCTGCGTTTAAAAAGCTCGTAGTTGAAGCGT
CAGACGCGGGAGGGCGGACGGCCGCACTGGTCGCTCTCGCCCGACCGAGCCAGGCGCA
GAGGGCGCGCGTCGCTTGCACTTCACCGTGTGAGCGCGCCGCCGCTGCACGTTCACCTT
GAAGAAATTGAACCGCTTAGAGGGGGCGAGCAGCTTGAACAGCTCAGCATGGTATGAT
GGAACATGGGCTCGTACTCATTTTGTTGGTTAGAGAGTCGGCGAGCCAATGATTAACAG
GGACTGCCGGGGGCATTAGTACTCGGACGGGAGAGGTGAAATTCAAGGATCGTCCGAA
GACTCCCTACTGCGAAAGCATTTGCCAAGAATGTTTTCATTAATCAAGAACGAAAGTCA
GAGGTGCGAAGGCGATCAGATACCGCCGTAGTTCTGACCGTAAACGTTGCCGACTGGCA
ATTGGGCGCACTTCTGTCGAAGTTGCCGCCAGCAGTATTGCCCGGGAAACCAAAGTCAT
TGGGTTCCGGGGGGAGTATGGTTGCAAAGCTGAAACTTAAAGGAATTGACGGAAGGGC
ACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAACCTCACCCGGCC
CGAACACTGTTATGACAGACAGGTTGAGAGCTCTTTCTCGATTCAGTGGTTGGTGGTGC
ATGGCCGTTCTTAGTTCGTGGAGCAATCTGTCTGGTTAATTCCGATAACGAACGAGATTC
TTGCCTGCTAAATAGACGGCGTCGACTGTGTGCGACGGCCGTTCGCTTCTTAGAGGGAC
AAGCGGCTTTTAGCCGCGTGAAGCTGAGAGCAATAACAGGTCAGTGATGCCCTCAGATG
TTCGGGGCCGCACGCGCGCTACACTGTTTGCATCAGCGTGTTTCTCCCGCGCCGGCCGG
CGCGGGCAACCCGTTGAACCGCAAACGTGCTAGGGATCGGAGATTGCAATTGTTCTCCG
TGAACGAGGAATTCCTTGTACTGGCGAGTCATCAGCTCGCGGGGAATCTGTCCCTGCCC
TTTGTACACACCGCCCGTCGCTACTACCGATTGAGTGGTTTAGTGAGGCTCACGGACCG

CGAGCGGCACCCGGCGGCCTCGGCCGTCCAAGGCGTCNGCGAGTGGGNAGTGAGACGA
ACTTGATCACTTTCTAGTAAGTATATCGTCGTAACGAGACCTAGTTCTCC

>*Crisia aculeata*

AAGATTAAGCCATGCATGTCTAAGTACGCATCTTAGAACGGTGAAACCGCGAATGGCTC
ATTAGATCGGTTGTGGTTCCTTGGATCGTACAAATCCTACTCGGATAACTGTGGTAATTC
TAGAGCTAATACGTGCGCAGAGGCCGGACCGCGAGGGATGGCTGCGTTTATTGTCTCAA
AACCAAACCGCGCTCTCGGGCGCGGTTCCTTTGGTGAACCTGGATAACTTTGGGCTGAT
CGCACGGGCTCGGTCCCGGCGACGTATCTTTCAAACGTCCGCCCTATCAACTGTCGATG
GTCGGCGACCTGCCTACCATGGTTGTAACGGGTAACGGGGAATCAGGGTTCGGTTCCGG
AGAGGGAGCATGAGAAACGGCTACCACTTCCAAGGAAGGCAGCAGGCGCGCAAATTAC
CCACTCCCGGCTCGGGGAGGTAGTGACGAAAAATAACAATGCGGGACTCTTTCGAGGC
CCCGTAATTGGAATGAGTACACTTTAAATCCTTTAACGAGGACCAACTGGAGGGCAAGT
CTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAGTAGCGTATATTAAAGTTGTTGCGGT
TAAAAAGCTCGTAGTTGGATCTCAGAGAGGCGTCGGCGGTCGGCCTTCGCGGTCTGACT
GCCTGCGCCTCTCACCGACCGCCGGTCGCGCATGCCCTTCGCTGAGCGTGCGTCGGCTG
CGGCCACGTTTACTTTGAAAAAATTAGAGTGTTCAAAGCAGGCGTCTCTTCGCCCGCAT
ACCCCAGCATGGAATAATGGAATAGGAGGCTGGTTCTATTTTGTTGGTTTTCGGAATCT
AGCCTAATGGTTAAGAGGGACGGCCGGGGGCATTCGTATTGCGGCGTTAGAGGTGAAA
TTCTTAGATCGCCGCAAGACGAACGAGTGCGAAAGCATTTGCCAAGAATGTTTTCATTA
ATCAAGAACGAAAGTCAGAGGCGCGAAGGCGATCAGATACCGCCGTAGTTCTGACCAT
AAACGATGCCAACTAGCGATCGGTCGGAGTTGCTCTAATGACCCGACCGGCAGCTGCCG
GGAAACCAAGCAAGTGTTTGGGTTCCGGGGGGAGTATGGTCGCAAGGCCGAAACTTAA
AGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAACA
CGGGAAAACTCACCCGGCCCGGACACTGTGAGGATTGACAGATTGATAGCTCTTTCTTG
ATTCGGTGGGTGGTGGTGCATGGCCGTTCTTAGTTGGTGGAGCGATTTGTCTGGTTAATT
CCGATAACGAACGAGACTCTAGCCTGCTAAATACGTCGGCGGATCCCCCGCGGTCCGCC
GCACACTTCTTAGAGGGACAAGCGGCGTATAGCCGCACGAGATTGAGCAATAACAGGT
CTGTGATGCCCTTAGATGTTCGGGGCCGCACGCGCGCTACACTGAAGGCATCAGCGTGT
CCTCCCTGCTCCGACAGGAGCGGGAAACCCGTTGAACCGCTTTCGTGCTAGGGATCGGG
GCTTGTAATTGTTCCCCGTGAACGAGGAATTCCCAGTAGGCGCAAGTCATAAGCTTGCG
TCGATTACGTCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGATTGAATGGTTTA
GTGAGGCCCTCGGACCTCCTGCCGGCGACTCGCGAGAGACGCTGGCGAGGGGGGAAGA
CGGTCAAACTTGATCATTTAGAGGAAGTAAAGTCGTAACATGGTATTNCC

>*Crisia denticulata*

GTAGGNATTAATATACCTCTGAGCACGGTGAAACCGCGAATGGCTCATTAGATCGGTTA
TGGTTCCTTGGATCTTACAAATCCTACTTGGATAACTGTGGTAATTCTAGAGCTAATACA
TGCGCAAAGGCCGGACCCGTGAGGGAATGGCTGCAGTTATGTTCTCAAAACCAAACCG
GGCTTGCTCGGTTCGTGTGGTGAATCTGGATAACTTTGAGCTGATCGCACGGGCTTTGAC

CCGGCGACGTATCTTTCAAATGTCCGCCCTATCAACTGTCGATGGTCGGCGACCTGCCTA
CCATGGTTGTAACGGGTAACGGGGAATCAGGGTTCGGTTCCGGAGAGGGAGCATGAGA
AACGGCTACCACTTCCAAGGAAGGCAGCAGGCGCGCAAATTACCCACTCCCGGCTCGG
GGAGGTAGTGACGAAAAATAACAATGCGGGACTCTTTCGAGGCCCCGTAATTGGAATG
AGTACACTTTAAATCCTTTAACGAGGACCAATTGGAGGGCAAGTCTGGTGCCAGCAGCC
GCGGTAATTCCAGCTCCAATAGCGTATATTAAAGTTGTTGCGGTTAAAAAGCTCGTAGT
TGGATCTCAGAGAGGCGTACGCGGTTGGCCTTCGCGGTCTTACTGCGTGCGCCTTTCACC
GACCGCCGGTCGCGCATGCCCTTCGCTGAGCGTGCGTCGGCTGCGGCCACGTTTACTTT
GAAAAAATTAGAGTGTTCAAAGCAGGCGTTTTCTCGCCCGCATACCCCAGCATGGAATA
ATGGAATAGGAGGCTGGTTCTATTTTGTTGGTTTTCGGAATCTAGCCTAATGGTTAAGAG
GGACGGCCGGGGGCATTCGTATTGCGGCGTTAGAGGTGAAATTCTTAGATCGCCGCAAG
ACGAACAAGTGCGAAAGCATTTGCCAAGAATGTTTTCATTAATCAAGAACGAAAGTCA
GAGGCGCGAAGGCGATCAGATACCGCCGTAGTTCTGACCATAAACGATGCCAACTAGC
GATCGGTCGGAGTTGCTCTAATGACCCGACCGGCAGCTGCCGGGAAACCAAGCAAGTCT
TTGGGTTCCGGGGGGAGTATGGTCGCAAGGCCGAAACTTAAAGGAATTGACGGAAGGG
CACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAAACTCACCCGGC
CCGGACACTGTGAGGATTGACAGATTGATAGCTCTTTCTTGATTCGGTGGGTGGTGGTG
CATGGCCGTTCTTAGTTGGTGGAGCGATTTGTCTGGTTAATTCCGATAACGAACGAGAC
TCTAGCCTGCTAAATACGTTGGCGGATCCCCCGCGGTCCGTCGTGAACTTCTTAGAGGG
ACAAGCGGCGTTTAGCCGCACGAGATTGAGCAATAACAGGTCTGTGATGCCCTTAGATG
TTCGGGGCCGCACGCGCGCTACACTGAAGGCATCAGCGTGTCCTCCCTGCTCCGACAGG
AGCGGGAAAATCCGTTGTAACCGCTTTCGTGCTAGGGATCGGGGCTTGCAATTGTTCCC
CGTGAACGAGGAATTCCCAGTAAGCGCAAGTCATAAGCTTGCGTTGATTACGTCCCTGC
CCTTTGTACACACCGCCCGTCGCTACTACCGATTGAATGGTTTAGTGAGGCCCTCGGACC
TCCTGCTGGCGACTCGCAAGAGAAGCCGGCGAGGGGGGTAGTTCGTGTCTCATCTCTGA
GCAAAATGACGTANAGTCGTACTGGGTTAACCCTCTTT

>*Crisia eburnea*

ATGGTTAAGCTACGCATCGTTAGCACGGTGAAACCGCGAATGGCTCATTAGATCGGTTA
TGGTTCCTTGGATCTTACAAATCCTACTCGGATAACTGTGGTAATTCTAGAGCTAATACG
TGCACCCAAGGCCGGACCCCGAGGGATGGCTGCGTTTATTGTCTCAAAACCAAACCGGG
TCTCGCAAGGGGGGCGGTTCTTTGGTGAACCTGGATAACTTTGGGCTGATCGCACGGGC
TTTGACCCGGCGACGTATCTTTCAAACGTCCGCCCTATCAACTGTCGATGGTCGGCGACC
TGCCTACCATGGTTGTAACGGGTAACGGGGAATCAGGGTTCGGTTCCGGAGAGGGAGC
ATGAGAAACGGCTACCACTTCCAAGGAAGGCAGCAGGCGCGCAAATTACCCACTCCCG
GCTCGGGGAGGTAGTGACGAAAAATAACAATGCGGGACTCTTTCGAGGCCCCGTAATT
GGAATGAGTACACTTTAAATCCTTTAACGAGGACCGAGTGGAGGGCAAGTCTGGTGCCA
GCAGCCGCGGTAATTCCAGCTCCACTAGCATATATTAAAGTTGTTGCGGTTAAAAAGCT
CGTAGTTGGATCTCAGAGGGGCGCGTGCGGTTGGCCTCCGTGGTTTTACTGCTCGTGCCT
TTCACCGACCGCTGGTCGCGCGTGCTCTTGACTGAGCGTGCGTCGGCTGCGGCCACGTTT

ACTTTGAAAAAATTAGAGTGTTCAAAGCAGGCGTCTTCTCGCCCGCATATCCCAGCATG
GAATAATGGAATAGGAGGCTGGTTCTATTTTGTTGGTTTTCGGAATCTAGCCTAATGGTT
AAGAGGGACAGCCGGGGGCATTCGTATTGTGGCGTTAGAGGTGAAATTCTTAGATCGCC
GCAAGACGAACGAGTGCGAAAGCATTTGCCAAGAATGTTTTCATTAATCAAGAACGAA
AGTCAGAGGCGCGAAGGCGATCAGATACCGCCGTAGTTCTGACCATAAACGATGCCAA
CTAGCGATCGGTCGGAGTTGCTTTTTTTGACCCGACCGGCAGCTGCCGGGAAACCAAGC
AAGTGTTTGGGTTCCGGGGGGAGTATGGTCGCAAGGCCGAAACTTAAAGGAATTGACG
GAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAAACTC
ACCCGGCCCGGACACTGTGAGGATTGACAGATTGATAGCTCTTTCTTGATTCGGTGGGT
GGTGGTGCATGGCCGTTCTTAGTTGGTGGAGCGATTTGTCTGGTTAATTCCGATAACGA
ACGAGACTCTAGCCTGCTAAATACGTCGGCGGATCCCCTCGCGGTCCGCCGCACACTTC
TTAGAGGGACAAGCGGCGTATAGCCGCACGAGATTGAGCAATAACAGGTCTGTGATGC
CCTTAGATGTTCGGGGCCGCACGCGCGCTACACTGAAGGCATCAGCGTGTCCTCCCTGC
TCCGACAGGAGCGGGAAACCCGTTGAACCGCTTTCGTGCTAGGGATCGGGGCTTGTAAT
TGTTCCCCGTGAACGAGGAATTCCCAGTAGGCGCAAGTCATAAGCTTGCGTCGATTACG
TCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGATTGAATGGTTTAGTGAGGCCC
TCGGACCTCCTGCCGGCGACTCGCGAGAGACGCTGGCGAGGGGGGAAGACGGTCAAAC
TTGATCATTTAGAGGAAGTAAAAGTCGTAACAAGGTTTCCGTAGTGACGCCCGGAGGGG
GGATATACCCCTACGGGAAACCTTGTTACGACTTTTACTTCCTCTAAATGATCAAGTTTG
ACGTCTTCCCCCCTCGCAGCGTCTCTGCGAATCGCGGCGGGAAGGTCCAGGGGCTCCTA
AACATTCNATCGGTAGTAGCGAACGGGGCGGNGGTGTTACAAAGGGCAGGGACGTAAA
TCGANNNNGGCTTTATGAANNTG

>*Escharella immersa*

GCAGTAATCATATGCTTGTCTCAAAGATTAAGCCATGCATGTCTAAGTACAAGCCGCGT
ACGCGGCGAGACTGCGGACGGCTCATTAAATCGGTTATGAATCCACTGGGGACAGACCT
ACCCGTGGATAACGTGCGGTAACTCCGGTGCTAATACATGCAACAAGGCTCCGACCCCG
TCTCCGGGCGGGGGAAGGGCGCACTTATTAGGCGAAAACCAATCGCTCGCTCGCGGGC
GTTGGTGGACGACACCCGAGTAATTGCCGCCGATCGCACGGCCTCCGAGCCGGCGACAC
CTTCACTGAGTTTCTGATCTATCATGCTGTCGACGGTTGGTGCCATGCCAACCGTGGCGT
TAACGGATAACAGAGAATCTGGGTTCGATTCTGGAGAGGCCGCATGAGAAACGGCGAC
CACTTCTAAGGAAGGCAGCAGGCACGCAAATTACCCACTTCGGACACCGAGAGGTAGT
GACGAACAATACCGATGCGGCGCACATTCGTGTCGCCGTAATCGGAATGAGTACACTTC
AAATCCTTTAACGAGGATCCACTGGAGGGCAAGCCTGGTGCCAGCAGCCGCGGTAATTC
CAGCTCCAGCAGCGTATATTTATATTGCTGCGTTTAAAAAGCTCGTAGTTGAAGCGTCA
GACGCGGGAGGGCGGACGGCCGCACTGGTCGCTCTCGCCCGACCGAGCCAGGCGCAGA
GGGCGCACGTCGCTTGCACTTCGCCGTGTGAGCGCGTCGCTACTGCACGTTCACCTTGA
AGAAATTGAACCGCTTAGAGGGGGCGAGCAGCTTGTACAGCTCAGCATGGTATGATGG
AAGACGGGCTCGTACTCATTTTGTTGGTTAGAGAGTCGGCGAGCCAATGATTAACAGGG
ACTGCCGGGGGCATTCGTACTCGGACGGGAGAGGTGAAATTCAAGGATCGTCCGAAGA

CGCCCTACTGCGAAAGCATTTGCCAAGAATGTTTTCATTAATCAAGAACGAAAGTCAGA
GGTGCGAAGGCGATCAGATACCGCCGTAGTTCTGACCGTAAACGATGCCGACTGGCAAT
TGGGCGCACTTCTGTCGAAGTTGCCGCCAGCAGTATTGCCCGGGAAACCAAAGTCATTG
GGTTCCGGGGGGAGTATGGTTGCAAAGCTGAAACTTAAAGGAATTGACGGAAGGGCAC
CACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAACCTCACCCGGCCCG
AACACTGTTATGACAGACAGGTTGAGAGCTCTTTCTCGATTCAGTGGTTGGTGGTGCAT
GGCCGTTCTTAGTTCGTGGAGCGATCTGTCTGGTTAATTCCGATAACGAACGAGACTCTT
GCCTGCTAAATAGACGGCGTCGACTTGAGCCGGCGACCGTTCGCTTCTTAGAGGGACAA
GCGGCTTTTAGCCGCGTGAAGCCGAGAGCAATAACAGGTCAGTGATGCCCTCAGATGTT
CGGGGCCGCACGCGCGCTACACTGTTTGCATCAGCGTGTTTCTCCCGCGCCGGCTGGCG
CGGGCAACCCGTTGAACCGCAAACGTGCTAGGGATCGGAGATTGCAATTGTTCTCCGTG
AACGAGGAATTCCTTGTACTGGCGAGTCATCAGCTCGCGGGGAATCTGTCCCTGCCCTT
TGTACACACCGCCCGTCGCTACTACCGATTGAGTGGTTTAGTGAGGCTCACGGACTGCG
CGCGGCACACGACGCTTTCGGGCGGCGCGTCGCGAGGCGGAAAGTGAGACGAACTTGA
TCACTTCTAGGAAGTAAAAGTCGTAACAAGGTTTTCGTAGTGACGCGGGAAAAGAGATC
ATACGAAAACCTTGTTACGACTTTTACTTCCTAGAAGTGATCAAGTTCGTCTCCTTTCGC
TCGCGACGCGCGCGAAGCGTGTGGTGCGCGCGCAGTCGTGAGCTTCCTAAACACTCATC
GGTAGTAGCGACGGGCGGGTGTGGTACAAACGGCAGGGTAAGATTTCGGTAGCTGATG
ACTGCGGTTAAGGAGTTTCTGTCGGCAGAACATTGCATCTCGCTCATCATTTGATGTACG
GGATTGGCGCGCCCTTAGACCATATGAACGAGTCTGTCGAGGCATCAAATATAGGTGAG
TGTCGTAGTGTCTCAGACGATAGCGATCTACGCAGA

>*Escharoides coccinea*

AGTGCTATAGCCATCAGCGCTTGTGCGGCGAGACTGCGGACGGCTCATTAAATCGGTTA
TGAATCCACTGGGGACAGACTCACCCGTGGATAACGTGCGGTAACTCCGGTGCTAATAC
ATGCAACAAGGCTCCGACCTCGTCTCCGGGCGGGGGAAGGGCGCACTTATTAGGCGAA
AACCAACGGCCGGCTTCGGTCGGCCTTGGTGGACGACACCCGAGTAATTGCCGCCGATC
GCACGGCCTCCGAGCCGGCGACACCTTCACTGAGTTTCTGATCTATCATGCTGACGACG
GTTGGTGCTATGCCAACCGTGGCGTTAACGGATAACAGAGAATCTGGGTTCGATTCTGG
AGAGGCCGCATGAGAAACGGCGACCACTTCTAAGGAAGGCAGCAGGCACGCAAATTAC
CCACTTCGGACACCGAGAGGTAGTGACGAATAATACCGATGCGGCGCACTTACGTGTCG
CCGTAATCGGAATGAGTACACTTCAAATCCTTTAACGAGGATCCACTGGAGGGCAAGCC
TGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAGCAGCGTATATTTATATTGCTGCGTTTA
AAAAGCTCGTAGTTGAAGCGTCAGACGCGGGAGGGCGGACGGCCGCACTGGTCGATCT
CGCCCGACCGAGCCAGGCGCAGAGGGCGCGCGTCGCTTGCACTTTACCGTGTGAGCGCG
CCGCCGCTGCACGTTCACCTTGAAGAAATTGAACCGCTCAGAGGAGGCGTGCAGCTTGT
ACAGCTCAGCATGGTATGATGGAACATGGGCTCGTACTCATTTTGTTGGTTAGAGAGTC
GGCGAGCCAATGATTAACAGGGCACTGCCGGGGGCATTCGTACTCGGCCGGGAGAGGT
GAAATTCAAGGATCGGCCGAAGACGCCCTACTGCGAAAGCATTTGCCAAGAATGTTTTC
ATTAATCAAGAACGAAAGTCAGAGGTGCGAAGGCGATCAGATACCGCCGTAGTTCTGA

CCGTAAACGATGCCGACTGGCAATTGGGCGCACTTCTGTCGAAGTTGCCGCCAGCAGTA
CTGCCCGGGATAACCAAAGTCATTGGGTTCCGGGGGGAGTATGGTTGCAAAGCTGAAA
CTTAAAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTC
AACACGGGAAACCTCACCCGGCCCGAACACTGTTATGACAGACAGGTTGAGAGCTCTTT
CTCGATTCAGTGGTTGGTGGTGCATGGCCGTTCTTAGTTCGTGGAGCGATCTGTCTGGTT
AATTCCGATAACGAACGAGACTCTTGCCTGCTAAATAGACGGCGCCGACTCTCGTCGGC
GACCGTTTGCTTCTTAGAGGGACAAGCGGCTTTTAGCCGCGTGAAGCCGAGAGCAATAA
CAGGTCAGTGATGCCCTCAGATGTTCGGGGCCGCACGCGCGCTACACTGTTTGCATCAG
CGTGTTTCTCCCGCGCCGGCTGGCGTGGGCAACCCGTTGAACCGCAAACGTGCTAGGGA
TCGGAGATTGCAATTGTTCTCCGTGAACGAGGAATTCCTTGTACTGGCGAGTCATCAGC
TCGCGGGGAATCTGTCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGATTGAGTG
GTTTAGTGAGGCTCACGGACTGCGCGCGGCACACGGCGTCTTCGGGAGTCGCGTCGCGA
GGCGGNAAGGTGAGACGAACTTGATCACTTCTAGGAGGTAAAAGTCCTACCAGGGTAA
C

>*Filicrisia geniculata*

GGGGACTCTTGTCCCTTGTCTGCAAAGATTAAGCCATGCATGTCTATGTACGCATCTTAG
CACGGTGAAACCGCGAATGGCTCATTAGATCGGTTATGGTTCCTTGGATCTTACAAATC
CTACTTGGATAACTGTGGTAATTCTAGAGCTAATACATGCGCAAAGGCCGGACCTCGCG
GGATGGCTGCGTTTATTGTCTCAAAACCAAACCGCTTTCGGGCGGTTCCTTTGGTGAATC
TGGATAACTTTGAGCTGATCGCACGGGCTTTGACCCGGCGACGTATCTTTCAAATGTCC
GCCCTATCAACTGTCGATGGTCGGCGACCTGCCTACCATGGTTGTAACGGGTAACGGGG
AATCAGGGTTCGGTTCCGGAGAGGGAGCATGAGAAACGGCTACCACTTCCAAGGAAGG
CAGCAGGCGCGCAAATTACCCACTCCCGGCTCGGGGAGGTAGTGACGAAAAATAACAA
TGCGGGACTCTTTCGAGGCCCCGTAATTGGAATGAGTACACTTTAAATCCTTTAACGAG
GACCAACTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAGTAGCGT
ATATTAAAGTTGTTGCGGTTAAAAAGCTCGTAGTTGGATCTCAGAGGGGCGTTCGCGGT
TGGCCTTCGCGGTCTTACTGCGTGCGCCCTTCACCGACCGCCGGTCGCGCATGCTCTTCA
CTGCGCGTGCGTCGGCTGCGGCCACGTTTACTTTGAAAAAATTAGAGTGTTCAAAGCAG
GCGTCTTCTCGCCCGCATACCCCAGCATGGAATAATGGAATAGGAGGCTGGTTCTATTT
TGTTGGTTTTCGGAATCTAGCCTAATGGTTAAGAGGGACGGCCGGGGGCATTCGTATTG
CGGCGTTAGAGGTGAAATTCTTAGATCGCCGCAAGACGAACAAGTGCGAAAGCATTTG
CCAAGAATGTTTTCATTAATCAAGAACGAAAGTCAGAGGCGCGAAGGCGATCAGATAC
CGCCGTAGTTCTGACCATAAACGATGCCAACTAGCGATCGGTCGGAGTTGCTCTAATGA
CCCGACCGGCAGCTGCCGGGAAACCAAGCAAGTCTTTGGGTTCCGGGGGGAGTATGGT
CGCAAGGCCGAAACTTAAAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCG
GCTTAATTTGACTCAACACGGGAAAACTCACCCGGCCCGGACACTGTGAGGATTGACAG
ATTGATAGCTCTTTCTTGATTCGGTGGGTGGTGGTGCATGGCCGTTCTTAGTTGGTGGAG
CAATTTGTCTGGTTAATTCCGATAACGAACGAGACTCTAGCCTGCTAAATACGTTGGCG
GATCCCTCGCGGTCCGTCTTTAACTTCTTAGAGGGACAAGCGGCGTTTAGCCGCACGAG

ATTGAGCAATAACAGGTCTGTGATGCCCTTAGATGTTCGGGGCCGCACGCGCGCTACAC
TGAAGGCATCAGCGTGTCCTCCCTGCTCCGACAGGAGCGGGAAACCCGTTGAACCGCTT
TCGTGCTAGGGATCGGGGCTTGCAATTGTTCCCCGTGAACGAGGAATTCCCAGTAAGCG
CAAGTCATAAGCTTGCGTTGATTACGTCCCTGCCCTTTGTACACACCGCCCGTCGCTACT
ACCGATTGAATGGTTTAGTGAGGCCCTCGGACCTCCTGCCGGCGACTCGTAAGAGATGC
TGGCGAGGGGGGAAGACGGTCAAACTTGATCATTTAGTAGGTAAGTAAAAGTCGTAAC
AAGGTTTCCGTAGGTGAACCTGCGGAAGATA

>*Flustrellidra hispida*

AGCGCTATGACTCAGACCGCACCCGGCGAGCGAAACTGCGAACGGTCTCATTAAATCG
GTTACGACTCCACTGGTGCCAGACGTCCAGGTGGATAACGTGCGGTAACTCCGGTGCTA
ATACATGCAACAAGCTCCGACCCGGCCGCGAGGCCGGGGAAGGGCGCACTTATTAGGC
GAAAACCAATGCCCGCTCAACGGCGGGCTTTGGCGGACGACACCCGAGTAATTGCCGC
CGATCGCACGGCCTCGCGCCGGCGACGCATCTTTCGAGTTTCTGATCTATCACGCTGAC
GACGGCTGGTGCTATGCCAACCGTGGCGTTTACGGATAACAGAGAATCTGGGTTCGATT
CTGGAGAGGCCGCCTGAGAAACGGCGACCACTTCCAAGGAAGGCAGCAGGCACGCAAA
TTACCCACTTCTGACACAGAGAGGTAGTGACGTGAAATACCGATGCGGCGCGCGCAAG
CGACGTCGCAATCGGAATGAGAACAGTTCAAAACCTTTAACGAGGATCCACTGGAGGG
AAAGCCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAGCAGCGTATATTATTGTTGCT
GCGTTTAAAAAGCTCGTAGTTGGCGTATCGACGCGTGACGGCGGTCGGCCTCGCGGCTC
GCACTGCCCGATCGCGCCATGCGCGAACGTCGCCGACGGTTGCGCTTCGCCGCGTGACC
TCGGCAGCGTCGCACGTTCACCTTGAGGAAACTGAACCGTTCATAGGAGGCGTGCGCGC
TCGAACAATACAGCATGGTATGATGGAACAGGCGCCGGTGGCCGTTTTGTTGGTTTTAG
ACCTCGCCGGCAAATGATTAATAGGGACTGCCGGGGGCATTCGTACTCGGGGGGGAGA
GGTGAAATTCAAGGATCCTCCGAAGACGACCTACTGCGAAAGCATTTGCCAAGAATGTT
TTCATTAATCAAGAACGAAAGTTGGAGGCGCGAAGGCGATCAGATACCGCCGTAGTTCC
AACCGTAAACGATGCCGACTGGCGATCGGAGGCTCTTTAAGTGAAGAAGCTTCCGGCA
GCACCCGGGAAACCAAAGTTATTGGGTTCCGGGGGGAGTATGGTTGCAAAATTGAAAC
TTAAAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCTTGCGGCTTAATTTGACTCA
ACACGGGAAACCTCACCCGGCCCGGACACTGCTATGACAGACAGGTTGAGAGCTCTTTC
TCGATTCAGTGGTTGGTGGTGCATGGCCGTTCTTAGTTCGTGGAGCGATTTGTCTGGTTA
ATTCCGATAACGAACGAGACTCACGCCTGCTAAATAGTGGCGGCCGCTCCGGCGGCTCG
CGCTTACTTCTTAGAGGGACAAGCGGCGCCCAGCCGCGTGAAGCTGTGAGCAATAACA
GGTCAGTGATGCCCTTAGATGTTCGGGGCCGCACGCGCGCTACACTGTCGGCATCAGCG
TGACACCCGCGCCGAAAAGCGCGGGCAACCCGTTGAACCGCCGACGTGCTAGGGATCG
GGGCTTGCAATTGTTCCCCGTGAACGAGGAATTCCTTGTAGGGGCGCGTCATCAGCGCG
CCCCGAATCCGTCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGATCGAATGGTT
TAGTGAGGCCCGCGGACCGCGCGCGCTTCGCCGGGTTCGCCCGGCGCCGCGCGAGCGG
GAAGCGGTGACGTACTTGTACTATTATTACCGAACGTATTACGTTCTTTCAANTGGATAA
AAGTTTCGTTNCGCCTTTCGCCGTGCGCCGGCGGCGGGGG

> *Haplopoma graniferum*

ATATCGTTTTACAGCCGTGGCACACGGACGAGACTGCGGACGGTTCATTAAATCGGTTA
TGACTCCACTGGGGCCAGACTCACCCGTGGATAACGTGCGGCAACTCCGGTGCTAATAC
ATGCAACAAGGCTCCGACCCGCCTCCGGGCGGGGAAGGGCGCACTTATTAGGCTAAAA
CCAATCGCCCGCCTCGGCGGGCGTTGGTGGACGACACCCGAGTAATTGCCGCCGATCGC
ACGGCCTCAGAGCCGGCGACGCCTACACTTAATTTCTGATCTATCATGCTGACGACGGT
TGGTGCTATGCCAACCGTGGCGTTAACGGATAACAGAGAATCTGGGTTCGATTCTGGAG
AGGCCGCATGAGAAACGGCGACCACTTCCAAGGAAGGCAGCAGGCACGCAAATTACCC
ACTTCGAACACCGAGAGGTAGTGACGAACAATACCGATGCGGCGCACATACGTGCCGC
CGTAATCGGAATGAGTACACTTCAAACCTTTTAACGAGGATCCACTGGAGGGCAAGCCT
GGTGCCAGCAGCCGCGGTAATTCCAGCTCCAGCAGCGTATATTTATATTGCTGCGTTTA
AAAAGCTCGTAGTTGAAGCGTCAGACGCGGGAGGGCGGACGGCCGCACTGGTCGCTCT
CGCCCGACCGAGCCAGGCGCAGAGGGCGCGCGTCGCTTGCACTTCGTCGTGTGAGCGCG
CCGCCGCTGCACGTTCACCTTGAAGAAATTGAACCGCTTAGAGGGGGCGAGCAGCTTGT
ACAGCTCAGCATGGTATGATGGAACATGGGCTCGTACTCATTTTGTTGGTTAGAGAGTC
GGCGAGCCAATGATTAACAGGGACTGCCGGGGGCATTCGTACTCGGACGGGAGAGGTG
AAATTCAAGGATCGTCCGAAGACGTCCAACTGCGAAAGCATTTGCCAAGAATGTTTTCA
TTAATCAAGAACGAAAGTCAGAGGTGCGAAGGCGATCAGATACCGCCGTAGTTCTGAC
CGTAAACGATGCCGACTGGCAATTGGGCGCACTTCTGTCGAAGTTGCCGCCAGCAGTAT
TGCCCGGGAAACCAAAGTCATTGGGTTCCGGGGGGAGTATGGTTGCAAAGCTGAAACTT
AAAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAA
CACGGGAAACCTCACCCGGCCCGAACACTGTTATGACAGACAGGTTGAGAGCTCTTTCT
CGATTCAGTGGTTGGTGGTGCATGGCCGTTCTTAGTTCGTGGAGCGATCTGTCTGGTTAA
TTCCGATAACGAACGAGACTCTTGCCTGCTAAATAGACGGCGTCGACATAGCGCGGCGA
CCGTTCGCTTCTTAGAGGGACAAGCGGCTTTTAGCCGCGTGAAGCTGAGAGCAATAACA
GGTCAGTGATGCCCTCAGATGTTCGGGGCCGCACGCGCGCTACACTGTTTGCATCAGCG
TGTTCTTCCCGCGCCGGCTGGCGCGGGCAACCCGTTGAACCGCAAACGTGCTAGGGATC
GGAGCTTGCAATTGTTCTCCGTGAACGAGGAATTCCTTGTACTGGCGAGTCATCAGCTC
GCGGGGAATCTGTCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGATTGAGTGGT
TTAGTGAGGCTCACGGACTGCGAGCGACATCCGGCGTTCGCGTCGCGTCGCGAGCGGAA
AGTGAGACGAACTTGATCACTTCTAGGAAGTAAAAGTCGTAACAAGGTTTTCGTAGGTG
AACCTGCGGGAAGGATCCTACGAAAACTTGTTACGACTTTTACTTCCTAGAAGTGATCA
GTTCGTCTCACTTCCGCTGCGAGCAAGCTAGCGGAGTCCTGCATCGTGAGCCTCATAAC
ACTCATCGGTAGTAGACGGGCGGTGTGTACAAGGGCAGGACGAATGCGCGTGATGGAT
CTTCGAGAAGTCTCGTACAAACATGAGCACCTACATTGCTTCAGGATTGTCGAGCGGCG
TAAACACTAGTACCGGTAGACGTGCCATCAGGGACGCGTAATGTCTCGCTACGTCTAGA
GAGGCCCGTTACCATAGGAAAGCTATGTACAACTCCTAAGCTGCCTGCGTATACGGA

>*Microporella ciliata*

TTCTGGTATCATATGCTTGTCTCAAAGATTAAGCCATGCATGTCTAAGTACAAGCCGCGT
ATGCGGCGAGACTGCGGACGGCTCATTAAATCGGTTATGAATCCACTGGGGACAGACTC
ACCCGTGGATAACGTGCGGTAACTCCGGTGCTAATACATGCAACAAGGCTCCGACCTCG
TCTTCGGGCGGGGGAAGGGCGCACTTATTAGGCGAAAACCAATCGCTCTCTGAGCGTTG
GTGGACGACACCCGAGTAATTGCCGCCGATCGCACGGCCTCCGAGCCGGCGACACCTGC
ACTGAGTTTCTGATCTATCATGCTGACGACGGTTGGTGCTATGCCAACCGTGGCGTTAAC
GGATAACAGAGAATCTGGGTTCGATTCTGGAGAGGCCGCATGAGAAACGGCGACCACT
TCCAAGGAAGGCAGCAGGCACGCAAATTACCCACTTCGGACACCGAGAGGTAGTGACG
AACAATACCGATGCGGCGCACTTACGTGTCGCCGTAATCGGAATGAGTACAATTCAAAT
CCTTTAACGAGGATCAACTGGAGGGCAAGCCTGGTGCCAGCAGCCGCGGTAATTCCAGC
TCCAGCAGCGTATATTTATATTGCTGCGTTTAAAAAGCTCGTAGTTGAAGCGTCAGACG
CGGGAGGGCGGACGGCCGCACAGGTCGCTCTCGCCCGACCGAGCCAGGCGCAGAGGGC
GCGCGTCGCTTGCACTTCGCCGTGTGAGCGCGCCGCCGCTGCACGTTCACCTTGAAGAA
ATTGAACCGCTTAGAGGGGGCGAGCAGCTTGAACAGCTCAGCATGGTATGATGGAACA
TGGGCTCGTACTCATTTTGTTGGTTAGAGAGTTGGCGAGCCAATGATTAACAGGGACTG
CCGGGGGCATTAGTACTCGGACGGGAGAGGTGAAATTCAAGGATCGTCCGAAGACTCC
CTACTGCGAAAGCATTTGCCAAGAATGTTTTCATTAATCAAGAACGAAAGTCAGAGGTG
CGAAGGCGATCAGATACCGCCGTAGTTCTGACCGTAAACGATGCCGACTGGCAATTGGG
CGCACTTCTGTCGAAGTTGCCGCCAGCAGTACTGCCCGGGAAACCAAAGTCATTGGGTT
CCGGGGGGAGTATGGTTGCAAAGCTGAAACTTAAAGGAATTGACGGAAGGGCACCACC
AGGAGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAACCTCACCCGGCCCGAAC
ACTGTTATGACAGACAGGTTGAGAGCTCTTTCTCGATTCAGTGGTTGGTGGTGCATGGC
CGTTCTTAGTTCGTGGAGCGATCTGTCTGGTTAATTCCGATAACGAACGAGACTCTTGCC
TGCTAAATAGACGGCGCCGACTCGGCACGGCGACCGTTTGCTTCTTAGAGGGACAAGCG
GCTTTTAGTCGCGTGAAGCCGAGAGCAATAACAGGTCAGTGATGCCCTCAGATGTTCGG
GGCCGCACGCGCGCTACACTGTTTGCATCAGCGTGTTACTCCCGCGCCGGCTGGCGTGG
GCAACCCGTTGAACCGCAAACGTGCTAGGGATCGGAGATTGCAATTGTTCTCCGTGAAC
GAGGAATTCCTTGTACTGGCGAGTCATCAGCTCGCGGGGAATCTGTCCCTGCCCTTTGTA
CACACCGCCCGTCGCTACTACCGATTGAGTGGTTTAGTGAGGCTCACGGACTGCGCGCG
GCACACGGCGTCTTCGGGCGTCGCGTCGCGAGGCGGAAAGTGAGACGAACTTGATCAC
TTCTAGGAAGTAAAAGTCGTAACAAGGTTTTCGTATGTGAACCTCGGAAAGGATCACAC
CCCTACGAAAACCCTTGTTACCACTTTTACTTCCAGAAGTGATCAGTTGCTTCATTTTGC
CTGCGCAGCGAGCCGAGACACCGGGGGCGGCGGCGCGAGCTACTAACATAATCGGTAT
AGCGAAGGGGGGTGGTGTTCTGTTGTAGAGTGAGAGGAGACCGACATCTGCATGACTC
CGCATCCAGGGAATTCTCGTGCACTATAAAAAATTGTATTCCCTCTCCACTATTGTGGGT
GACGCGCGGACGAGCTCACTACAC

>*Phaeostachys spinifera*

TAGTATCAAGCCGCGTACGACGGCGAGACTGCGGACGGCTCATTAAATCGGTTATGAAT
CCACTGGGGACAGACTGACCCGTGGATAACGTGCGGTAACTCCGGTGCTAATACATGCA

ACAAGGCTCCGACCCCGTCCGCTCGCGGGCGGGGGAAGGGCGCACTTATTAGGCGAAA
ACCAATCGCTCGCTCGCGGGCGTTGGTGGACGACACCCGAGTAATTGCCGCCGATCGCA
CGGCCTCCGAGCCGGCGACACCTACACTGAGTTTCTGATCTATCATGCTGACGACGGTT
GGTGCTATGCCAACCGTGGCGTTAACGGATAACAGAGAATCTGGGTTCGATTCTGGAGA
GGCCGCATGAGAAACGGCGACCACTTCCAAGGAAGGGCAGCAGGCACGCAAATTACCC
ACTTCGGACACCGAGAGGTAGTGACGAACAATACCGATGCGGCGCACTTACGTGTCGCC
GTAATCGGAATGAGTACACTTCAAATCCTTTAACGAGGATCCACTGGAGGGCAAGCCTG
GTGCCAGCAGCCGCGGTAATTCCAGCTCCAGCAGCGTATATTTATATTGCTGCGTTTAA
AAAGCTCGTAGTTGAAGCGTCAGACGCGGGAGGGCGGACGGCCGTACAGGTCGCTCTC
GCCCGACCGAGCCAGGCGCAGAGGGCGCGCGTCGCTTGCACTTCGCCGTGTGAGCGCG
CCGCCGCTGCACGTTCACCTTGAAGAAATTGAACCGCTTAGAGGGGGCGAGCAGCTTGG
ACAGCTCAGCATGGTATGATGGAACATGGGCTCGTACTCATTTTGTTGGTTAGAGAGTC
GGCGAGCCAATGATTAACAGGGACTGCCGGGGGCATTCGTACTCGGACGGGAGAGGTG
AAATTCAAGGATCGTCCGAAGACGCCCTACTGCGAAAGCATTTGCCAAGAATGTTTTCA
TTAATCAAGAACGAAAGTCAGAGGTGCGAAGGCGATCAGATACCGCCGTAGTTCTGAC
CGTAAACGATGCCGACTGGCAATTGGGCGCACTTCTGTCGAAGTTGCCGCCAGCAGTAT
TGCCCGGGAAACCAAAGTCATTGGGTTCCGGGGGGAGTATGGTTGCAAAGCTGAAACTT
AAAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAA
CACGGGAAACCTCACACCGGCCCGAACACTGTTATGACAGACAGGTTGAGAGCTCTTTC
TCGATTCAGTGGTTGGTGGTGCATGGCCGTTCTTAGTTCGTGGAGCGATCTGTCTGGTTA
ATATCCGATAACGAACGAGACTCTTGCCTGCTAAATAGACGGCGCCGACTCGGTACGGC
GTCCGTTTGCTTCTTAGAGGGACAGGCGGCTTTTAGCCGCAGTAAGCCGAGAGCAATAA
CAGGTCAGTGATGCCCTCAGATGTTCGGGGCCGCACGCGCGCTACACTGTTTGCATCAG
CGTGTTACTCCCGCGCCGGCTGGCGTGGGCAACCCGTTGAACCGCAAACGTGCTAGGGA
TCGGAGATTGCAATTGTTCTCCGTGAACGAGGAATTCCTTGTACTGGCGAGTCATCAGC
TCGCGGGGAATCTGTCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGATTGAGTG
GTTTAGTGAGGCTCACGGACTGCGCGCGGCACTCGGCGTCTTCGGGCGTCGAAGCCGCG
TCGCGGAAAGTGATACGCTACCTGATCGACTTACTAGGAAGTTAAAAGTCGTACAGGGT
TGGAAACATAAAAAAGTGANTCT

>*Schizomavella linearis*

ACCAACTGGGTGCACCTGCAGTCTGGTNGATCNNNCCGGNGGNTGAAACTGCGGANNG
CTCATTAAATCAGTTATGTTACNTGATTGTACAATGTTTACATGGATATCTGTGGTAATT
CTAGAGCTAATACATGCGAAAGGTCCCGACCTCTGGAAGGGATGTATTTATCAGCTTTA
AAACCAATGGAGTCCTTGTGTCTCGCATTATTGACGAATCATGATAACTGTTCGAATCG
CACGGCCTCGTGCTGGCGATGTTTCTTTCAAATTTCTGCCCTATCAACTGTCGATGGTAC
GGTAGTGGCCTACCATGGTTTTTACGGGTGACGGAGAATCAGGGTTCGGTTCCGGAGAG
GGAGCCTGAGAAACGGCTACCACATCTAAGGAAGGCAGCAGGCACGGAAATTACCCAA
TCCCAATTCGGAGAGGTAGTGACGAACAATACCGATGCGGCGCACACATGTGCCGCCGT
AATCGGAATGAGTACACTTCAAATCCTTTAACGAGGATCCACTGGAGGGCAAGCCTGGT

GCCAGCAGCCGCGGTAATTCCAGCTCCAGCAGCGTATATTTATATTGCTGCGTTTAAAA
AGCTCGTAGTTGAAGCGTCAGACGCGGGAGGGCGGACGGCCGCATTGGTCAGTCGCGC
CCGACCGAGCCAGGCGCAGAATGGCGCGCGTCGCTTGCACTTCACCGTGTGAGCGCGCC
GCCGCTGCACGTTCACCTTGAAGAAATTGAACCGCTTAGAGGGGGCGAGCAGCTTGGAC
AGCTCAGCATGGTATGATGGAACATGGGCTCGTACTCATTTTGTTGGTTAGAGAGTTGG
CGAGCCAATGATTAACAGGGACTGCCGGGGGCATTCGTACTCGGACGGGAGAGGTGAA
ATTCAAGGATCGTCCGAAGACGCCCTACTGCGAAAGCATTTGCCAAGAATGTTTTCATT
AATCAAGAACGAAAGTCAGAGGTGCGAAGGCGATCAGATACCGCCGTAGTTCTGACCG
TAAACGATGCCGACTGGCAATTGGGCGCACTTCTATAGAAGTTGCCGCCAGCAGTATTG
CCCGGGAAACCAAAGTCATTGGGTTCCGGGGGGAGTATGGTTGCAAAGCTGAAACTTA
AAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAAC
ACGGGGAACCTCACCCGGCCCGAACACTGTTATGACAGACAGGTTGAGAGCTCTTTCTC
GATTCAGTGGTTGGTGGTGCATGGCCGTTCTTAGTTCGTGGAGCGATCTGTCTGGTTAAT
TCCGATAACGAACGAGACTCTTGCCTGCTAAATAGACGGCGTCGACACGAGCGGCGAC
CGTTCGCTTCTTAGAGGGACAAGCGGCTTTTAGCCGCGTGAAGCAGAGAGCAATAACAG
GTCAGTGATGCCCTCAGATGTTCGGGGCCGCACGCGCGCTACACTGTTTGCATCAGCGT
GTTTCTCCCGCGCCGGCTGGCGTGGGCAACCCGTTGAACCGCAAACGTGCTAGGGATCG
GAGATTGCAATTGTTCTCCGTGAACGAGGAATTCCTTGTACTGGCGAGTCATCAGCTCG
CGGGGAATCTGTCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGATTGAGTGGTT
TAGTGAGGCTCACGGACTGCGCGCGGCACTCGGCAGCTAATCGGGCGTCGCGCCGCGAT
GTGGAAAGTGAGACGAACTTGTCACTTNTAGGANGTNAAAAGNCNTNACAAGGTTTTC
GTAGTGAACTGCCGGGAAGGATCATA

>*Scruparia chelata*

GACTCAAATGGCGGGTCAGATAGCCATGCATGTCTAGTTCACACCCTCGTATGGTGAAA
CCGCGAATGGCTCATTAAATCAGTCGAGGTTCCTTAGATGATCCAAATCTACTTGGATA
ACTGTGGTAATTCTAGAGCTAATACATGCCAACCAGCTCCGACCCGCAAGGGAAAGAG
CGCTTTTATTAGTTCAAAACCAGTCGGGCTTCGGTCCGTCCTTTGGTGACTCTGGATAAC
TTTGTGCCGATCGCAGGGCCTTGTGCCGGCGACGCATCTTTCAAATGTCTGCCCTATCAA
ATGACGATGGTACGTGATCTGCCTACCATGTTAGCAACGGGTAGCGGGGAATCAGGGTT
CGATTCCGGAGAGGGAGCATGAGAAACGGCTACCACATCCAAGGAAGGCAGCAGGCGC
GCAACTTACCCACTCCTGGCACGGGGAGGTAGTGACGAAAAATAACAATACGGAACTC
TTTTGAGGCTCCGTAATTGGAATGAGTACACTTTAAACCCTTTAACGAGGATCTATTGGA
GGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATACTAAAGTT
GTTGCGATTAAAAAGCTCGTAGTTGGATCTCAGGCATGGGCGCACGGTCCGCCTCACGG
CGGTCACTGTGTGTATTTTCCCATCCTACGCTTCCCGGTTGTTCAGCCCATGGTGCTCTTC
ATTGAGCGTTTTGGGTGGCCGGAAATTTTACTTTGAAGAAATTAGAGTGTTCAAAGCAG
GCACGTCGCCTGAATAATGGTGCATGGAATAATGGAATAGGACCTCGGTTCTATTTTGC
TGGTTTTCGGAACACGAGGTAATGATTAAGAGGGACAGACGGGGGCATCCGTATTGCG
GTGTTAGAGGTGAAATTCTTGGATCATCGCAAGACGAACAACTGCGAAAGCATTTGCCA

AGAATGTTTTCATTAGTCAAGAACGAAAGTCAGAGGTTCGAAGACGATCAGATACCGTC
GTAGTTCTGACCATAAACGATGCCAACTAGCGATTCGCTGGTGTTGCTTCATCGACTCTG
CGGGCAGCTTCCGGGAAACCAAAGTTTTCGGGTTCCGGGGGAAGTATGGTTGCAAAGCT
GAAACTTAAAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTT
GACTCAACACGGGAAAACTCACCCGGTCCGGACACTGTAAGGATTGACAGATTGATAG
CTCTTTCTTGATTCGGTGGGTGGTGGTGCATGGCCGTTCTTAGTTGGTGGAGCGATTTGT
CTGGTTAATTCCGATAACGAACGAGACTCTAGCCTACTAAATAGTTCGTCGATCCTTTAT
GCGTCGGCGTTAACTTCTTAGAGGGACAAGTGGCTTTTAGTCACACGAGATTGAGCAAT
AACAGGTCTGTGATGCCCTTAGATGTCCGGGGCCGCACGCGCGCTACACTGAAGGAGAC
AGCGTGGCTTCTTCCCTGATTCGAAAGGATTGGGAAACCCGTTGAATTTCCTTCGTGATA
GGGATTGGGGCTTGAAATTCTTCCCCATGAACGAGGAATTCCCAGTAAGCGCGAGTCAT
AAGCTCGCGTTGATTACGTCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGATTG
AACGGTTTAGTGAGGGCCTCGGATTGATCTCGGCCCGCCCTTCACTGGGCGGCGCCGTT
GATCGAGAAGACGCTCGAACTTGATCGTTTAGAGGAAGTAAAAGTCGTAACAAGGTTTC
CGTAGTGAACTGTCGGCGGAAGGGATATAC

>*Tubulipora liliacea*

GAACTTTGTCTTGTCTGCAAAGATTAAGCCATGCATGTCTATGTCTAAACCTTAGTACGG
TGAAACCGCGAATGGCTCATTAGATCGGTTATGGTTCCTTGGATCTTACAAATCCTACTC
GGATAACTGTGGTAATTCTAGAGCTAATACGTGCTACCAAGGCCCGACCTCACGGGACG
GCTGCGTTTATTTTCTCAAAACCAAACCGCTTTCGGGCGGTTCCTTTGGTGATTCTGGAT
AACTTAGAGCTGATCGCACAGGCTAGTCCTGGCGACGTATCTTTCAAATGTCCGCCCTA
TCAACTTTCGATGGTCGGTGCCTTGCCTACCATGGTTGTAACGGGTAACGGGGAATCAG
GGTTCGGTTCCGGAGAGGGAGCATGAGAAACGGCTACCACTTCCAAGGAAGGCAGCAG
GCGCGCAAATTACCCACTCCCAGCTCGGGGAGGTAGTGACGAAAAATAACAATGCGGG
ACTCTTTCGAGGCCCCGTAATTGGAATGAGTACACTTTAAAGCCTTTAGCGAGGATCAA
TTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATTA
AAGTTGTTGCGGTTAAAAAGCTCGTAGTTGGATCCCAGAGAGGCGCACGCGGTTGGCCC
ACGCGGTCTTACTTCGTGCGTGTCTCACCGACCGCTGGTCGCGAGTGCTCTTTACTGAGT
GTTCGTCGACTGCGGCCACGTTTACTTTGAAAAAATTAGAGTGTTCAAAGCAGGCGTCT
TTCGCCCGCATACCCCAGCATGGAATAATGGAATAGGAGGCTGGTTCTATTTTGTTGGTT
TTCGGAATCTAGCCTAATGGTTAAGAGGGACGGCCGGGGGCATTCGTATTGTGGCGTTA
GAGGTGAAATTCTTAGATCGCCGCAAGACGAACAAGTGCGAAAGCATTTGCCAAGAAT
GTTTTCATTAATCAAGAACGAAAGTCAGAGGCGCGAAGGCGATCAGATACCGCTGTAGT
TCTGACCATAAACGATGCCAACTAGCGATCGGTCGGAGTTGCTCTAATGACCCGACCGG
CAGCTGCCGGGAAACCAAGCAAGTGTTTGGGTTCCGGGGGGAGTATGGTCGCAAGGCC
GAAACTTAAAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTT
GACTCAACACGGGAAAACTCACCCGGCCCGGACACTGTGAGGATTGACAGATTGATAG
CTCTTTCTTGATTCGGTGGGTGGTGGTGCATGGCCGTTCTTAGTTGGTGGAGCGATTTGT
CTGGTTAATTCCGATAACGAACGAGACTCTAGCCTGCTAAATAGTAGGCGGATCCCCTG

TGGTCCGTCGTGAACTTCTTAGAGGGACAAGCGGCGTTCAGCCGCACGAGATTGAGCAA
TAACAGGTCTGTGATGCCCTTAGATGTTCGGGGCCGCACGCGCGCTACACTGAAGGCAT
CAGCGTGTCCTTCCTGCTCCGACAGGAGCGGGTAACCCGTTGAACCGCTTTCGTGCTAG
GGATCGGGGCTTGCAATTCTTCCCCGTGAACGAGGAATTCCCAGTAAGCGCAAGTCATA
AGCTTGCGTTGATTACGTCCCTGCCCTTTGTACACACCGCCCGTCGCTACTACCGATTGA
ATGGTTTAGTGAGGCCCTCGGACGTTCTGTCCGCCGCCTGTCAAGGGCGGTAGACGAGA
GTGGAAGACGGTCAAACTTGATCATTTAGAGGAAGTAAAAGTC

>*Umbonula littoralis*

ATGCATAGCCGGTCCAGATAAGACATGCATGTCTAAGTACAAGCCGTGCACACGGCGA
GACTGCGGACGGCTCATTAAATCGGTTATGACTCCACTGGGGACAGACTCACCCGTGGA
TAACGTGCGGTAACTCCGGTGCTAATACATGCAACAAGGCTCCGACCCGGCCGTCAAAC
GCTGGGGAAGGGCGCACTTATTAGGCGAAAACCAATCGCTCGCTCGCGGGCGTTGGTG
GACGACACCCGAGTAATTGCCGCCGATCGCACGGCCTCGCGCCGGCGACACCTTCACTG
AGTTTCTGATCTATCATGCTGACGACGGTTGGTGCTATGCCAACCGTGGCGTTAACGGA
TAACAGAGAATCTGGGTTCGATTCTGGAGAGGCCGCATGAGAAACGGCGACCACTTCTA
AGGAAGGCAGCAGGCACGCAAATTACCCACTTCGGACACCGAGAGGTAGTGACGAACA
ATACCGATGCGGCGCACACATGTGCCGCCGTAATCGGAATGAGTACACTTCAAATCCTT
TAACGAGGATCCACTGGAGGGCAAGCCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCC
AGCAGCGTATATTTATATTGCTGCGTTTAAAAAGCTCGTAGTTGAAGCGTCAGACGCGG
GAGGGCGGACGGCCGCAATGGTCGCTCTCGCCCGACCGAGCCAGGCGCAGAGGGCGCG
CGTCGCTTGCACTTCGCCGTGTGAGCGCGTCGCCGCTGCACGTTCACCTTGAAGAAATT
GAACCGCTTAGAGGGGGCGAGCAGCTTGAACAGCTCAGCATGGTATGATGGAACATGG
GCTCGTACTCATTTTGTTGGTTAGAGAGTCGGCGAGCCAATGATTAACAGGGACTGCCG
GGGGCATTCGTACTCGGACGGGAGAGGTGAAATTCAAGGATCGTCCGAAGACGCCCTA
CTGCGAAAGCATTTGCCAAGAATGTTTTCATTAATCAAGAACGAAAGTCAGAGGTGCGA
AGGCGATCAGATACCGCCGTAGTTCTGACCGTAAACGATGCCGACTGGCAATTGGGCGC
ACTTCTATAGAAGTTGCCGCCAGCAGTATTGCCCGGGAAACCAAAGTCATTGGGTTCCG
GGGGGAGTATGGTTGCAAAGCTGAAACTTAAAGGAATTGACGGAAGGGCACCACCAGG
AGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAACCTCACCCGGCCCGAACACTG
TTATGACAGACAGGTTGAGAGCTCTTTCTCGATTCAGTGGTTGGTGGTGCATGGCCGTTC
TTAGTTCGTGGAGCGATCTGTCTGGTTAATTCCGATAACGAACGAGACTCTTGCCTGCTA
AATAGACGGCGTCGACCTCGTCGGCGACCGTTCGCTTCTTAGAGGGACAAGCGGCTTTT
AGCCGCGTGAAGCCGAGAGCAATAACAGGTCAGTGATGCCCTCAGATGTTCGGGGCCG
CACGCGCGCTACACTGTTTGCATCAGCGTGTTTCTCCCGCGCCGGCCGGCGTGGGCAAC
CCGTTGAACCGCAAACGTGCTAGGGATCGGAGATTGCAATTGTTCTCCGTGAACGAGGA
ATTCCTTGTACTGGCGAGTCATCAGCTCGCGGGGAATCTGTCCCTGCCCTTTGTACACAC
CGCCCGTCGCTACTACCGATTGAGTGGTTTAGTGAGGCTCACGGACTGCGCACGGCTCA
CGGCGCCTCCGGGCGTCGCGTCGCGATGCGGAAAGTGAGACGAACTGATCACTCTAGA
AGTAAAAGTCGACCATGCCAAGTAGTGACCTGCGGAGGATCAGCCCTTACGAAAACTT

GTTACGAACTTTTACTTTCTTAGAAAGTGATCAAGTTCGTCTTCACTTTTTCGCATCGCG
GACAGGAAA

>*Walkeria uva*

CCACCCATCATTTCCTAGATTAGCCATGCATGTCTAAGTACAAGCCGCGCACGCGGCGA
GACTGCGGACGGCTCATTAAATCGGTTACGACTCCGCTGGGGCCAGACTCCTACGTGGA
TAACGTGCGGTAACTCCGGTGCTAATACATGCAACCAGGCTCCGACCGCGTCTTCGGGC
GCGGGAAGGGCGCACTTATTAGGCGAAAACCAATCGCCGGCCTCCGGGTCGGCGTTGG
CGGACGACACCCGAGTAATTGCCGCCGATCGCACGGCCTCGAGCCGGCGACGCTTCCAT
CGAGTTTCTGATCTATCATGCTGACGACGGTTGGCGCTATGCCAACCGTGGCGTTGACG
GATAACAGAGAATCTGGGTTCGATTCTGGAGAGGCCGCATGAGAAACGGCGACCACTT
CCAAGGAAGGCAGCAGGCACGCAAATTACCCACTTCGGACACCGAGAGGTAGTGACGA
ACAATACCGATGCGGCGCGCTCACGCGTCGCCGTAATCGGAATGAGTACACTTCAAATC
CTTTAACGAGGATCCACTGGAGGGCAAGCCTGGTGCCAGCAGCCGCGGTAATTCCAGCT
CCAGCAGCGTATATTTATATTGCTGCGTTTAAAAAGCTCGTAGTTGAAGCGTCAGACGC
GGGAGGGCGGACGGCCGCACTGGTCGCTCTCGCCCGACCGAGCCAGGCGCAGAGGGCG
TCCGTCGCTTGCACTTCGCCGTGTGAGCGCGGCGCCGCTGCACGTTCACCTTGAGGAAA
TTGAACCGCTCAGAGGGGGCGAGCAGCTTGCACAGCTCAGCATGGTATGATGGAACAC
GGGCTCGTACTCGTTTTGTTGGTTTTAGAGTCGGCGAGCCAATGATTAAGAGGGACTGC
CGGGGGCATTCGTACTCGGACGGGAGAGGTGAAATTCAAGGATCGTCCGAAGACGCCC
TACTGCGAAAGCATTTGCCAAGAATGTTTTCATTAATCAAGAACGAAAGTCAGAGGTGC
GAAGGCGATCAGATACCGCCGTAGTTCTGACCGTAAACGATGCCGACTGGCAATTGGGC
GCACTTCTGTAGAAGTTGCTGCCAGCAGCGCGTCCCGGGAAACCAAAGTCATTGGGTTC
CGGGGGGAGTATGGTTGCAAAGCTGAAACTTAAAGGAATTGACGGAAGGGCACCACCA
GGAGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAACCTCACCCGGCCCGAACAC
TGTTATGACAGACAGGTTGAGAGCTCTTTCTCGATTCAGTGGTTGGTGGTGCATGGCCGT
TCTTAGTTCGTGGAGCGATCTGTCTGGTTAATTCCGATAACGAACGAGACTCTCGCCTGC
TAAATAGACGGCGCCGACGTACGCGCGGCGACCGCGCAGCTTCTTAGAGGGACAAGCG
GCGTTTAGCCGCGTGAAGCCGAGAGCAATAACAGGTCAGTGATGCCCTCAGATGTTCGG
GGCCGCACGCGCGCTACACTGTTTGCATCAGCGTGTCTCTCCCTCGCCGGCGGGCGCGG
GCAACCCGTTGAAACGCAAACGTGCTAGGGATCGGAGATTGCAATTGTTCTCCGTCAAC
GAGGAATTCCTTGTACTTGCGAGTCATCAGCTCGCGGGGAATCTGTCCCTGCCCTTTGTA
CACACCGCCCGTCGCTACTACCGATTGAGTGGTTTAGTGAGGCTCGCGGACGGCGCGCG
GCAACCGTCGGGTTCGCCCGTCGCGTCGCGCGCCGGAAGCGAGACGAACTTGATCACTC
TAGGAAGTAAAGCGACATGCAGCATCGATAAGGGTACTGGGGGTAGGGACCCCCCTCC
GAAAACTTGGTTTCGAATTTTTTCTTTCTCAGGAGG

>*Alcyonidium gelatinosum*

TGGGGGCGAGTCGCATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTTACCTGGT
TGATCCTGCCAGTAGTCACATGNTCAACCCAAATGCTAAGCCATGCATGCGTAAGTGTT

GGTCGTTCACGCGGCGTACACGGCAAAGAGCTCATTAAATCGGTTAGAAATCTACTGGC
TCCGAGAGCCTGTTAGTGGATAAGATCGGAAACTCTGGTATTAATACATGCAACGATGC
GAGACGGCCGACCGCCGCCTCGCGCACTTACTCTTCGTAGCCGGCCGATGCGTCAGCGA
TGGCGCGTCGCGATGGAGGGGATGCCCGAGTAAACTGGCCGATCGCATAGCATAGCCT
GCGCGGCGACGTTTCGATTGTCTTCCCGCTCCATCAGCCCGAGATGGCTGTGACGTCGA
CGGCCCTGGCGTTCACGGATAACAGGGAATCGGGGTTCGGTTCTGGAGAGGACGCCTG
AGAAACAGCGACCATTTCCAAGGAGGGCAGCAGGCGCGCAACTTACCCAATTCTGCGG
CAGAGAGGTAGTGACGAGTAATAACGGTACGCGGGATGCCGCGGCCTCCGGGTCGCGG
CTACCGCAATACAGTCATGAGAGCGGTCGAAACCACCCAGCGAGGAGCTACTGGAGGG
CAAGCCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAGGAGCGTATATTAATGCTGTT
GCGTCTGAAACGTTCCTAGCCGATGCGGGTGCCGCGTCCGAATCGTGTGGGACCGCTTC
CGACGGTCGCGTCGACCTTGTTTGCACTGGTCGGCTGCGATCCAAGTCGGCTCGGGGCG
GGAAGACCTAGCGAAGACGGCGCGGTTCGTCGCGGCGGTCTCGACCGTCCGCTCGTCCG
AGCGCGCGCGGCGAGGCCGCTGCCGTTCACCCTGAAGAAAGCGGATTGCCCAACGAAG
GCATACGGCCTGCACCTTAAAGCAGGGCACGATGCAATACGCGAGACTGATCGGTGCC
GGCTCGGAGCCAGTACGCCATGAGAGATGGGGATGGCAGGCGAAAGCTGCTGTCGGAG
CGGCTAGTACGACGGTCCGAGTGGTGAAATGCAGTGACTGTCGTTTGACTGCCCGAGGC
GAAGGCGAGCCCCGCGTACGCCTCCACCGGCCAAGAATGAAGGTTGGGGTAGTCAAGG
CGATTAGATACCGCCGTAGTCCCGACGGTAAACTATGCCAACCGGGGATCCGGTCTCGC
CAGGCTTGCCGTGAAAAATCGTGTCGGCTCGAGATTGTGGGAAACCCGCGGGAAACCG
GAGTGATCGGGTTACGGGGGGAGTATGCTCGCAAGGGTGAAACTTAAAGGAATTGACG
GAAGAGCACCACAAGGCGTGGAGCTTGCGGCTTAATTTGACTCAACACGGGAAAACTT
ACCCGGTCAGGACACTGTAATGACAGACAGGTTGAGAGATCTTTCTCGATTGCAAACGG
CAGTTGGCCGACTCTCGACGCGACCGAGAGTTTTACCCGTGAGGTGTTTGAACCGTGAC
CGCGAGGTTATGTAAATTCAGCTTTCACAGACCGCCCCGGCTGTTCGCTCCGGGGCGGT
TTCGATGGCCAAAGACACGCGAAGAGGCGAGCAAAAGCAGGTCAGTGATGCCCCTAGA
TGTTCGGGGCCGCACGCGCGCCACACTGCCTGCGGCAGCGTGTCATCGTTAGACCGGCC
GTCTTCGTACGGCTGGTGAAGCAGAACTGACGACCAACATCAAAAGCTCGCGTGGTCGG
GCTCGGGGGTTGCAATTGTCCCCCGCAAACGCGGAATGCCTTGTATGGGCGCCTCATCA
GGGCGCGCCGAATACGTCCCTGTTCTTTGTACACACCGCCCGTCGCTACTACCGATCGA
CGACCCAGTGAGGCACGCGGACCGTCTCGGAGGTGTAGCAAATCATTTTAACCAAGGCT
AAAGTAAAGTCGTTAACCAGGTTGTAAACGACTTTTTATTTCTTTAGAACGCAGGTCCGT
AGCGCTTTCTCCAAGAGTCCGTACTCACTGTATCTCTGACTGAATATACGCGTGCGCTGC
TGTACAAGAAAGACGAGACTATCCGGACCGTAATAACGGCATCAAGATGACTTTTGCA
GTGCAGAGTCGACAGACAGCCTTGAGTACTGTACATTGGCTTAAACAGAGTCC

>*Alcyonidium polyoum*

AGGGCGAAGTCGCATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTTACCTGGTT
GATCCTGCCAGTAGTCACATGCTCAACCCAAAGGCTAAGCCATGCATGCGTAAGTGTTG
GTCGTTCACGCGGCGTACACGGCAAAGAGCTCATTAAATCGGTTAGAAATCTACTGGCT

CCGAGAGCCTGTTAGTGGATAAGATCGGAAACTCTGGTATTAATACATGCAACGATGCG
AGACGGCCGACCGCCGCCTCGCGCACTTACTCTTCGTAGCCGGCCGATGCGTCAGCGAT
GGCGCGTCGCGATGGAGGGGATGCCCGAGTAAACTGGCCGATCGCATAGCATAGCCTG
CGCGGCGACGTTTCGATTGTCTTCCCGCTCCATCAGCCCGAGATGGCTGTGACGTCGAC
GGCCCTGGCGTTCACGGATAACAGGGAATCGGGGTTCGGTTCTGGAGAGGACGCCTGA
GAAACAGCGACCATTTCCAAGGAGGGCAGCAGGCGCGCAACTTACCCAATTCTGCGGC
AGAGAGGTAGTGACGAGTAATAACGGTACGCGGGATGCCGCGGCCTCCGGGTCGCGGC
TACCGCAATACAGTCATGAGAGCGGTCGAAACCACCCAGCGAGGAGCTACTGGAGGGC
AAGCCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAGGAGCGTATATTAATGCTGTTG
CGTCTGAAACGTTCCTAGCCGATGCGGGTGCCGCGTCCGAATCGTGTGGGACCGCTTCC
GACGGTCGCGTCGACCTTGTTTGCACTGGTCGGCTGCGATCCAAGTCGGCTCGGGGCGG
GAAGACCTAGCGAAGACGGCGCGGTTCGTCGCGGCGGTCTCGACCGTCCGCTCGTCCGA
GCGCGCGCGGCGAGGCCGCTGCCGTTCACCCTGAAGAAAGCGGATTGCCCAACGAAGG
CATACGGCCTGCACCTTAAAGCAGGGCACGATGCAATACGCGAGACTGATCGGTGCCG
GCTCGGAGCCAGTACGCCATGAGAGATGGGGATGGCAGGCGAAAGCTGCTGTCGGAGC
GGCTAGTACGACGGTCCGAGTGGTGAAATGCAGTGACTGTCGTTTGACTGCCCGAGGCG
AAGGCGAGCCCCGCGTACGCCTCCACCGGCCAAGAATGAAGGTTGGGGTAGTCAAGGC
GATTAGATACCGCCGTAGTCCCGACGGTAAACTATGCCAACCGGGGATCCGGTCTCGCC
AGGCTTGCCGTGAAAAATCGTGTCGGCTCGAGATTGTGGGAAACCCGCGGGAAACCGG
AGTGATCGGGTTACGGGGGGAGTATGCTCGCAAGGGTGAAACTTAAAGGAATTGACGG
AAGAGCACCACAAGGCGTGGAGCTTGCGGCTTAATTTGACTCAACACGGGAAAACTTA
CCCGGTCAGGACACTGTAATGACAGACAGTTTGAGAGACCTTTCTCGATTCGGTGATTG
GTGGTGCATGGCCGCTCATAGTTCGTGGAATGATTTGTCAGGTTAAATCCGGTAACGGG
CGGAACTCGCACCTGCTAAAAAGACGGCCGAACGCCAATGTACGGCGCCGTTCGGCTG
CTCCCGCGACCGTAAGAGCTTCTTAGAGGGACCAATGTCACGGCTACCGATTTCTCGCT
TGCAAACGGCAGTTGGCCGACTCTCGACGCGACCGAGAGTTTTACCCGTGAGGTGTTTG
AACCGTGACCGCGAGGTTATGTAAATTCAGCTTTCACAGACCGCCCCGGCTGTTCGCTC
CGGGGCGGTTTCGATGGCCAAAGACACGCGAAGAGGCGAGCAAAAGCAGGTCAGTGAT
GCCCCTAGATGTTCGGGGCCGCACGCGCGCCACACTGCCTGCGGCAGCGTGTCATCGTT
AGACCGGCCGTCTTCGTACGGCTGGTGAAGCAGAACTGACGACCAACATCAAAAGCTC
GCGTGGTCGGGCTCGGGGGTTGCAATTGTCCCCCGCAAACGCGGAATGCCTTGTATGGG
CGCCTCATCAGGGCGCGCCGAATACGTCCCTGTTCTTTGTACACACCGCCCGTCGCTACT
ACCGATCGACGACCCAGTGAGGCACGCGGACCTTCTCGGAGGAAAGCGCTCCGAACTT
GGCTGTTCTAATAAGG

>*Alcyonidium hirsutum*

TGGGGCGAGTCGCATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTTACCTGGTT
GATCCTGCCAGTAGTCACATGCTCAACCCAAAGGCTAAGCCATGCATGCGTAAGTGTTG
GTCGTTCACGCGGCGTACACGGCAAAGAGCTCATTAAATCGGTTAGAAATCTACTGGCT
CCGAGAGCCTGTTAGTGGATAAGATCGGAAACTCTGGTATTAATACATGCAACGATGCG

AGACGGCCGACCGCCGCCTCGCGCACTTACTCTTCGTAGCCGGCCGATGCGTCAGCGAT
GGCGCATCGCGATGGAGGGGATGCCCGAGTAAACTGGCCGATCGCATAGCATAGCCTG
CGCGGCGACGTTTCGATTGTCTTCCCGCTCCATCAGCCCGAGATGGCTGTGACGTCGAC
GGCCCTGGCGTTCACGGATAACAGGGAATCGGGGTTCGGTTCTGGAGAGGACGCCTGA
GAAACAGCGACCATTTCCAAGGAGGGCAGCAGGCGCGCAACTTACCCAATTCTGCGGC
AGAGAGGTAGTGACGAGTAATAACGGTACGCGGGATGCCGCGGCCTCCGGGTCGCGGC
TACCGCAATACAGTCATGAGAGCGGTCGAAACCACCCAGCGAGGAGCTACTGGAGGGC
AAGCCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAGGAGCGTATATTAATGCTGTTG
CGTCTGAAACGTTCCTAGCCGATGCGGGTGCCGCGTCCGAATCGTGTGGGACCGCTTCC
GACGGTCGCGTCGACCTTGTTTGCACTGGTCGGCTGCGATCCAAGTCGGCTCGGGGCGG
GAAGACCTAGCGAAGACGGCGCGGTTCGTCGCGGCGGTCTCGACCGTCCGCTCGTCCGA
GCGCGCGCGGCGAGGCCGCTGCCGTTCACCCTGAAGAAAGCGGATTGCCCAACGAAGG
CATACGGCCTGCACCTTAAAGCAGGGCACGATGCAATACGCGAGACTGATCGGTGCCG
GCTCGGAGCCAGTACGCCATGAGAGATGGGGATGGCAGGCGAAAGCTGCTGTCGGAGC
GGCTAGTACGACGGTCCGAGTGGTGAAATGCAGTGACTGTCGTTTGACTGCCCGAGGCG
AAGGCGAGCCCCGCGTACGCCTCCACCGGCCAAGAATGAAGGTTGGGGTAGTCAAGGC
GATTAGATACCGCCGTAGTCCCGACGGTAAACTATGCCAACCGGGGATCCGGTCTCGCC
AGGCTTGCCGTGAAAAATCGTGTCGGCTCGAGATTGTGGGAAACCCGCGGGAAACCGG
AGTGATCGGGTTACGGGGGGAGTATGCTCGCAAGGGTGAAACTTAAAGGAATTGACGG
AAGAGCACCACAAGGCGTGGAGCTTGCGGCTTAATTTGACTCAACACGGGAAAACTTA
CCCGGTCCGGACACTGTAATGACAGACAGTTTGAGAGACCTTTCTCGATTCGGTGGTTG
GTGGTGCATGGCCGTTCATAGTTCGTGGAATGATCTGTCAGGTTAAATCCGGTAACGGG
CGGAACTCGACCCTGCTAAAAAGACGGCCGAACGCCCGATCTCGGGTGGGAGGCTGCT
CGCGTGACCGTAAGAGCTTCTTAGAGGGACCAATGTCACGGCCACTGATTTCCCGCTTG
CGAGGTAGACGACCCGGCTCTCAGCGCTCGTCGGGAGGATTTGCATGATGCAATTTCGC
TACTCGATCGAATCACAGACGGCTCCGCGATTCGTCGCGGAATTCGCCGTTTCGATGGC
TATAGACAAGTGAAGATCCGAGCAAAAGCAGGTCAGTGATGCCCCTAGATGTTCGGGG
CCGCACGCGCGCTACACTGCCCGCCGCAGAAAGAGATCGTCAGCGCCCGGAGGCGTTT
ACGCGTCTTCGGACCCGGAGAACTGACGAATATCCGCAATGGTGATTGGCGTGGTTGGG
ATCGGGGTTTGCAATTGTCTCCCGTGAACGCGGAATGCCTTGTATGGGCGCGTCATCAG
CGCGCGCCGAATACGTCCCTGTTCTTTGTACACACCGCCCGTCGCTACTTCCGATCGATG
GCCCAGTGAGGACCGCGGACTCCGGAAAACGGCTCGAACCTTTCTATTGTCTGGTAAGC
GTACAAAGAGTGTTAACGGATTTATTCCGAAGGAGACAGTCCAGCCGTCTCCGGATCGG
GTCTCCCTGGGCATCGATCGGAAGTATGACGGCGGTGTGTTACAGAA

APPENDIX B

The reference alignment sequences stored in the ERRD database were in the DCSE file format (DeRijk and Dewachter 1993). This file format (see Figure 70) was designed for the DCSE (see main text) software package, which was used during the time of the creation of the database. The following symbols are used to indicate secondary structure elements in this format (DeRijk and Dewachter 1993):

[ ] – symbolises beginning and end of one strand of a helix.
^ – symbolises a new helix starting immediately after the previous one (would be identical to using][ ).
{ } – symbolises beginning and end of an internal loop or bulge loop.
( ) – symbolises enclosed base forming part of a non–canonical pair.

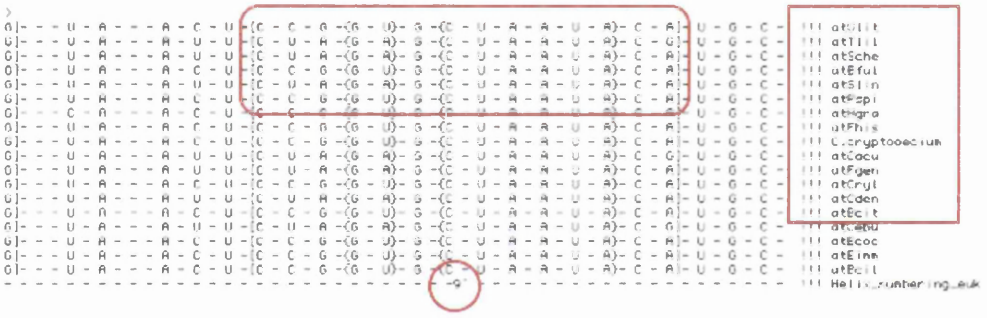


Figure 70 An example of DCSE file format. Red rectangle marks list of sequence names (corresponding to the species names); Red oval shows a fragment of Helix 9′-stem with DCSE recognised symbols (see text); Red circle shows helix numbering mask placer.

As mentioned, the package was not usable due to its age and compatibility issues on top of which this software did not appear to allow any data export to be made for the use with contemporary alignment programs. Another version for this software—source code in C language for Solaris Sun UNIX operating systems—was obtained through an anonymous ftp server in China: (Peking University ftp://ftp.cbi.pku.edu.cn/pub/software/unix/dcse/). This software also failed to compile (under FreeBSD recent update and Fedora Linux Core 6) and after some consideration it was decided not to use it. Although several software packages claim to assist or aid in secondary structure alignment none (to the knowledge of the author) allowed any automation of the alignment process apart from that provided in

the ARB project (Ludwig *et al.* 2004), which unfortunately is mainly aimed at 16S rRNA genes.

APPENDIX C

The Swansea University supercomputer cluster "Blue C" consists of 16 IBM eServer LPARs (logical partitions) nodes. Each server (or a node) contains 16 Power5 64-bit RISC 1.5 GHz clock rate processors. Each processor has level 1, 2 and 3 caches, with a total memory available to each node – 64 Gbytes. The theoretical peak performance of the cluster is 6.8 Gflop/s[45]. As mentioned in chapter 5 MrBayes was compiled for the use as an MPI version. This required some modifications to the Makefile script used for compilation on an AXI Blue-C cluster. Below is shown a user modifiable part of the Makefile as it was used here, with modifications specific to the Blue C system marked in bold. Comments included in an original Makefile (which is available as part of the source code package), and non-user changeable parts are omitted to preserve space, but could easily be found on the MrBayes website.

```
# CONFIG
ARCHITECTURE = unix
MPI = yes
DEBUG = no
#OPTFLAGS = -O2 -march=pentium4 -mfpmath=sse -fomit-frame-
pointers
OPTFLAGS = -O3 -qstrict -qtune=pwr5 -qarch=pwr5 -
DUNIX_VERSION -DMPI_ENABLED -qMAXMEM=-1
CC = mpcc
ifeq ($(strip $(ARCHITECTURE)), unix)
   USEREADLINE = no
else
   USEREADLINE = no
endif
# End of user configuration
ifeq ($(strip $(ARCHITECTURE)),mac)
     CFLAGS += -DMAC_VERSION
else
ifeq ($(strip $(ARCHITECTURE)), windows)
```

[45] All hardware parameters and performance values are taken from the Blue C user guide distributed internally to the users of the cluster by IBM.

```
     CFLAGS += -DWIN_VERSION
else
ifeq ($(strip $(ARCHITECTURE)), unix)
     CFLAGS += -DUNIX_VERSION
else
     ARCHITECTURE = none
endif
endif
endif
ifeq  ($(strip $(USEREADLINE)),yes)
    CFLAGS += -DUSE_READLINE
     LIBS += -lncurses -lreadline
endif
ifeq  ($(strip $(MPI)),yes)
     CFLAGS += -DMPI_ENABLED
     CC = mpcc
endif
ifeq ($(strip $(DEBUG)), yes)
     CFLAGS += -ggdb
else
     CFLAGS += $(OPTFLAGS)
endif
#CFLAGS   +=   -Wall
LIBS     +=   -lm
LDFLAGS  =   $(CFLAGS)
LDLIBS   =   $(LIBS)
OBJECTS     =     bayes.o  command.o  mbmath.o  mcmc.o  model.o
plot.o sump.o sumt.o
PROGS     =   mb
ifeq ($(ARCHITECTURE), none)
missing:
     @echo
     @echo Please set compilation target in this Makefile.
     @echo set ARCHITECTURE to one of \"mac\", \"windows\" or
\"unix\"
     @echo set MPI to yes if you want to use the parallel
version
```

```
	@echo set DEBUG to generate a debug version of MrBayes
	@echo
endif
all:	$(PROGS)
clean:
	rm -f *.o
showdep:
	@$(CC) -MM bayes.c command.c mbmath.c mcmc.c model.c
plot.c sump.c sumt.c
mb: mb.o bayes.o command.o mbmath.o mcmc.o model.o plot.o
sump.o sumt.o
# dependencies are generated by make showdep
bayes.o: bayes.c mb.h globals.h bayes.h command.h mcmc.o
command.o: command.c mb.h globals.h command.h bayes.h model.h
mcmc.h \
   plot.h sump.h sumt.h
mbmath.o: mbmath.c mb.h globals.h mbmath.h bayes.h
mcmc.o: mcmc.c mb.h globals.h bayes.h mcmc.h model.h
command.h mbmath.h \
   sump.h sumt.h plot.h
model.o: model.c mb.h globals.h bayes.h model.h command.h
plot.o: plot.c mb.h globals.h command.h bayes.h plot.h sump.h
sump.o: sump.c mb.h globals.h command.h bayes.h sump.h mcmc.h
sumt.o: sumt.c mb.h globals.h command.h bayes.h mbmath.h
sumt.h mcmc.h
```

Once compiled the software was used in a standard non-interactive way by means of a batch file submission. Initiations of the calculations were done using a Load Leveller file (see example below), which would control the way the cluster's operating system, interacts with MrBayes program.

```
#!/bin/ksh
#
# @ error   = Error
# @ output  = Output
# @ output  = parap.$(jobid).out
# @ error  = parap.$(jobid).err
```

```
# @ notification  = always
# @ notify_user =
# @ wall_clock_limit=96:00:00
# @ job_type = parallel
# @ node = 1
# @ class = par1_96
# @ tasks_per_node = 16
# @ network.mpi = csss,not_shared,US,HIGH
# @ node_usage = not_shared
# @ bulkxfer = yes
# @ queue
export MP_SHARED_MEMORY=yes
set -x
./mb mrbatchfile.nex > output_results.out
```

The details of how to create and modify the above file can be found in the IBM manuals, but of importance here are the lines marked in bold. The first four lines marked state that the job is to be run for 96 hours on 16 CPUs (referring to the total number of chains of MrBayes). This was found to be optimal to load all 16 chains on one node thus not requiring spreading individual chain calculation over the different nodes of the cluster. If more than 16 chains were to be used then additional resources of the second node (i.e. more than 16 CPUs) would have to be called up – these would put a considerable demand on the time required for the calculations to be done and thus was rejected. In total 16 chains (4 chains for 4 parallel runs were used). The final line of the executable script (marked in bold as well) specifies the input file (i.e. MrBayes batch file) and the output file to which screen data was redirected—these are standard UNIX pipes.

MrBayes batch file consisted of the main data set in NEXUS format and an additional MrBayes command executable block that is required for a non-interactive run in the UNIX environment. This block is given below, and it follows the guidelines given in the MrBayes manual.

Batch file ending for the GTR+I+Γ model (no data partitioning) is shown below. Comments are given in square brackets to clarify lines.

```
begin mrbayes;
```

```
[to ensure that !MrBayes does not stop during an analysis to
wait for confirmation from the user]
set autoclose=yes  nowarn = yes  ;
[set the evolutionary model to the GTR model with gamma-
distributed rate variation alfa across sites, and invariable
sites i]
lset  nst=6  rates=invgamma;
[specify number of individual runs, chains and generations;
also tell MrBayes how many samples to record in the out put
file]
mcmcp  nruns=4  nchains=4  ngen =1000000     samplefreq = 100
printfreq = 100;

[begin to run MrBayes]
mcmc;
[summarize the parameter values]
sump burnin = 2500;
[summarize the trees]
sumt burnin = 2500;
[quit automatically when the analysis is done]
quit;
end;
```

Batch file ending for the RNA16BHKY+I+Γ (stem) and GTR+I+Γ (loop) models is given below with data partitioned into two partitions.

```
begin mrbayes;
set autoclose=yes  nowarn = yes ;
[specify two partitions - one for loops and one for stems]
partition loopstem = 2:LOOP,STEM;
set partition = loopstem;
[specify individual models for partitions]
lset applyto = (1) nst = 6 rates = invgamma ;
lset  applyto  =  (2)  nst  =  2  nucmodel  =  doublet  rates  =
invgamma;
[unlink all parameters - therefore allow parameters to be
estimated individually for loop and stem partitions]
```

```
unlink shape = (all);
unlink revmat = (all);
unlink statefreq = (all);
unlink pinvar = (all);
mcmcp nruns=4 nchains=4 ngen = 1000000 samplefreq = 100
printfreq = 100 savebrlens = yes;
mcmc;
sump burnin = 2500;
sumt burnin = 2500;
quit;
end;
```

The above two batch files could be used for any similar runs on MrBayes. In addition, complete files including dataset can be obtained from the author upon request.

REFERENCES

(2004) *Admiralty tide tables : European waters*, Great Britain. Hydrographic Department.

Abd-Elsalam, K. A. (2003) Bioinformatic tools and guideline for PCR primer design. *African Journal of Biotechnology,* 2(5): pp. 91-95.

Abouheif, E., Zardoya, R. & Meyer, A. (1998) Limitations of metazoan 18S rRNA sequence data: Implications for reconstructing a phylogeny of the animal kingdom and inferring the reality of the Cambrian explosion. *Journal of Molecular Evolution,* 47(4): pp. 394-405.

Adi, S. S. & Ferreira, C. E. (2001) DNA Fragments Assembly Programs: a comparative study In Brazilian Summer School on Combinatorics and Algorithms (GRACO). *Electronic Notes in Discrete Mathematics,* 7.

Adoutte, A., Balavoine, G., Lartillot, N. & de Rosa, R. (1999) Animal evolution - the end of the intermediate taxa? *Trends in Genetics,* 15(3): pp. 104-108.

Adoutte, A., Balavoine, G., Lartillot, N., Lespinet, O., Prud'homme, B. & de Rosa, R. (2000) The new animal phylogeny: Reliability and implications. *Proceedings of the National Academy of Sciences of the United States of America,* 97(9): pp. 4453-4456.

Ahola, V., Aittokallio, T., Vihinen, M. & Uusipaikka, E. (2006) A statistical score for assessing the quality of multiple sequence alignments. *Bmc Bioinformatics,* 7.

Akaike, H. (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6): pp. 716-723.

Alberts, B. (2002) *Molecular biology of the cell,* New York, Garland Science.

Alfaro, M. E. & Holder, M. T. (2006) The posterior and the prior in Bayesian phylogenetics. *Annual Review of Ecology Evolution and Systematics,* 37: pp. 19-42.

Alvarez, L. W., Alvarez, W., Asaro, F. & Michel, H. V. (1980) Extraterrestrial Cause for the Cretaceous-Tertiary Extinction - Experimental Results and Theoretical Interpretation. *Science,* 208(4448): pp. 1095-1108.

Alvesgomes, J. A., Orti, G., Haygood, M., Heiligenberg, W. & Meyer, A. (1995) Phylogenetic Analysis of the South-American Electric Fishes (Order Gymnotiformes) and the Evolution of Their Electrogenic System - a Synthesis Based on Morphology, Electrophysiology, and Mitochondrial Sequence Data. *Molecular Biology and Evolution,* 12(2): pp. 298-318.

Amos, W. B. & White, J. G. (2003) How the Confocal Laser Scanning Microscope entered Biological Research. *Biology of the Cell,* 95(6): pp. 335-342.

Amos, W. B., White, J. G. & Fordham, M. (1987) Use of Confocal Imaging in the Study of Biological Structures. *Applied Optics,* 26(16): pp. 3239-3243.

Anderson, C. L., Canning, E. U. & Okamura, B. (1999) 18S rDNA sequences indicate that PKX organism parasitizes bryozoa. *Bulletin of the European Association of Fish Pathologists,* 19(3): pp. 94-97.

Anderson, F. E., Cordoba, A. J. & Thollesson, M. (2004) Bilaterian phylogeny based on analyses of a region of the sodium-potassium ATPase beta-subunit gene. *Journal of Molecular Evolution,* 58(3): pp. 252-268.

Anderson, F. E. & Swofford, D. L. (2004) Should we be worried about long-branch attraction in real data sets? Investigations using metazoan 18S rDNA. *Molecular Phylogenetics and Evolution,* 33(2): pp. 440-451.

Andronescu, M., Aguirre-Hernandez, R., Condon, A. & Hoos, H. H. (2003) RNAsoft: a suite of RNA secondary structure prediction and design software tools. *Nucleic Acids Research,* 31(13): pp. 3416-3422.

Anton, H. (1995) *Calculus with analytic geometry,* New York ; Chichester, Wiley.

Archibald, J. K., Mort, M. E. & Crawford, D. J. (2003) Bayesian inference of phylogeny: a non-technical primer. *Taxon,* 52(2): pp. 187-191.

Arntzen, J. W. & Garciaparis, M. (1995) Morphological and Allozyme Studies of Midwife Toads (Genus Alytes), Including the Description of 2 New Taxa from Spain. *Contributions to Zoology,* 65(1): pp. 5-34.

Ashelford, K. E., Chuzhanova, N. A., Fry, J. C., Jones, A. J. & Weightman, A. J. (2005) At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. *Applied and Environmental Microbiology,* 71(12): pp. 7724-7736.

Ashelford, K. E., Chuzhanova, N. A., Fry, J. C., Jones, A. J. & Weightman, A. J. (2006) New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. *Applied and Environmental Microbiology,* 72(9): pp. 5734-5741.

Atkins, D. (1955) The Cyphonautes larvae of the Plymouth area and the metamorphosis of Membranipora membranacea (L.). *Journal of the Marine Biological Association of the United Kingdom,* 34: pp. 441-449.

Ballard, J. W. O. (1996) Combining data in phylogenetic analysis. *Trends in Ecology & Evolution,* 11(8): pp. 334-334.

Barnes, R. S. K. (2001) *The invertebrates : a synthesis,* Oxford, Blackwell Pub.

Becker, W. M., Kleinsmith, L. J., Hardin, J. & Becker, W. M. (2003) *Guide to microscopy,* San Francisco, Benjamin Cummings Pub. Co.

Belshaw, R. & Katzourakis, A. (2005) BlastAlign: a program that uses blast to align problematic nucleotide sequences. *Bioinformatics,* 21(1): pp. 122-123.

Ben Ali, A., Wuyts, J., De Wachter, R., Meyer, A. & Van de Peer, Y. (1999) Construction of a variability map for eukaryotic large subunit ribosomal RNA. *Nucleic Acids Research,* 27(14): pp. 2825-2831.

Benevolenskaya, E. V., Kogan, G. L., Tulin, A. V., Philipp, D. & Gvozdev, V. A. (1997) Segmented gene conversion as a mechanism of correction of 18S rRNA pseudogene located outside of rDNA cluster in D-melanogaster. *Journal of Molecular Evolution,* 44(6): pp. 646-651.

Bensasson, D., Zhang, D. X., Hartl, D. L. & Hewitt, G. M. (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology & Evolution,* 16(6): pp. 314-321.

Benton, M. J. (1994) Paleontological Data and Identifying Mass Extinctions. *Trends in Ecology & Evolution,* 9(5): pp. 181-185.

Benton, M. J. (1998) Molecular and morphological phylogenies of mammals: Congruence with stratigraphic data. *Molecular Phylogenetics and Evolution,* 9(3): pp. 398-407.

Benton, M. J., Hitchin, R. & Wills, M. A. (1999) Assessing congruence between cladistic and stratigraphic data. *Systematic Biology,* 48(3): pp. 581-596.

Benton, M. J., Wills, M. A. & Hitchin, R. (2000) Quality of the fossil record through time. *Nature,* 403(6769): pp. 534-537.

Bergsten, J. (2005) A review of long-branch attraction. *Cladistics,* 21(2): pp. 163-193.

Biffin, E., Harrington, M. G., Crisp, M. D., Craven, L. A. & Gadek, P. A. (2007) Structural partitioning, paired-sites models and evolution of the ITS transcript in Syzygium and Myrtaceae. *Molecular Phylogenetics and Evolution,* 43(1): pp. 124-139.

Blanchette, M., Green, E. D., Miller, W. & Haussler, D. (2004) Reconstructing large regions of an ancestral mammalian genome in silico. *Genome Res.,* 14(12): pp. 2412-2423.

Blanchette, M., Kent, W. J., Riemer, C., Elnitski, L., Smit, A. F. A., Roskin, K. M., Baertsch, R., Rosenbloom, K., Clawson, H., Green, E. D., Haussler, D. & Miller, W. (2004) Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Research,* 14(4): pp. 708-715.

Boore, J. L. & Brown, W. M. (1998) Big trees from little genomes: mitochondrial gene order as a phylogenetic tool. *Current Opinion in Genetics & Development,* 8(6): pp. 668-674.

Box, G. E. P. (1976) Science and Statistics. *Journal of the American Statistical Association,* 71(356): pp. 791-799.

Bradbury, S. & Royal Microscopical Society (Great Britain). (1989) *An introduction to the optical microscope,* Oxford, Oxford University Press [for] Royal Microscopical Society.

Bridge, D., Cunningham, C. W., Desalle, R. & Buss, L. W. (1995) Class-Level Relationships in the Phylum Cnidaria - Molecular and Morphological Evidence. *Molecular Biology and Evolution,* 12(4): pp. 679-689.

Brodsky, L. I., Ivanov, V. V., Kalaydzidis, Y. L., Leontovich, A. M., Nikolaev, V. K., Feranchuk, S. I. & Drachev, V. A. (1995) Genebee-Net - Internet-Based Server for Analyzing Biopolymers Structure. *Biochemistry-Moscow,* 60(8): pp. 923-928.

Brower, A. (1996) Combining data in phylogenetic analysis. *Trends in Ecology & Evolution,* 11(8): pp. 334-335.

Brower, A. V. Z. & Desalle, R. (1994) Practical and Theoretical Considerations for Choice of a DNA- Sequence Region in Insect Molecular Systematics, with a Short Review of Published Studies Using Nuclear Gene Regions. *Annals of the Entomological Society of America,* 87(6): pp. 702-716.

Brown, R. P. & Pestano, J. (1998) Phylogeography of skinks (Chalcides) in the Canary Islands inferred from mitochondrial DNA sequences. *Molecular Ecology,* 7(9): pp. 1183-1191.

Buckley, T. R., Simon, C., Flook, P. K. & Misof, B. (2000) Secondary structure and conserved motifs of the frequently sequenced domains IV and V of the insect mitochondrial large subunit rRNA gene. *Insect Molecular Biology,* 9(6): pp. 565-580.

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): pp. 384-397.

Burbrink, F. T. (2001) Systematics of the eastern ratsnake complex (Elaphe obsoleta). *Herpetological Monographs,* (15): pp. 1-53.

Cadman, P. S. & Ryland, J. S. (1996) Redescription of Alcyonidium mytili Dalyell, 1848 (Bryozoa: Ctenostomatida). *Zoological Journal of the Linnean Society,* 116(4): pp. 437-450.

Cancino, J. M., Hughes, R. N. & Orellana, M. C. (1994) A comparative study of larval release in bryozoans. IN: Hayward, P. J., Ryland, J. S. & Taylor, P. D. (Eds.) *Biology and PAleobiology of Bryozoans.* Fredensborg, Olsen and Olsen. pp. 41-46.

Carranza, S., Baguna, J. & Riutort, M. (1997) Are the Platyhelminthes a monophyletic primitive group? An assessment using 18S rDNA sequences. *Molecular Biology and Evolution,* 14(5): pp. 485-497.

Chastain, M. & Tinoco, I. (1991) Structural Elements in RNA. *Progress in Nucleic Acid Research and Molecular Biology,* 41: pp. 131-177.

Cheetham, A. H. (1968) Morphology and systematics of the bryozoan genus Metrarabdotos, Washington, Smithsonian Institution.

Cheetham, A. H. & Lorenz, D. M. (1976) A vector approach to size and shape comparisons among zooids in cheilostome bryozoans, Washington, Smithsonian Institution Press.

Cheetham Alan, H. & Cook, P. L. (1983) General features of class Gymnolaemata. IN: Boardman, R. S., Blake, D. B., Utgaard, J., Karklins, O. L., Cook, P. L., Sandberg, P. A., Lutaud, G. & Wood, T. S. (Eds.) *Treatise on Invertebrate Paleontology Part G, Bryozoa.* Boulder, Colo., Geological Society of America Lawrence. pp. 138-207.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. & Thompson, J. D. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Research,* 31(13): pp. 3497-3500.

Christensen, I., Haug, T. & Wiig, O. (1990) Morphometric Comparison of Minke Whales Balaenoptera- Acutorostrata from Different Areas of the North-Atlantic. *Marine Mammal Science,* 6(4): pp. 327-338.

Claxton, N. S., Fellers, T. J. & Davidson, M. W. ([no date]) Laser Scanning Confocal Microscopy. Olympus America.

Clendenon, J. L., Phillips, C. L., Sandoval, R. M., Fang, S. F. & Dunn, K. W. (2002) Voxx: a PC-based, near real-time volume rendering system for biological microscopy. *American Journal of Physiology-Cell Physiology,* 282(1): pp. C213-C218.

Cohen, B. L., Gawthrop, A. & Cavalier-Smith, T. (1998) Molecular phylogeny of brachiopods and phoronids based on nuclear-encoded small subunit ribosomal RNA gene sequences. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences,* 353(1378): pp. 2039-2061.

Collura, R. V., Auerbach, M. R. & Stewart, C. B. (1996) A quick, direct method that can differentiate expressed mitochondrial genes from their nuclear pseudogenes. *Current Biology,* 6(10): pp. 1337-1339.

Cook, P. L. (1973) Settlement and early colony development in some Cheilostomata. IN: Larwood, G. P. (Ed.) *Living and Fossil Bryozoa.* London, Academic Press. pp. 65-71.

Cook, P. L. & Chimonides, P. J. (1981) Morphology and Systematics of Some Rooted Cheilostome Bryozoa. *Journal of Natural History,* 15(1): pp. 97-134.

Cooper, M. S., D'Amico, L. A. & Henry, C. A. (1999) Confocal microscopic analysis of morphogenetic movements. *Methods In Cell Biology, Vol 59.* pp. 179-+.

Cooper, M. S., Szeto, D. P., Sommers-Herivel, G., Topezewski, J., Solnica-Krezel, L., Kang, H. C., Johnson, I. & Kimelman, D. (2005) Visualizing morphogenesis in transgenic zebrafish embryos using BODIPY TR methyl ester dye as a vital counterstain for GFP. *Developmental Dynamics,* 232(2): pp. 359-368.

Copley, R. R., Aloy, P., Russell, R. B. & Telford, M. J. (2004) Systematic searches for molecular synapomorphies in model metazoan genomes give some support for Ecdysozoa after accounting for the idiosyncrasies of Caenorhabditis elegans. *Evolution & Development,* 6(3): pp. 164-169.

Costello, D. P. & Henley, C. (1971) *Methods for obtaining and handling marine eggs and embryos.,* Woods Hole, MA, Marine Biological Laboratory.

Craig, S. F., D'Amato, M. E., Harley, M., Bishop, M., Bishop, J., Hughes, R. & Carvalho, G. R. (2001) Isolation and characterization of microsatellites in the bryozoan Crisia denticulata. *Molecular Ecology Notes,* 1(4): pp. 281-282.

Cummings, M. P., Otto, S. P. & Wakeley, J. (1995) Sampling Properties of DNA-Sequence Data in Phylogenetic Analysis. *Molecular Biology and Evolution,* 12(5): pp. 814-822.

Dams, E., Hendriks, L., Vandepeer, Y., Neefs, J. M., Smits, G., Vandenbempt, I. & Dewachter, R. (1988) Compilation of Small Ribosomal-Subunit RNA Sequences. *Nucleic Acids Research,* 16: pp. R87-R173.

Davis, M. (2005) *Scientific papers and presentations,* New York, Academic Press.

De Rijk, P., Wuyts, J. & De Wachter, R. (2003) RnaViz 2: an improved representation of RNA secondary structure. *Bioinformatics,* 19(2): pp. 299-300.

Delest, D. A. M. & Dulucq, J.-P. D. S. (2006) Efficient drawing of RNA secondary structure. *Journal of Graph Algorithms and Applications,* 10(2): pp. 329-351.

Derijk, P. & Dewachter, R. (1993) Dcse, an Interactive Tool for Sequence Alignment and Secondary Structure Research. *Computer Applications in the Biosciences,* 9(6): pp. 735-740.

Derijk, P., Neefs, J. M., Vandepeer, Y. & Dewachter, R. (1992) Compilation of Small Ribosomal-Subunit RNA Sequences. *Nucleic Acids Research,* 20: pp. 2075-2089.

Dick, M. H., Freeland, J. R., williams, L. P. & Coggeshall-Burr, M. (2000) Use of 16S mitochondrial ribosomal DNA sequences to investigate sister-group relationships among gymnolaemate bryozoans. IN: Herrera-Cubilla, A. & Jackson, J. B. C. (Eds.) *Proceedings of the 11th International Bryozoology Association Conference.* Smithsonian Tropical Research Institute, Panamá.

Dick, M. H., Herrera-Cubilla, A. & Jackson, J. B. C. (2003) Molecular phylogeny and phylogeography of free-living Bryozoa (Cupuladriidae) from both sides of the Isthmus of Panama. *Molecular Phylogenetics and Evolution,* 27(3): pp. 355-371.

Dick, M. H. & Mawatari, S. F. (2005) Morphological and molecular concordance of Rhynchozoon clades (Bryozoa, Cheilostomata) from Alaska. *Invertebrate Biology,* 124(4): pp. 344-354.

Diekmann, O. E., Bak, R. P. M., Stam, W. T. & Olsen, J. L. (2001) Molecular genetic evidence for probable reticulate speciation in the coral genus Madracis from a Caribbean fringing reef slope. *Marine Biology,* 139(2): pp. 221-233.

Diekmann, O. E., Olsen, J. L., Stam, W. T. & Bak, R. P. M. (2003) Genetic variation within Symbiodinium clade B from the coral genus Madracis in the Caribbean (Netherlands Antilles). *Coral Reefs,* 22(1): pp. 29-33.

Dunn, K. W., Sandoval, R. M., Kelly, K. J., Dagher, P. C., Tanner, G. A., Atkinson, S. J., Bacallao, R. L. & Molitoris, B. A. (2002) Functional studies of the kidney of living animals using multicolor two-photon microscopy. *American Journal of Physiology-Cell Physiology,* 283(3): pp. C905-C916.

Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research,* 32(5): pp. 1792-1797.

Ehrenberg, C. G. (1831) Symbolae Physicae, seu Icones et descriptiones Corporum Naturalium novorum aut minus cognitorum, quae ex itineribus per Libyam,

Aegyptum, Nubiam, Dongalam, Syriam, Arabiam et Habissiniam ... studio annis 1820-25 redierunt ...Pars Zoologica Vol 4 Animalia Evertebrata exclusis, Berlin.

Eldredge, N. & Gould, S. J. (1972) Punctuated equilibria: an alternative to phyletic gradualism. IN: Schopf, T. J. M. (Ed.) *Models in paleobiology.* San Francisco, Freeman, Cooper & Co. pp. vi, 250.

Ellis, R. E., Sulston, J. E. & Coulson, A. R. (1986) The Rdna of C-Elegans - Sequence and Structure. *Nucleic Acids Research,* 14(5): pp. 2345-2364.

Felsenstein, J. (1981) Evolutionary Trees from DNA-Sequences - a Maximum-Likelihood Approach. *Journal of Molecular Evolution,* 17(6): pp. 368-376.

Felsenstein, J. (2004) *Inferring phylogenies,* Sunderland, Mass., Sinauer Associates.

Field, A. P. (2000) Discovering statistics using SPSS for Windows : advanced techniques for the beginner, London, Sage.

Field, K. G., Olsen, G. J., Lane, D. J., Giovannoni, S. J., Ghiselin, M. T., Raff, E. C., Pace, N. R. & Raff, R. A. (1988) Molecular Phylogeny of the Animal Kingdom. *Science,* 239(4841): pp. 748-753.

Flynn, J. J. & Nedbal, M. A. (1998) Phylogeny of the Carnivora (Mammalia): Congruence vs incompatibility among multiple data sets. *Molecular Phylogenetics and Evolution,* 9(3): pp. 414-426.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology,* 3(5): pp. 294-299.

Friedlander, T. P., Regier, J. C. & Mitter, C. (1992) Nuclear Gene-Sequences for Higher Level Phylogenetic Analysis - 14 Promising Candidates. *Systematic Biology,* 41(4): pp. 483-490.

Friedlander, T. P., Regier, J. C. & Mitter, C. (1994) Phylogenetic Information-Content of 5 Nuclear Gene-Sequences in Animals - Initial Assessment of Character Sets from Concordance and Divergence Studies. *Systematic Biology,* 43(4): pp. 511-525.

Frost, S. R., Marcus, L. F., Bookstein, F. L., Reddy, D. P. & Delson, E. (2003) Cranial allometry, phylogeography, and systematics of large- bodied papionins (Primates : Cercopithecinae) inferred from geometric morphometric analysis of landmark data. *Anatomical Record Part a-Discoveries in Molecular Cellular and Evolutionary Biology,* 275A(2): pp. 1048-1072.

Gallardo, M. H. & Palma, E. (1990) Systematics of Oryzomys-Longicaudatus (Rodentia, Muridae) in Chile. *Journal of Mammalogy,* 71(3): pp. 333-342.

Gappa, J. J. L. (1986) A New Bryozoan Genus from the Weddell Sea, Antarctica. *Polar Biology,* 6(2): pp. 103-105.

Gerbi, S. A. (1986) The Evolution of Eukaryotic Ribosomal DNA. *Biosystems,* 19(4): pp. 247-258.

Giese, A. C., Pearse, J. S. & Pearse, V. B. (1991) *Reproduction of marine invertebrates,* California, Boxwood Press.

Gillespie, J. J., McKenna, C. H., Yoder, M. J., Gutell, R. R., Johnston, J. S., Kathirithamby, J. & Cognato, A. I. (2005) Assessing the odd secondary structural properties of nuclear small subunit ribosomal RNA sequences (18S) of the twisted-wing parasites (Insecta : Strepsiptera). *Insect Molecular Biology,* 14(6): pp. 625-643.

Gillespie, J. J., Yoder, M. J. & Wharton, R. A. (2005) Predicted secondary structure for 28S and 18S rRNA from Ichneumonoidea (Insecta : Hymenoptera : Apocrita): Impact on sequence alignment and phylogeny estimation. *Journal of Molecular Evolution,* 61(1): pp. 114-137.

Giribet, G. (2002) Current advances in the phylogenetic reconstruction of metazoan evolution. A new paradigm for the Cambrian explosion? *Molecular Phylogenetics and Evolution,* 24(3): pp. 345-357.

Giribet, G., Carranza, S., Baguna, J., Riutort, M. & Ribera, C. (1996) First molecular evidence for the existence of a Tardigrada plus Arthropoda clade. *Molecular Biology and Evolution,* 13(1): pp. 76-84.

Giribet, G., Carranza, S., Riutort, M., Baguna, J. & Ribera, C. (1999) Internal phylogeny of the Chilopoda (Myriapoda, Arthropoda) using complete 18S rDNA and partial 28S rDNA sequences. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences,* 354(1380): pp. 215-222.

Giribet, G., Distel, D. L., Polz, M., Sterrer, W. & Wheeler, W. C. (2000) Triploblastic relationships with emphasis on the acoelomates and the position of Gnathostomulida, Cycliophora, Platyhelminthes, and Chaetognatha: A combined approach of 18S rDNA sequences and morphology. *Systematic Biology,* 49(3): pp. 539-562.

Glenner, H., Hansen, A. J., Sorensen, M. V., Ronquist, F., Huelsenbeck, J. P. & Willerslev, E. (2004) Bayesian inference of the metazoan phylogeny: A combined molecular and morphological approach. *Current Biology,* 14(18): pp. 1644-1649.

Glenner, H., Hansen, A. J., Sorensen, M. V., Ronquist, F., Huelsenbeck, J. P. & Willerslev, E. (2005) Bayesian inference of the metazoan phylogeny: A combined

molecular and morphological approach (vol 12, pg 1828, 2004). *Current Biology,* 15(4): pp. 392-393.

Glenner, H., Lutzen, J. & Takahashi, T. (2003) Molecular and morphological evidence for a monophyletic clade of asexually reproducing Rhizocephala: Polyascus, new genus (Cirripedia). *Journal of Crustacean Biology,* 23(3): pp. 548-557.

Godfray, H. C. J. (2002) Towards taxonomy's 'glorious revolution'. *Nature,* 420(6915): pp. 461-461.

Godfray, H. C. J. (2002) Challenges for taxonomy - The discipline will have to reinvent itself if it is to survive and flourish. *Nature,* 417(6884): pp. 17-19.

Goffredi, S. K., Jones, W. J., Scholin, C. A., Marin, R. & Vrijenhoek, R. C. (2006) Molecular detection of marine invertebrate larvae. *Marine Biotechnology,* 8(2): pp. 149-160.

Goldman, N., Anderson, J. P. & Rodrigo, A. G. (2000) Likelihood-based tests of topologies in phylogenetics. *Systematic Biology,* 49(4): pp. 652-670.

Gordon, D. P. (2000) Towards a phylogeny of Cheilostomes - Morphological Models of frontal wall/shield evolution. IN: Herrera-Cubilla, A. & Jackson, J. B. C. (Eds.) *Proceedings of the 11th International Bryozoology Association Conference.* Smithsonian Tropical Research Institute, Panamá.

Gowri-Shankar, V. & Jow, H. (2006) PHASE: a software Package for Phylogenetics and Sequnce Evolution. 2.0 ed., Unniversity of Manchester.

Graham, A. (2000) Animal phylogeny: Root and branch surgery. *Current Biology,* 10(1): pp. R36-R38.

Graur, D. & Martin, W. (2004) Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends In Genetics,* 20(2): pp. 80-86.

Gutell, R. R., Larsen, N. & Woese, C. R. (1994) Lessons from an Evolving Ribosomal-RNA - 16s and 23s Ribosomal-RNA Structures from a Comparative Perspective. *Microbiological Reviews,* 58(1): pp. 10-26.

Gutell, R. R., Power, A., Hertz, G. Z., Putz, E. J. & Stormo, G. D. (1992) Identifying Constraints on the Higher-Order Structure of RNA - Continued Development and Application of Comparative Sequence-Analysis Methods. *Nucleic Acids Research,* 20(21): pp. 5785-5795.

Gutell, R. R., Weiser, B., Woese, C. R. & Noller, H. F. (1985) Comparative Anatomy of 16-S-Like Ribosomal-RNA. *Progress in Nucleic Acid Research and Molecular Biology,* 32: pp. 155-216.

Hageman, S. J., Bayers, M. M. & Todd, C. D. (1999) Partitioning phenotypic variation: genotypic, environmental and residual components from bryozoan skeletal morphology. *Journal of Natural History,* 33(11): pp. 1713-1735.

Halanych, K. M. (1995) Evidence from 18s Ribosomal DNA That the Lophophorates Are Protostome Animals. *Science,* 268(5210): pp. 485-485.

Halanych, K. M., Bacheller, J. D., Aguinaldo, A. M. A., Liva, S. M., Hillis, D. M. & Lake, J. A. (1995) Evidence from 18s Ribosomal DNA That the Lophophorates Are Protostome Animals Inarticulate. *Science,* 267(5204): pp. 1641-1643.

Halanych, K. M., Lutz, R. A. & Vrijenhoek, R. C. (1998) Evolutionary origins and age of vestimentiferan tube-worms. *Cahiers De Biologie Marine,* 39(3-4): pp. 355-358.

Halanych, K. M. & Passamaneck, Y. (2001) A brief review of metazoan phylogeny and future prospects in Hox-research. *American Zoologist,* 41(3): pp. 629-639.

Halanych, K. M. & Passamaneck, Y. (2003) Recent progress on resolving lophotrochozoan phylogeny. *Integrative and Comparative Biology,* 43(6): pp. 1074-1074.

Hall, B. G. (2004) Phylogenetic trees made easy: a how-to manual for molecular biologists., Sunderland, Massachusetts, Sinauer Associates.

Hall, S. R., Taylor, P. D., Davis, S. A. & Mann, S. (2002) Electron diffraction studies of the calcareous skeletons of bryozoans. *Journal of Inorganic Biochemistry,* 88(3-4): pp. 410-419.

Hao, J. S., Li, C. X., Sun, X. Y. & Yang, Q. (2005) Phylogeny and divergence time estimation of cheilostome bryozoans based on mitochodrial 16S rRNA sequences. *Chinese Science Bulletin,* 50(12): pp. 1205-1211.

Hao, J. S., Yang, Q., Li, C.-X., Zhang, D. X. & Sun, X. Y. (2002) Preliminary study on molecular phylogeny of bryozoans based on 18S rRNA gene sequences. *Acta Micropalaeontologica Sinica,* 19: pp. 199-205.

Hasegawa, M., Kishino, H. & Yano, T. A. (1985) Dating of the Human-Ape Splitting by a Molecular Clock of Mitochondrial-DNA. *Journal of Molecular Evolution,* 22(2): pp. 160-174.

Hatton-Ellis, T. W., Noble, L. R. & Okamura, B. (1998) Genetic variation in a freshwater bryozoan. I: Populations in the Thames basin, UK. *Molecular Ecology,* 7(11): pp. 1575-1585.

Haugland, R. P. (2002) Handbook of Fluorescent Probes and Research Products., Molecular Probes.

Hausdorf, B., Helmkampf, M., Meyer, A., Witek, A., Herlyn, H., Bruchhaus, I., Hankeln, T., Struck, T. H. & Lieb, B. (2007) Spiralian Phylogenomics Supports the Resurrection of Bryozoa Comprising Ectoprocta and Entoprocta. *Mol Biol Evol,* 24(12): pp. 2723-2729.

Hayward, P. J. (1985) *Ctenostome bryozoans : keys and notes for the identification of the species,* London ; E.J. Brill/W. Backhuys for the Linnean Society of London and the Estuarine and Brackish-Water Sciences Association,.

Hayward, P. J. (1992) Studies on the Bryozoa. *School of Biological Sciences.* Swansea, University of Wales Swansea.

Hayward, P. J. (1993) New Species of Cheilostomate Bryozoa from Antarctica and the Sub-Antarctic Southwest Atlantic. *Journal of Natural History,* 27(6): pp. 1409-1430.

Hayward, P. J., Muir, A. I., Linnean Society of London., Estuarine and Coastal Sciences Association. & Field Studies Council (Great Britain) (1998) *Cheilostomatous Bryozoa : part 1 Aeteoidea -Cribrilinoidea,* Shrewsbury, Published for the Linnean Society of London and the Estuarine and Coastal Sciences Association by Field Studies Council.

Hayward, P. J. & Ryland, J. S. (1979) *British ascophoran bryozoans : keys and notes for the identification of the species,* London, Academic Press for the Linnean Society of London and the Estuarine and Brackish-water Sciences Association.

Hayward, P. J. & Ryland, J. S. (1993) Taxonomy of 6 Antarctic Anascan Bryozoa. *Antarctic Science,* 5(2): pp. 129-136.

Hayward, P. J. & Ryland, J. S. (1999) *Cheilostomatous Bryozoa : part 2 Hippothooidea -Celleporoidea,* Shrewsbury, Published for the Linnean Society of London and the Estuarine and Coastal Sciences Association by Field Studies Council.

Hayward, P. J. & Ryland, J. S. (2000) *Handbook of the marine fauna of north-west Europe,* Oxford, Oxford University Press.

Hayward, P. J., Ryland, J. S., Estuarine and Brackish-water Sciences Association. & Linnean Society of London. (1985) *Cyclostome Bryozoans : keys and notes for identification of the species,* London, Published for the Linnean Society of London and the Estuarine and Brackish-Water Sciences Association by E.J. Brill/W. Backhuys.

Hayward, P. J., Taylor, P. D. & Ryland, J. S. (1994) Biology and palaeobiology of bryozoans : Proceedings of the 9th International Bryozoology Conference, School of Biological Sciences, University of Wales, Swansea, 1992, Denmark, Olsen and Olsen.

Hayward, P. J. & Thorpe, J. P. (1995) Some British Species of Schizomavella (Bryozoa, Cheilostomatida). *Journal of Zoology,* 235: pp. 661-676.

Hendriks, L., Debaere, R., Vandepeer, Y., Neefs, J., Goris, A. & Dewachter, R. (1991) The Evolutionary Position Of The Rhodophyte Porphyra-Umbilicalis And The Basidiomycete Leucosporidium-Scottii Among Other Eukaryotes As Deduced From Complete Sequences Of Small Ribosomal-Subunit RNA. *Journal Of Molecular Evolution,* 32(2): pp. 167-177.

Hendriks, L., Van Broeckhoven, C., Vandenberghe, A., Van de Peer, Y. & De Wachter, R. (1988) Primary and secondary structure of the 18S ribosomal RNA of the bird spider Eurypelma californica and evolutionary relationships among eukaryotic phyla. *European Journal of Biochemistry,* 177: pp. 15-20.

Hepperle, D., Nozaki, H., Hohenberger, S., Huss, V. A. R., Morita, E. & Krienitz, L. (1998) Phylogenetic position of the phacotaceae within the chlamydophyceae as revealed by analysis of 18S rDNA and rbcL sequences. *Journal of Molecular Evolution,* 47(4): pp. 420-430.

Heuermann, K. & Cosgrove, J. (2001) S-Gal (TM): An autoclavable dye for color selection of cloned DNA inserts. *Biotechniques,* 30(5): pp. 1142-+.

Hickson, R. E., Simon, C., Copper, 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): pp. 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): pp. 530-539.

Higgins, D. G. & Sharp, P. M. (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene,* 73(1): pp. 237-244.

Hillis, D. M. (1987) Molecular Versus Morphological Approaches to Systematics. *Annual Review of Ecology and Systematics,* 18: pp. 23-42.

Hillis, D. M. & Dixon, M. T. (1991) Ribosomal DNA - Molecular Evolution and Phylogenetic Inference. *Quarterly Review of Biology,* 66(4): pp. 410-453.

Hillis, D. M., Mable, B. K. & Moritz, C. (1996) *Molecular methods; systematics,* Sunderland, Mass., Sinauer Associates.

Hillis, D. M., Pollock, D. D., McGuire, J. A. & Zwickl, D. J. (2003) Is sparse taxon sampling a problem for phylogenetic inference? *Systematic Biology,* 52(1): pp. 124-126.

Höchsmann, M., Töller, T., Giegerich, R. & Kurtz, S. (2003) Local Similarity in RNA Secondary Structures. *Proceedings of the IEEE Bioinformatics Conference 2003(CSB 2003).*

Hofacker, I. L. (2003) Vienna RNA secondary structure server. *Nucleic Acids Research,* 31(13): pp. 3429-3431.

Holton, T. A. & Graham, M. W. (1991) A Simple and Efficient Method for Direct Cloning of Pcr Products Using Ddt-Tailed Vectors. *Nucleic Acids Research,* 19(5): pp. 1156-1156.

Hondt, J. L. d. (1977) Valeur systématique de la structure larvaire et des particularités de la morphogenése post-larvaire chez les Bryozoaires Gymnolaemates. *Gegenbaurs Morphologisches Jahrbuch,* 123(3): pp. 463-483.

Hua-Bin, Z., Qun, Y. & Jia-Sheng, H. (2006) 28s rDNA molecular morphology of the main bryozoan lineages and its phylogenetic significance. *Acta Zootaxonomica Sinica,* 31(2): pp. 247-255.

Hudelot, C., Gowri-Shankar, V., Jow, H., Rattray, M. & Higgs, P. G. (2003) RNA-based phylogenetic methods: application to mammalian mitochondrial RNA sequences. *Molecular Phylogenetics And Evolution,* 28(2): pp. 241-252.

Hudson, J. D. (1998) Discussion on the Cretaceous-Tertiary biotic transition. *Journal of the Geological Society,* 155: pp. 413-415.

Huelsenbeck, J. P., Bull, J. J. & Cunningham, C. W. (1996) Combining data in phylogenetic analysis. *Trends in Ecology & Evolution,* 11(4): pp. 152-158.

Huelsenbeck, J. P., Bull, J. J. & Cunningham, C. W. (1996) Combining data in phylogenetic analysis - Reply. *Trends in Ecology & Evolution,* 11(8): pp. 335-335.

Huelsenbeck, J. P., Larget, B., Miller, R. E. & Ronquist, F. (2002) Potential applications and pitfalls of Bayesian inference of phylogeny. *Systematic Biology,* 51(5): pp. 673-688.

Huelsenbeck, J. P. & Ronquist, F. (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics,* 17(8): pp. 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(5550): pp. 2310-2314.

Humphries, E. M. (1977) Larval Behavior and Post-Larval Development in Parasmittina nitida Morphotype B (Bryozoa: Cheilostomata). *Amer. Zool.,* 17(1): pp. 5-20.

Hunter, E. & Fusetani, N. (1995) Studies on the effects of larval swimming time on settlement, metamorphosis and post-larval development of Bugula neretina (Cheilostomatida). IN: Gordon, D. P., Smith, A. M. & Grant-Makie, J. A. (Eds.) *Bryozoans in Space and Time: Proceedings of the 10th International Bryozoology Conference.* Wellington, New Zealand, National Institute of Water & Atmospheric Research Ltd, Wellington.

Hunter, E., Okano, K., Tomono, Y. & Fusetani, N. (1998) Functional partitioning of energy reserves by larvae of the marine bryozoan Bugula neritina (L.). *Journal Of Experimental Biology,* 201(20): pp. 2857-2865.

Hunter, E., Shimizu, K. & Fusetani, N. (1999) Role of protein in larval swimming and metamorphosis of Bugula neritina (Bryozoa : Cheilostomatida). *Marine Biology,* 133(4): pp. 701-707.

Huysmans, E. & Dewachter, R. (1986) Compilation of Small Ribosomal-Subunit RNA Sequences. *Nucleic Acids Research,* 14: pp. R73-R118.

Hyman, L. H. (1959) *Invertebrates:Smaller Coelomate Groups,* New York, McGraw-Hill.

Jackson, J. B. C. (1986) Modes of Dispersal of Clonal Benthic Invertebrates - Consequences for Species Distributions and Genetic-Structure of Local-Populations. *Bulletin of Marine Science,* 39(2): pp. 588-606.

Jackson, J. B. C. & Cheetham, A. H. (1990) Evolutionary Significance of Morphospecies - a Test with Cheilostome Bryozoa. *Science,* 248(4955): pp. 579-583.

Jacobs, D. K., Hughes, N. C., Fitz-Gibbon, S. T. & Winchell, C. J. (2005) Terminal addition, the Cambrian radiation and the Phanerozoic evolution of bilaterian form. *Evolution & Development,* 7(6): pp. 498-514.

Jaffe, S. (2003) Putting a pretty face on multiple sequence alignment. *Scientist,* 17(19): pp. 33-33.

Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G. & Gibson, T. J. (1998) Multiple sequence alignment with Clustal X. *Trends in Biochemical Sciences,* 23(10): pp. 403-405.

Jebram, D. (1973) The importance of Different Growth Directions in Phylactolaema and Gymnolaemata for reconstructing the Phylogeny of the Bryozoa. IN: Larwood, G.

P. (Ed.) *Living and Fossil Bryozoa.* London & New York, Academic Press. pp. 565-592.

Jefferson, T. A. (2002) Preliminary analysis of geographic variation in cranial morphometrics of the finless porpoise (Neophocaena phocaenoides). *Raffles Bulletin of Zoology*: pp. 3-14.

Jenner, R. A. (2000) Evolution of animal body plans: the role of metazoan phylogeny at the interface between pattern and process. *Evolution & Development,* 2(4): pp. 208-221.

Jow, H., Hudelot, C., Rattray, M. & Higgs, P. G. (2002) Bayesian phylogenetics using an RNA substitution model applied to early mammalian evolution. *Molecular Biology and Evolution,* 19(9): pp. 1591-1601.

Karlin, S. & Altschul, S. F. (1990) Methods for Assessing the Statistical Significance of Molecular Sequence Features by Using General Scoring Schemes. *Proceedings of the National Academy of Sciences of the United States of America,* 87(6): pp. 2264-2268.

Kass, R. E. & Raftery, A. E. (1995) Bayes Factors. *Journal Of The American Statistical Association,* 90(430): pp. 773-795.

Katoh, K., Kuma, K., Toh, H. & Miyata, T. (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research,* 33(2): pp. 511-518.

Katoh, K., Misawa, K., Kuma, K. & Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research,* 30(14): pp. 3059-3066.

Katoh, K. & Toh, H. (2007) PartTree: an algorithm to build an approximate tree from a large number of unaligned sequences. *Bioinformatics,* 23(3): pp. 372-374.

Kim, H. G., Kim, H. S., Hwang, H. J., Chung, S. K., Lee, J. M. & Chung, D. K. (2004) Construction of a pTOC-T vector using GST-ParE toxin for direct cloning and selection of PCR products. *Biotechnology Letters,* 26(21): pp. 1659-1663.

Kimura, M. (1968) Evolutionary rate at the Molecular Level. *Nature,* 217: pp. 624-626.

Kimura, M. (1981) Estimation of Evolutionary Distances between Homologous Nucleotide-Sequences. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences,* 78(1): pp. 454-458.

Kjer, K. M. (1995) Use of Ribosomal-RNA 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): pp. 314-330.

Kjer, K. M. (1997) Conserved primary and secondary structural motifs of amphibian 12S rRNA, domain III. *Journal of Herpetology,* 31(4): pp. 599-604.

Kjer, K. M. (2004) Aligned 18S and insect phylogeny. *Systematic Biology,* 53(3): pp. 506-514.

Kjer, K. M., Baldridge, G. D. & Fallon, A. M. (1994) Mosquito Large Subunit Ribosomal-RNA - Simultaneous Alignment of Primary and Secondary Structure. *Biochimica Et Biophysica Acta-Gene Structure and Expression,* 1217(2): pp. 147-155.

Kluge, G. A. (1975) *Мшанки северных морей СССР (Bryozoa of the northern seas of the USSR),* New Delhi, Published for the Smithsonian Institution Washington by Amerind Pub. Co.

Knudsen, B. & Hein, J. (2003) Pfold: RNA secondary structure prediction using stochastic context-free grammars. *Nucleic Acids Research,* 31(13): pp. 3423-3428.

Kosakovsky-Pond, S. L. K., Mannino, F. V., Gravenor, M. B., Muse, S. V. & Frost, S. D. W. (2007) Evolutionary model selection with a genetic algorithm: A case study using stem RNA. *Molecular Biology and Evolution,* 24(1): pp. 159-170.

Kowalewski, M., Dyreson, E., Marcot, J. D., Vargas, J. A., Flessa, K. W. & Hallman, D. P. (1997) Phenetic discrimination of biometric simpletons: paleobiological implications of morphospecies in the lingulide brachiopod Glottidia. *Paleobiology,* 23(4): pp. 444-469.

Kuittinen, H., Aguade, M., Charlesworth, D., Haan, A. D. E., Lauga, B., Mitchell-Olds, T., Oikarinen, S., Ramos-Onsins, S., Stranger, B., Van Tienderen, P. & Savolainen, O. (2002) Primers for 22 candidate genes for ecological adaptations in Brassicaceae. *Molecular Ecology Notes,* 2(3): pp. 258-262.

Kwak, J. H. & Kim, M. Y. (1995) Construction of T-Vector for Direct Cloning and Expression of Cloned Genes in Escherichia coli. *Analytical Biochemistry,* 228(1): pp. 178-180.

Kwon, O. Y., Ogino, K. & Ishikawa, H. (1991) The Longest 18S Ribosomal-RNA Ever Known - Nucleotide-Sequence and Presumed Secondary Structure of the 18S Ribosomal-RNA of the Pea Aphid, Acyrthosiphon-Pisum. *European Journal of Biochemistry,* 202(3): pp. 827-833.

Lafontaine, D. & Tollervey, D. (2001) Ribosomal RNA. *Encyclopaedia of life sciences.* Chichester Nature Publishing Group.

Larsen, N. (1992) Higher-Order Interactions in 23s Ribosomal-RNA. *Proceedings of the National Academy of Sciences of the United States of America,* 89(11): pp. 5044-5048.

Larsen, N. & Zwieb, C. (1991) SRP-RNA Sequence Alignment and Secondary Structure. *Nucleic Acids Research,* 19(2): pp. 209-215.

Larwood, G. P. & Taylor, P. D. (1979) Early Structural and Ecological Diversification in the Bryozoa. IN: House, M. R. (Ed.) *The Origin of Major Invertebrate Groups.* London, New York and San Francisco, Academic Press. pp. 209-234.

Lee, C. E. & Frost, B. W. (2002) Morphological stasis in the Eurytemora affinis species complex (Copepoda : Temoridae). *Hydrobiologia,* 480(1-3): pp. 111-128.

Leise, E. M. (1996) Selective retention of the fluorescent dye DASPEI in a larval gastropod mollusc after paraformaldehyde fixation. *Microscopy Research and Technique,* 33(6): pp. 496-500.

Leontis, N. B., Stombaugh, J. & Westhof, E. (2002) The non-Watson-Crick base pairs and their associated isostericity matrices. *Nucleic Acids Research,* 30(16): pp. 3497-3531.

Lewis, J. R. (1964) *The ecology of rocky shores,* London, English Universities Press.

Lewis, P. O. (2001) Phylogenetic systematics turns over a new leaf. *Trends in Ecology & Evolution,* 16(1): pp. 30-37.

Lidgard, S. & Buckley, G. A. (1994) Toward a morphological species concept in cheilostomates: phenotypic variation in Adeonellopsis yarraensis (Waters). IN: Hayward, P. J., Taylor, P. D. & Ryland, J. S. (Eds.) *Biology and palaeobiology of bryozoans : proceedings of the 9th International Bryozoology Conference, School of Biological Sciences, University of Wales, Swansea, 1992.* Denmark, Olsen and Olsen. pp. 101-106.

Lincoln, R. J., Sheals, J. G. & British Museum (Natural History). (1979) *Invertebrate animals : collection and preservation,* Cambridge, British Museum (Natural History) and Cambridge University Press.

Lockhart, 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): pp. 605-612.

Longin, A., Souchier, C., Ffrench, M. & Bryon, P. A. (1993) Comparison of Anti-Fading Agents Used in Fluorescence Microscopy - Image-Analysis and Laser Confocal Microscopy Study. *Journal of Histochemistry & Cytochemistry,* 41(12): pp. 1833-1840.

Lopez-Gonzalez, C., Presley, S. J., Owen, R. D. & Willig, M. R. (2001) Taxonomic status of Myotis (Chiroptera : Vespertilionidae) in Paraguay. *Journal of Mammalogy,* 82(1): pp. 138-160.

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., Lai, T., Steppi, S., Jobb, G., Forster, W., Brettske, I., Gerber, S., Ginhart, A. W., Gross, O., Grumann, S., Hermann, S., Jost, R., Konig, A., Liss, T., Lussmann, R., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A. & Schleifer, K. H. (2004) ARB: a software environment for sequence data. *Nucleic Acids Research,* 32(4): pp. 1363-1371.

Lutzoni, F., Wagner, P., Reeb, V. & Zoller, S. (2000) Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses without violating positional homology. *Systematic Biology,* 49(4): pp. 628-651.

Lynch, W. F. (1947) The behaviour and metamorphosis of the larva of Bugula neritina (Linnaeus): Experimental modification of the length of the free-swimming period and the responces of the larvae to light and gravity. *Biological Bulletin,* 92(February-June): pp. 115-150.

Mackey, L. Y., Winnepenninckx, B., DeWachter, R., Backeljau, T., Emschermann, P. & Garey, J. R. (1996) 18S rRNA suggests that Entoprocta are protostomes, unrelated to Ectoprocta. *Journal of Molecular Evolution,* 42(5): pp. 552-559.

Mackie, J. A., Keough, M. J., Norman, J. A. & Christidis, L. (2001) Mitochondrial evidence of geographical isolation within Bugula dentata. IN: Wyse Jackson, P. N., Buttler, C. J. & Spencer Jones, M. E. (Eds.) *Bryozoan studies.* Lisse, Swets & Zeitlinger. pp. 199-206.

MacLeod, N., Rawson, P. F., Forey, P. L., Banner, F. T., BoudagherFadel, M. K., Bown, P. R., Burnett, J. A., Chambers, P., Culver, S., Evans, S. E., Jeffery, C., Kaminski, M. A., Lord, A. R., Milner, A. C., Milner, A. R., Morris, N., Owen, E., Rosen, B. R., Smith, A. B., Taylor, P. D., Urquhart, E. & Young, J. R. (1997) The Cretaceous-Tertiary Biotic Transition. *Journal of the Geological Society,* 154: pp. 265-292.

Maidak, B. L., Cole, J. R., Lilburn, T. G., Parker, C. T., Saxman, P. R., Farris, R. J., Garrity, G. M., Olsen, G. J., Schmidt, T. M. & Tiedje, J. M. (2001) The RDP-II (Ribosomal Database Project). *Nucleic Acids Research,* 29(1): pp. 173-174.

Maidak, B. L., Cole, J. R., Lilburn, T. G., Parker, C. T., Saxman, P. R., Stredwick, J. M., Garrity, G. M., Li, B., Olsen, G. J., Pramanik, S., Schmidt, T. M. & Tiedje, J. M. (2000) The RDP (Ribosomal Database Project) continues. *Nucleic Acids Research,* 28(1): pp. 173-174.

Maidak, B. L., Cole, J. R., Parker, C. T., Garrity, G. M., Larsen, N., Li, B., Lilburn, T. G., McCaughey, M. J., Olsen, G. J., Overbeek, R., Pramanik, S., Schmidt, T. M., Tiedje, J. M. & Woese, C. R. (1999) A new version of the RDP (Ribosomal Database Project). *Nucleic Acids Research,* 27(1): pp. 171-173.

Maidak, B. L., Olsen, G. J., Larsen, N., Overbeek, R., McCaughey, M. J. & Woese, C. R. (1996) The Ribosomal Database Project (RDP). *Nucleic Acids Research,* 24(1): pp. 82-85.

Maidak, B. L., Olsen, G. J., Larsen, N., Overbeek, R., McCaughey, M. J. & Woese, C. R. (1997) The RDP (Ribosomal Database Project). *Nucleic Acids Research,* 25(1): pp. 109-110.

Maley, L. E. & Marshall, C. R. (1998) Evolution - The coming of age of molecular systematics. *Science,* 279(5350): pp. 505-506.

Marchuk, D., Drumm, M., Saulino, A. & Collins, F. S. (1991) Construction of T-Vectors, a Rapid and General System for Direct Cloning of Unmodified Pcr Products. *Nucleic Acids Research,* 19(5): pp. 1154-1154.

Mariscal, R. N. (1965) The Adult and Larval Morphology and Life History of the Entoproct Barentsia gracilis (M. Sars, 1835). *Journal of Morphology,* 116: pp. 311-338.

Mathews, D. H., Sabina, J., Zuker, M. & Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *Journal of Molecular Biology,* 288(5): pp. 911-940.

Mattern, D. & Schlegel, M. (2001) Molecular evolution of the small subunit ribosomal DNA in woodlice (Crustacea, Isopoda, Oniscidea) and implications for Oniscidean phylogeny. *Molecular Phylogenetics and Evolution,* 18(1): pp. 54-65.

Maturo, F. J. S. & Thorpe, J. P. (1979) The relationship of Schizomavella linearis (Hassal) and S. hastata (Hincks). IN: Larwood, G. P. & Abbott, M. B. (Eds.) *Advances in Bryozoology.* London, Academic Press. pp. 431-442.

Mayr, E. (1968) Bryozoa versus Ectoprocta. *Systematic Zoology,* 17(2): pp. 213-217.

Mayr, E. (1992) Speciational Evolution or Punctuated Equilibria. IN: Somit, A. & Peterson Steven, A. (Eds.) *The dynamics of Evolution.* New York, Cornell University Press. pp. 21-48.

McGovern, T. M. & Hellberg, M. E. (2002) Cryptic species, cryptic endosymbionts, and geographic variation in chemical defenses in the bryozoan Bugula neritina. *Integrative and Comparative Biology,* 42(6): pp. 1276-1276.

McGovern, T. M. & Hellberg, M. E. (2003) Cryptic species, cryptic endosymbionts, and geographical variation in chemical defences in the bryozoan Bugula neritina. *Molecular Ecology,* 12(5): pp. 1207-1215.

McKinney, F. K. & Jackson, J. B. C. (1989) *Bryozoan evolution,* Boston, Mass ; London, Unwin Hyman.

McKinney, F. K. & Taylor, P. D. (2001) Bryozoan generic extinctions and originations during the last one hundred million years. *Palaeontologia Electronica,* 4(1): pp. 26.

McPherson, M. J. & Møller, S. G. (2000) *PCR,* Oxford, Bios.

Medina, M., Collins, A. G., Silberman, J. D. & Sogin, M. L. (2001) Evaluating hypotheses of basal animal phylogeny using complete sequences of large and small subunit rRNA. *Proceedings of the National Academy of Sciences of the United States of America,* 98(17): pp. 9707-9712.

Miller, R. E., McDonald, J. A. & Manos, P. S. (2004) Systematics of Ipomoea subgenus quamoclit (Convolvulaceae) based on its sequence data and a Bayesian phylogenetic analysis. *American Journal of Botany,* 91(8): pp. 1208-1218.

Minsky, M. (1988) Memoir on Inventing the Confocal Scanning Microscope. *Scanning,* 10(4): pp. 128-138.

Morris, S. C. (2000) Evolution: Bringing molecules into the fold. *Cell,* 100(1): pp. 1-11.

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): pp. 428-441.

Munday, P. L. & Wilson, S. K. (1997) Comparative efficacy of clove oil and other chemicals in anaesthetization of Pomacentrus amboinensis, a coral reef fish. *Journal of Fish Biology,* 51(5): pp. 931-938.

Mundy, S. P., Taylor, P. D. & Thorpe, J. P. (1981) A reinterpretation of Phylactolaemate phylogeny. IN: Larwood, G. P. & Nielsen, C. (Eds.) *Recent and Fossil Bryozoa.* Fredensborg, Olsen & Olsen. pp. 185-190.

Murphy, D. B. & Rousseau, V. (1980) *Foundations of college chemistry,* New York ; Chichester, Wiley.

Neefs, J. M., Vandepeer, Y., Derijk, P., Chapelle, S. & Dewachter, R. (1993) Compilation of Small Ribosomal-Subunit RNA Structures. *Nucleic Acids Research,* 21(13): pp. 3025-3049.

Neefs, J. M., Vandepeer, Y., Derijk, P., Goris, A. & Dewachter, R. (1991) Compilation Of Small Ribosomal-Subunit RNA Sequences. *Nucleic Acids Research,* 19: pp. 1987-2015.

Nelles, L., Fang, B. D., Volckaert, G., Vandenberghe, A. & Dewachter, R. (1984) Nucleotide-Sequence of a Crustacean-18s Ribosomal-RNA Gene and Secondary Structure of Eukaryotic Small Subunit Ribosomal-Rnas. *Nucleic Acids Research,* 12(23): pp. 8749-8768.

Nelson-Smith, A. (1965) Marine Biology of Milford Haven: the physical environment. *Field Studies,* 2(2): pp. 155-174.

Nielsen, C. (1971) Entoproct life-cycles and the entoproct/ectoproct relationship. *Ophelia,* 9: pp. 209-341.

Nielsen, C. (2000) The phylogenetic position of Entoprocta and Ectoprocta. IN: Herrera-Cubilla, A. & Jackson, J. B. C. (Eds.) *Proceedings of the 11th International Bryozoology Association Conference.* Smithsonian Tropical Research Institute, Panamá.

Nielsen, C. (2001) Animal evolution : interrelationships of the living phyla, Oxford, Oxford University Press.

Nielsen, C. & Riisgard, H. U. (1998) Tentacle structure and filter-feeding in Crisia eburnea and other cyclostomatous bryozoans, with a review of upstream-collecting mechanisms. *Marine Ecology-Progress Series,* 168: pp. 163-186.

Nikon (2003) Introduction to Confocal Microscopy. Nikon.

Nikon (2003) Introduction to Differential Interference Contrast Microscopy. Nikon.

Nikulina, E. A., Hanel, R. & Schafer, P. (2007) Cryptic speciation and paraphyly in the cosmopolitan bryozoan Electra pilosa - Impact of the Tethys closing on species evolution. *Molecular Phylogenetics and Evolution,* 45(3): pp. 765-776.

Nitsche, H. (1870) Beitrage zur Kenntnis der Bryozoen. *Zeitschr. wissensch. Zool.,* 20.

Nixon, K. C. & Carpenter, J. M. (1993) On Outgroups. *Cladistics,* 9(4): pp. 413-426.

Nomura, M. (2001) Ribosomal RNA genes, RNA polymerases, nucleolar structures, and synthesis of rRNA in the yeast Saccharomyces cerevisiae. *Cold Spring Harbor Symposia On Quantitative Biology,* 66: pp. 555-565.

Notredame, C. (2002) Recent progress in multiple sequence alignment: a survey. *Pharmacogenomics,* 3(1): pp. 131-144.

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): pp. 205-217.

Notredame, C., Obrien, E. A. & Higgins, D. G. (1997) RAGA:RNA sequence alignment by genetic algorithm. *Nucleic Acids Research,* 25(22): pp. 4570-4580.

Nuin, P. A. S., Wang, Z. Z. & Elisabeth, R. M. (2006) The accuracy of several multiple sequence alignment programs for proteins. *Bmc Bioinformatics,* 7.

Nylander, J. A. A. (2004) MrModeltest 2ed., Evolutionary Biology Centre, Uppsala University.

Okano, K., Hunter, E. & Fusetani, N. (1996) Morphology and behavior of isolated living cells from Bugula neritina larvae (Bryozoa: Cheilostomata). *Journal Of Experimental Zoology,* 276(2): pp. 138-150.

Okuyama, M., Wada, H. & Ishii, T. (2006) Phylogenetic relationships of freshwater bryozoans (Ectoprocta, Phylactolaemata) inferred from mitochondrial ribosomal DNA sequences. *Zoologica Scripta,* 35(3): pp. 243-249.

Olsen, G. J., Matsuda, H., Hagstrom, R. & Overbeek, R. (1994) Fastdnaml - a Tool for Construction of Phylogenetic Trees of DNA-Sequences Using Maximum-Likelihood. *Computer Applications in the Biosciences,* 10(1): pp. 41-48.

Olsen, G. J. & Woese, C. R. (1993) Ribosomal RNA: a key to phylogeny. *FASEB J.,* 7(1): pp. 113-123.

Olson, L. E. & Yoder, A. D. (2002) Using secondary structure to identify ribosomal numts: Cautionary examples from the human genome. *Molecular Biology and Evolution,* 19(1): pp. 93-100.

Orti, G. & Meyer, A. (1997) The radiation of characiform fishes and the limits of resolution of mitochondrial ribosomal DNA sequences. *Systematic Biology,* 46(1): pp. 75-100.

Ostrovsky, A. N., Grischenko, A. V., Taylor, P. D., Bock, P. & Mawatari, S. F. (2006) Comparative anatomical study of internal brooding in three anascan bryozoans (Cheilostomata) and its taxonomic and evolutionary implications. *Journal of Morphology,* 267(6): pp. 739-749.

Ouvrard, D., Campbell, B. C., Bourgoin, T. & Chan, K. L. (2000) 18S rRNA secondary structure and phylogenetic position of Peloridiidae (Insecta, Hemiptera). *Molecular Phylogenetics and Evolution,* 16(3): pp. 403-417.

Packroff, G., Lawrence, J. R. & Neu, T. R. (2002) In situ confocal laser scanning microscopy of protozoans in cultures and complex biofilm communities. *Acta Protozoologica,* 41(3): pp. 245-253.

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): pp. 3839-3845.

Page, R. D. M. (2000) On the dangers of aligning RNA sequences using "Conserved" Motifs. *Technical Reports in Taxonomy series (University of Glasgow.).*

Page, R. D. M. & Holmes, E. C. (2002) *Molecular evolution: a phylogenetic approach*, Blackwell Science.

Parry-Smith, D. J., Payne, A. W. R., Michie, A. D. & Attwood, T. K. (1998) CINEMA - a novel colour interactive editor for multiple alignments (Reprinted from Gene, vol 221, pg GC57-GC63, 1998). *Gene,* 221(1): pp. GC57-GC63.

Passamaneck, Y. & Halanych, K. M. (2006) Lophotrochozoan phylogeny assessed with LSU and SSU data: Evidence of lophophorate polyphyly. *Molecular Phylogenetics and Evolution,* 40(1): pp. 20-28.

Passamaneck, Y. J. (2003) Mitochondrial genome of the bryozoan Membranipora membranacea and its phylogenetic implications. *Integrative and Comparative Biology,* 43(6): pp. 1074-1074.

Passamaneck, Y. J. & Halanych, K. M. (2002) Bryozoan affinities accessed using Hox genes and ribosomal RNA genes. *Integrative and Comparative Biology,* 42(6): pp. 1292-1292.

Passamaneck, Y. J. & Halanych, K. M. (2004) Evidence from Hox genes that bryozoans are lophotrochozoans. *Evolution & Development,* 6(4): pp. 275-281.

Passamaneck, Y. J., Schander, C. & Halanych, K. M. (2004) Investigation of molluscan phylogeny using large-subunit and small-subunit nuclear rRNA sequences. *Molecular Phylogenetics and Evolution,* 32(1): pp. 25-38.

Pawlik, J. R. (1992) Chemical Ecology Of The Settlement Of Benthic Marine-Invertebrates. *Oceanography And Marine Biology,* 30: pp. 273-335.

Peck, L. S., Hayward, P. J. & Spencerjones, M. E. (1995) A Pelagic Bryozoan from Antarctica. *Marine Biology,* 123(4): pp. 757-762.

Pei, J. M. & Grishin, N. V. (2006) MUMMALS: multiple sequence alignment improved by using hidden Markov models with local structural information. *Nucleic Acids Research,* 34(16): pp. 4364-4374.

Pei, J. M., Sadreyev, R. & Grishin, N. V. (2003) PCMA: fast and accurate multiple sequence alignment based on profile consistency. *Bioinformatics,* 19(3): pp. 427-428.

Peterson, K. J. & Addis, J. S. (2000) Clypeatula cooperensis gen. n., sp n., a new freshwater sponge (Porifera, Spongillidae) from the Rocky Mountains of Montana, USA. *Zoologica Scripta,* 29(3): pp. 265-274.

Peterson Steven, A. & Somit, A. (1992) The Dynamics of evolution : the punctuated equilibrium debate in the natural and social sciences, Ithaca ; London, Cornell University Press.

Petrucci, R. H. & Harwood, W. S. (1997) *General chemistry : principles and modern applications,* Upper Saddle River, N.J., Prentice Hall.

Philippe, H., Snell, E. A., Bapteste, E., Lopez, P., Holland, P. W. H. & Casane, D. (2004) Phylogenomics of eukaryotes: Impact of missing data on large alignments. *Molecular Biology And Evolution,* 21(9): pp. 1740-1752.

Phillips, A., Janies, D. & Wheeler, W. (2000) Multiple sequence alignment in phylogenetic analysis. *Molecular Phylogenetics and Evolution,* 16(3): pp. 317-330.

Phillips, C. L., Arend, L. J., Filson, A. J., Kojetin, D. J., Clendenon, J. L., Fang, S. & Dunn, K. W. (2001) Three-dimensional imaging of embryonic mouse kidney by two-photon microscopy. *American Journal of Pathology,* 158(1): pp. 49-55.

Pikaard, C. S. (1999) Nucleolar dominance and silencing of transcription. *Trends In Plant Science,* 4(12): pp. 478-483.

Pikaard, C. S. (2000) The epigenetics of nucleolar dominance. *Trends In Genetics,* 16(11): pp. 495-500.

Poe, S. (1998) The effect of taxonomic sampling on accuracy of phylogeny estimation: Test case of a known phylogeny. *Molecular Biology and Evolution,* 15(8): pp. 1086-1090.

Poe, S. (1998) Sensitivity of phylogeny estimation to taxonomic sampling. *Systematic Biology,* 47(1): pp. 18-31.

Porter, J. S. (2004) Morphological and genetic characteristics of erect subtidal species of Alcyonidium (Ctenostomata : Bryozoa). *Journal of the Marine Biological Association of the United Kingdom,* 84(1): pp. 243-252.

Porter, J. S. & Hayward, P. J. (2004) Species of Alcyonidium (Bryozoa : Ctenostomata) from Antarctica and Magellan Strait, defined by morphological, reproductive and molecular characters. *Journal of the Marine Biological Association of the United Kingdom,* 84(1): pp. 253-265.

Porter, J. S., Hayward, P. J. & Jones, M. E. S. (2001) The identity of Alcyonidium diaphanum (Bryozoa : Ctenostomatida). *Journal of the Marine Biological Association of the United Kingdom,* 81(6): pp. 1001-1008.

Porter, J. S., Ryland, J. S. & Carvalho, G. R. (2002) Micro- and macrogeographic genetic structure in bryozoans with different larval strategies. *Journal of Experimental Marine Biology and Ecology,* 272(2): pp. 119-130.

Porter, J. S., Skibinski, D. O. F., Leamon, J. & Hayward, P. J. (2001) Technique for analysis of bryozoan mitochondrial DNA. *Molecular Ecology Notes,* 1(1-2): pp. 103-105.

Porter, J. S. & Spencer-Jones, M. E. (2000) The use of fluorescent microscopy as a taxonomic tool in larval ctenostome Bryozoa. Swansea, Unpublished.

Porter, J. S. & University of Wales Swansea. School of Biological Sciences. (1999) *Reproduction and speciation in marine bryozoa of the genus Alcyonidium (lamouroux).*

Posada, D. (2006) ModelTest Server: a web-based tool for the statistical selection of models of nucleotide substitution online. *Nucleic Acids Research,* 34: pp. W700-W703.

Posada, D. & Buckley, T. R. (2004) Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology,* 53(5): pp. 793-808.

Posada, D. & Crandall, K. A. (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics,* 14(9): pp. 817-818.

Prestwich, S., Higgins, D. & O'Sullivan, O. (2003) A SAT-based approach to multiple sequence alignment. *Principles and Practice of Constraint Programming - Cp 2003, Proceedings.* Berlin, SPRINGER-VERLAG BERLIN. pp. 940-944.

Randle, C. P., Mort, M. E. & Crawford, D. J. (2005) Bayesian inference of phylogenetics revisited: developments and concerns. *Taxon,* 54(1): pp. 9-15.

Reed, C. G. (1977) Larval morphology and settlement of the bryozoan, Bowerbankia gracilis (Vesicularioidea, Ctenostomata): structure and eversion of the internal sac. IN: Chia, F.-S. & Rice, M. E. (Eds.) *Settlement and metamorphosis of marine invertebrate larvae: Proceedings of the symposium on settlement and metamorphosis of marine invertebrate larvae.* Toronto, Ontario, Canada, Elsevier, New York.

Reed, C. G. (1988) The Reproductive Biology of the Gymnolaemate Bryozoan Bowerbankia gracilis (Ctenostomata, Vesiculariidae). *Ophelia,* 29(1): pp. 1-23.

Reed, C. G. (1991) Bryozoa. IN: Giese, A. C., Pearse, J. S. & Pearse, V. B. (Eds.) *Reproduction of marine Invertebrates.* Pacific Grove, California, The Boxwood Press. pp. 85-245.

Reed, C. G. & Cloney, R. A. (1982) The Larval Morphology of the Marine Bryozoan Bowerbankia gracilis (Ctenostomata, Vesicularioidea). *Zoomorphology,* 100(1): pp. 23-54.

Reed, C. G., Ninos, J. M. & Woollacott, R. M. (1988) Bryozoan Larvae as Mosaics of Multifunctional Ciliary Fields - Ultrastructure of the Sensory Organs of Bugula stolonifera (Cheilostomata, Cellularioidea). *Journal of Morphology,* 197(2): pp. 127-145.

Reed, C. G. & Woollacott, R. M. (1982) Mechanisms Of Rapid Morphogenetic Movements In The Metamorphosis Of The Bryozoan Bugula neritina (Cheilostomata, Cellularioidea) .1. Attachment To The Substratum. *Journal Of Morphology,* 172(3): pp. 335-348.

Reed, C. G. & Woollacott, R. M. (1983) Mechanisms Of Rapid Morphogenetic Movements In The Metamorphosis Of The Bryozoan Bugula neritina (Cheilostomata, Cellularioidea) .2. The Role Of Dynamic Assemblages Of Microfilaments In The Pallial Epithelium. *Journal Of Morphology,* 177(2): pp. 127-143.

Reeder, J. & Giegerich, R. (2004) Design, implementation and evaluation of a practical pseudoknot folding algorithm based on thermodynamics. *BMC Bioinformatics,* 5(1): pp. 104.

Reeder, J., Hochsmann, M., Rehmsmeier, M., Voss, B. & Giegerich, R. (2006) Beyond Mfold: recent advances in RNA bioinformatics. *Journal of Biotechnology,* 124(1): pp. 41-55.

Rehm, B. H. A. (2001) Bioinformatic tools for DNA/protein sequence analysis, functional assignment of genes and protein classification. *Applied Microbiology and Biotechnology,* 57(5-6): pp. 579-592.

Rigoutsos, I., Huynh, T., Miranda, K., Tsirigos, A., McHardy, A. & Platt, D. (2006) Short blocks from the noncoding parts of the human genome have instances within nearly all known genes and relate to biological processes. *Proceedings of the National Academy of Sciences of the United States of America,* 103(17): pp. 6605-6610.

Rincon, P. A. (2000) Big fish, small fish: still the same species. Lack of morphometric evidence of the existence of two sturgeon species in the Guadalquivir river. *Marine Biology,* 136(4): pp. 715-723.

Rivera, M. C. & Lake, J. A. (2004) The ring of life provides evidence for a genome fusion origin of eukaryotes. *Nature,* 431(7005): pp. 152-155.

Rochow, T. G. & Tucker, P. A. (1994) Introduction to microscopy by means of light, electrons, X rays, or acoustics, New York, Plenum Press.

Rodriguez, F., Oliver, J. L., Marin, A. & Medina, J. R. (1990) The General Stochastic-Model of Nucleotide Substitution. *Journal of Theoretical Biology,* 142(4): pp. 485-501.

Ronquist, F. (2004) Bayesian inference of character evolution. *Trends in Ecology & Evolution,* 19(9): pp. 475-481.

Ronquist, F. & Huelsenbeck, J. P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics,* 19(12): pp. 1572-1574.

Ross, L. G. & B., R. (1999) Anaesthetic and Sedative techniques for Aquatic Animals, Blackwell Science.

Ryland, J. S. (1960) Light and Larval behaviour. *Journal of Experimental Biology,* 37: pp. 783-800.

Ryland, J. S. (1960) The British species of Bugula (Polyzoa). *Proceedings of the Zoological Society of London,* 134: pp. 65-105.

Ryland, J. S. (1970) *Bryozoans,* London, Hutchinson.

Ryland, J. S. (1976) Physiology and Ecology of Marine Bryozoans. *Advances in marine biology,* 14: pp. 285-443.

Ryland, J. S. (1977) Taxes and Tropisms of Bryozoans. IN: Woolacott, R. & Zimmer, R. L. (Eds.) *Biology of Bryozoans.* New York, Academic Press. pp. 411-436.

Ryland, J. S., Hayward, P. J. & Linnean Society of London. (1977) *British anascan bryozoans : Cheilostomata : Anasca ; keys and notes for the identification of the species,* London, Academic Press for the Linnean Society of London.

Ryland, J. S. & Porter, J. S. (2000) Alcyonidium reticulum sp nov., a common intertidal bryozoan from south-west Britain. *Journal of the Marine Biological Association of the United Kingdom,* 80(3): pp. 563-564.

Ryland, J. S. & Porter, J. S. (2003) The identity of Alcyonidium gelatinosum (Linnaeus, 1761) (Bryozoa : Ctenostomatida). *Journal of Natural History,* 37(18): pp. 2179-2189.

Ryland, J. S. & Porter, J. S. (2006) The identification, distribution andbiology of encrusting species of Alcyonidium (Bryozoa: Ctenostomatida) around the coasts of Ireland. *Biology and Environment: Proceedings of the Royal Irish Academy,* 106B(1): pp. 19-33.

Rzhetsky, A. & Nei, M. (1995) Tests of Applicability of Several Substitution Models for DNA-Sequence Data. *Molecular Biology and Evolution,* 12(1): pp. 131-151.

Sambrook, J. & Russell, D. W. (2001) *Molecular cloning : a laboratory manual,* Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press.

Sanderson, M. J. & Doyle, J. A. (2001) Sources of error and confidence intervals in estimating the age of angiosperms from rbcL and 18S rDNA data. *American Journal of Botany,* 88(8): pp. 1499-1516.

Sanderson, M. J. & Shaffer, H. B. (2002) Troubleshooting molecular phylogenetic analyses. *Annual Review of Ecology and Systematics,* 33: pp. 49-72.

Santagata, S. (2002) Structure and metamorphic remodeling of the larval nervous system and musculature of Phoronis pallida (Phoronida). *Evolution & Development,* 4(1): pp. 28-42.

Santagata, S. (2008) The morphology and evolutionary significance of the ciliary fields and musculature among marine bryozoan larvae. *Journal of Morphology,* 269(3): pp. 349-364.

Santagata, S. & Zimmer, R. L. (2000) Comparing cell patterns of coronate bryozoan larvae with fluorescent probes. IN: Herrera-Cubilla, A. & Jackson, J. B. C. (Eds.) *Proceedings of the 11th International Bryozoology Association Conference.* Smithsonian Tropical Research Institute, Panamá.

Santagata, S. & Zimmer, R. L. (2002) Comparison of the neuromuscular systems among actinotroch larvae: systematic and evolutionary implications. *Evolution & Development,* 4(1): pp. 43-54.

Sargis, E. J. (2002) A multivariate analysis of the postcranium of tree shrews (Scandentia, Tupaiidae) and its taxonomic implications. *Mammalia,* 66(4): pp. 579-598.

Savill, N. J., Hoyle, D. C. & Higgs, P. G. (2001) RNA sequence evolution with secondary structure constraints: Comparison of substitution rate models using maximum-likelihood methods. *Genetics,* 157(1): pp. 399-411.

Schoniger, M. & Von Haeseler, A. (1994) A Stochastic Model for the Evolution of Autocorrelated DNA Sequences. *Molecular Phylogenetics and Evolution,* 3(3): pp. 240-247.

Schwaninger, H. R. (1999) Population structure of the widely dispersing marine bryozoan Membranipora membranacea (Cheilostomata): implications for population history, biogeography, and taxonomy. *Marine Biology,* 135(3): pp. 411-423.

Sheffield, V. C., Cox, D. R., Lerman, L. S. & Myers, R. M. (1989) Attachment of a 40-Base-Pair G+C-Rich Sequence (Gc-Clamp) to Genomic DNA Fragments by the Polymerase Chain-Reaction Results in Improved Detection of Single-Base Changes. *Proceedings of the National Academy of Sciences of the United States of America,* 86(1): pp. 232-236.

Shimizu, K., Hunter, E. & Fusetani, N. (2000) Localisation of biogenic amines in larvae of Bugula neritina (Bryozoa : Cheilostomatida) and their effects on settlement. *Marine Biology,* 136(1): pp. 1-9.

Silén, L. (1972) Fertilization in bryozoa. *Ophelia,* 10: pp. 27-34.

Silva, E. P. & Russo, C. A. M. (2000) Techniques and statistical data analysis in molecular population genetics. *Hydrobiologia,* 420: pp. 119-135.

Simmons, M. P., Pickett, K. M. & Miya, M. (2004) How meaningful are Bayesian support values? *Molecular Biology and Evolution,* 21(1): pp. 188-199.

Skelton, P., Smith, A. & Monks, N. (2002) *Cladistics: a practical primer on CD-ROM,* Cambridge, Cambridge University Press.

Skibinski, D. O. F., Gallagher, C. & Beynon, C. M. (1994) Mitochondrial-DNA Inheritance. *Nature,* 368(6474): pp. 817-818.

Smaldon, G. & Lee, E. W. (1979) A synopsis of Methods for the narcotisation of marine invertebrates. *Natural History 6.* Edinburgh, Royal Scotish Museum.

Smith, A. B. (1994) Rooting Molecular Trees - Problems And Strategies. *Biological Journal Of The Linnean Society,* 51(3): pp. 279-292.

Smith, A. D., Lui, T. W. H. & Tillier, E. R. M. (2004) Empirical models for substitution in ribosomal RNA. *Molecular Biology and Evolution,* 21(3): pp. 419-427.

Sokal, R. R. & Rohlf, F. J. (1995) Biometry : the principles and practice of statistics in biological research, New York ; Oxford, Freeman.

Sorensen, M. V., Funch, P., Willerslev, E., Hansen, A. J. & Olesen, J. (2000) On the phylogeny of the Metazoa in the light of Cycliophora and Micrognathozoa. *Zoologischer Anzeiger,* 239(3-4): pp. 297-318.

Speakman, J. R. & Webb, P. I. (1993) Taxonomy, Status and Distribution of the Azorean Bat (Nyctalus- Azoreum). *Journal of Zoology,* 231: pp. 27-38.

Spencer, D. F., Schnare, M. N. & Gray, M. W. (1984) Pronounced structural similarities between the small subunit ribosomal RNA genes of wheat mitochondria and Escherichia coli. *Proc Natl Acad Sci U S A,* 81(2): pp. 493-497.

Spencer, M. & International Union for Pure and Applied Biophysics. (1982) *Fundamentals of light microscopy,* Cambridge ; New York, Cambridge University Press.

Spencer, M. A., Vick, M. C. & Dick, M. W. (2002) Revision of Aplanopsis, Pythiopsis, and 'subcentric' Achlya species (Saprolegniaceae) using 18S rDNA and morphological data. *Mycological Research,* 106: pp. 549-560.

Springer, M. S., DeBry, R. W., Douady, C., Amrine, H. M., Madsen, O., de Jong, W. W. & Stanhope, M. J. (2001) Mitochondrial versus nuclear gene sequences in deep-level mammalian phylogeny reconstruction. *Molecular Biology and Evolution,* 18(2): pp. 132-143.

Steffen, P., Voss, B., Rehmsmeier, M., Reeder, J. & Giegerich, R. (2006) RNAshapes: an integrated RNA analysis package based on abstract shapes. *Bioinformatics,* 22(4): pp. 500-503.

Stricker, S. A., Reed, C. G. & Zimmer, R. L. (1988) The Cyphonautes Larva of the Marine Bryozoan Membranipora membranacea .1. General Morphology, Body Wall, and Gut. *Canadian Journal of Zoology-Revue Canadienne De Zoologie,* 66(2): pp. 368-383.

Stricker, S. A., Reed, C. G. & Zimmer, R. L. (1988) The Cyphonautes Larva of the Marine Bryozoan Membranipora membranacea .2. Internal Sac, Musculature, and Pyriform Organ. *Canadian Journal of Zoology-Revue Canadienne De Zoologie,* 66(2): pp. 384-398.

Strom, R. (1977) Brooding Patterns of Bryozoans. IN: Woollacott, R. M. & Zimmer, R. L. (Eds.) *Biology of Bryozoans.* New York & London, Academic Press. pp. 23-56.

Sullivan, J. & Joyce, P. (2005) Model selection in phylogenetics. *Annual Review of Ecology Evolution and Systematics,* 36: pp. 445-466.

Sultan, F., Czubayko, U. & Thier, P. (2003) Morphological classification of the rat lateral cerebellar nuclear neurons by principal component analysis. *Journal of Comparative Neurology,* 455(2): pp. 139-155.

Susko, E., Inagaki, Y. & Roger, A. J. (2004) On inconsistency of the Neighbor-Joining, least squares, and minimum evolution estimation when substitution processes are incorrectly modeled. *Molecular Biology And Evolution,* 21(9): pp. 1629-1642.

Suter, S. J. (1994) Cladistic-Analysis of Cassiduloid Echinoids - Trying to See the Phylogeny for the Trees. *Biological Journal of the Linnean Society,* 53(1): pp. 31-72.

Sutherland, B., Stewart, D., Kenchington, E. R. & Zouros, E. (1998) The fate of paternal mitochondrial DNA in developing female mussels, Mytilus edulis: Implications for the mechanism of doubly uniparental inheritance of mitochondrial DNA. *Genetics,* 148(1): pp. 341-347.

Sutovsky, P., Tengowski, M. W., Navara, C. S., Zoran, S. S. & Schatten, G. (1997) Mitochondrial sheath movement and detachment in mammalian, but not nonmammalian, sperm induced by disulfide bond reduction. *Molecular Reproduction and Development,* 47(1): pp. 79-86.

Svennblad, B., Erixon, P., Oxelman, B. & Britton, T. (2006) Fundamental differences between the methods of maximum likelihood and maximum posterior probability in phylogenetics. *Systematic Biology,* 55(1): pp. 116-121.

Swofford, D. L. (2003) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). 4 ed. Sunderland, Massachusetts, Sinauer Associates.

Tabaska, J. E., Cary, R. B., Gabow, H. N. & Stormo, G. D. (1998) An RNA folding method capable of identifying pseudoknots and base triples. *Bioinformatics,* 14(8): pp. 691-699.

Tarkhnishvili, D. N., Thorpe, R. S. & Arntzen, J. W. (2000) Pre-Pleistocene refugia and differentiation between populations of the Caucasian salamander (Mertensiella caucasica). *Molecular Phylogenetics and Evolution,* 14(3): pp. 414-422.

Taylor, D. J. & Piel, W. H. (2004) An assessment of accuracy, error, and conflict with support values from genome-scale phylogenetic data. *Molecular Biology and Evolution,* 21(8): pp. 1534-1537.

Taylor, H. L. & Cooley, C. R. (1995) A Multivariate-Analysis of Morphological Variation among Parthenogenetic Teiid Lizards of the Cnemidophorus-Cozumela Complex. *Herpetologica,* 51(1): pp. 67-76.

Taylor, P. D. (1990) The impact of SEM in studies of living and fossil bryozoans. IN: Claugher, D. (Ed.) *Scanning electron microscopy in taxonomy and functional morphology.* Oxford, Clarendon Press. pp. 259-280.

Taylor, P. D. (1993) Bryozoa. IN: Benton, M. J. (Ed.) *The Fossil Record.* London, Chapman & Hall. pp. 87-103.

Taylor, P. D. (2000) Origin of the Modern Bryozoan Fauna. IN: Culver, S. J. & Rawson, P. F. (Eds.) *Biotic response to global change. The last 145 million years.* Cambridge, Natural History Museum/Cambridge University Press. pp. 195-209.

Taylor, P. D. (2000) Cyclostome Systematics: Phylogeny, Suborders and the Problem of Skeletal Organization. IN: Cubilla, A. H. & Jackson, J. B. C. (Eds.) *Proceedings of the 11th International Bryozoology Association Conference.* Smithsonian Tropical Research Institute, Balboa, Republic of Panama.

Taylor, P. D. & Ernst, A. (2004) Bryozoans. IN: Webby, B. D., Paris, F., Droser, M. L. & Percival, I. G. (Eds.) *The Great Ordovician Biodiversification Event.* New York, Columbia University Press. pp. 147-156.

Taylor, P. D. & Larwood, G. P. (1990) Major Evolutionary Radiations in the Bryozoa. IN: Taylor, P. D. & Larwood, G. P. (Eds.) *Major Evolutionary Radiations.* Oxford, Clarendon Press. pp. 209-233.

Telford, M. J., Wise, M. J. & Gowri-Shankar, V. (2005) Consideration of RNA secondary structure significantly improves likelihood-based estimates of phylogeny: Examples from the bilateria. *Molecular Biology And Evolution,* 22(4): pp. 1129-1136.

Thorpe, J. P. (1979) Enzyme variation and taxonomy: the estimation of sampling errors in measurements of interspecific genetic similarity. *Biological Journal of the Linnean Society,* 11(4): pp. 369-386.

Thorpe, J. P. & Beardmore, J. A. (1981) Genetic variation in natural populations of marine bryozoans. IN: Larwood, G. P. & Nielsen, C. (Eds.) *Recent and fossil Bryozoa.* Fredensborg, Olsen & Olsen. pp. 263-272.

Thorpe, J. P., Beardmore, J. A. & Ryland, J. S. (1978) Genetic evidence for cryptic specieation in the marine bryozoan Alcyonidium gelatinosum. *Marine Biology,* 49: pp. 27-32.

Thorpe, J. P. & Mundy, S. P. (1981) Biochemical genetics and taxonomy in Plumatella emarginata and P. repens (Bryozoa: Phylactolaemata). *Freshwater Biology,* 10(2): pp. 361-366.

Thorpe, R. S. (1988) Multiple Group Principal Component Analysis and Population Differentiation. *Journal of Zoology,* 216: pp. 37-40.

Thorpe, R. S. & Leamy, L. (1983) Morphometric Studies in Inbred and Hybrid House Mice (Mus Sp) - Multivariate-Analysis of Size and Shape. *Journal of Zoology,* 199(APR): pp. 421-432.

Tillier, E. R. M. & Collins, R. A. (1995) Neighbor Joining and Maximum-Likelihood with RNA Sequences - Addressing the Interdependence of Sites. *Molecular Biology and Evolution,* 12(1): pp. 7-15.

Tinoco, I. & Bustamante, C. (1999) How RNA folds. *Journal of Molecular Biology,* 293(2): pp. 271-281.

Tisdall, J. D. (2001) *Beginning Perl for bioinformatics,* Beijing ; Sebastopol, CA, O'Reilly.

Todd, J. A. (2000) The central role of ctenostomes in Bryozoan Phylogeny. IN: Herrera-Cubilla, A. & Jackson, J. B. C. (Eds.) *Proceedings of the 11th International Bryozoology Association Conference.* Smithsonian Tropical Research Institute, Panamá.

Todd, J. A. & Taylor, P. D. (1992) The First Fossil Entoproct. *Naturwissenschaften,* 79(7): pp. 311-314.

Tsyganov, A. (2003) Some aspects of biology and genetics of Bryozoa. *School of Biological Sciences.* Swansea, University of Wales.

Ueda-Nishimura, K. & Mikata, K. (2000) Two distinct 18S rRNA secondary structures in Dipodascus (Hemiascomycetes). *Microbiology-Uk,* 146: pp. 1045-1051.

Valentine, J. W., Jablonski, D. & Erwin, D. H. (1999) Fossils, molecules and embryos: new perspectives on the Cambrian explosion. *Development,* 126(5): pp. 851-859.

Valles, Y. & Boore, J. L. (2006) Lophotrochozoan mitochondrial genomes. *Integrative and Comparative Biology,* 46(4): pp. 544-557.

Van de Peer, Y. (1999) Molecular evolution and the incorporation of site-to-site rate variation in distance tree construction methods. *Belgian Journal of Zoology,* 129(1): pp. 5-15.

Van de Peer, Y., Baldauf, S. L., Doolittle, W. F. & Meyer, A. (2000) An updated and comprehensive rRNA phylogeny of (crown) eukaryotes based on rate-calibrated evolutionary distances. *Journal of Molecular Evolution,* 51(6): pp. 565-576.

Van de Peer, Y., De Rijk, P., Wuyts, J., Winkelmans, T. & De Wachter, R. (2000) The European Small Subunit Ribosomal RNA database. *Nucleic Acids Research,* 28(1): pp. 175-176.

Van de Peer, Y., Neefs, J. M., Derijk, P. & Dewachter, R. (1993) Reconstructing Evolution from Eukaryotic Small-Ribosomal-Subunit RNA Sequences - Calibration of the Molecular Clock. *Journal of Molecular Evolution,* 37(2): pp. 221-232.

Van de Peer, Y., Robbrecht, E., de Hoog, S., Caers, A., De Rijk, P. & De Wachter, R. (1999) Database on the structure of small subunit ribosomal RNA. *Nucleic Acids Research,* 27(1): pp. 179-183.

Van de Peer, Y., Vandenbroeck, I., Derijk, P. & Dewachter, R. (1994) Database on the Structure of Small Ribosomal-Subunit RNA. *Nucleic Acids Research,* 22(17): pp. 3488-3494.

Varberg, D. E., Purcell, E. J. & Rigdon, S. (2000) *Calculus,* Upper Saddle River, N.J., Prentice Hall.

Varón, A., L. S. , Vinh, I., Bomash, W. C. & Wheeler, Q. D. (2007) POY. 4.0 ed., American Musem of Natural History.

Voight, J. R. (1994) Morphological Variation in Shallow-Water Octopuses (Mollusca, Cephalopoda). *Journal of Zoology,* 232: pp. 491-504.

Voight, J. R. (2002) Morphometric analysis of male reproductive features of octopodids (Mollusca : Cephalopoda). *Biological Bulletin,* 202(2): pp. 148-155.

Voigt, E. (1991) Mono- or Polyphyletic Evolution of Chelostomatous Bryozoan Divisions? IN: Bigey, F. P. (Ed.) *Bryozaires Actuels et Fossiles: Bryozoa living and Fossil.* Nantes (France), Bulletin de la Société des Sciences Naturelles de l'Ouest de la France. pp. 505-522.

Waeschenbach, A. (2003) Molecular Evolutionary Processes in Ctenostome Bryozoans. *School of Biological Scinces.* Swansea, University of Wales.

Waeschenbach, A., Telford, M. J., Porter, J. S. & Littlewood, D. T. J. (2006) The complete mitochondrial genome of Flustrellidra hispida and the phylogenetic position of Bryozoa among the Metazoa. *Molecular Phylogenetics and Evolution,* 40(1): pp. 195-207.

Wagele, J. W. & Misof, B. (2001) On quality of evidence in phylogeny reconstruction: a reply to Zrzavy's defence of the 'Ecdysozoa' hypothesis. *Journal of Zoological Systematics and Evolutionary Research,* 39(3): pp. 165-176.

Warnke, K., Plötner, J., Santana, J. I., Rueda, M. J. & Llinas, O. (2002) Reflections on the phylogenetic position of Spirula (Cephalopoda): preliminary evidence from the 18S ribosomal RNA gene. IN: Warnke, K., Keupp, H. & von Boletzly, S. (Eds.) *Coleoid Cephalopods through time.* Berlin, Weinert GmbH.

Watts, P. C. & Thorpe, J. P. (2006) Influence of contrasting larval developmental types upon the population-genetic structure of cheilostome bryozoans. *Marine Biology,* 149(5): pp. 1093-1101.

Watts, P. C., Thorpe, J. P. & Taylor, P. D. (1998) Natural and anthropogenic dispersal mechanisms in the marine environment: a study using cheilostome Bryozoa. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences,* 353(1367): pp. 453-464.

Wenne, R. & Skibinski, D. O. F. (1995) Mitochondrial-DNA Heteroplasmy In European Populations Of The Mussel Mytilus-Trossulus. *Marine Biology,* 122(4): pp. 619-624.

Wheeler, Q. D., Gladstein, D. & De Laet, J. (2003) POY. 3.0.11 ed., American Museum of Natural History.

Wilgenbuch, J. C., Warren, D. L. & Swofford, D. L. (2004) AWTY: A system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference.

Williams, S. T. & Knowlton, N. (2001) Mitochondrial pseudogenes are pervasive and often insidious in the snapping shrimp genus Alpheus. *Molecular Biology and Evolution,* 18(8): pp. 1484-1493.

Wills, M. A. (1998) Crustacean disparity through the Phanerozoic: comparing morphological and stratigraphic data. *Biological Journal of the Linnean Society,* 65(4): pp. 455-500.

Wills, M. A. (2007) Fossil ghost ranges are most common in some of the oldest and some of the youngest strata. *Proceedings of the Royal Society B-Biological Sciences,* 274(1624): pp. 2421-2427.

Wilm, A., Mainz, I. & Steger, G. (2006) An enhanced RNA alignment benchmark for sequence alignment programs. *Algorithms for Molecular Biology,* 1(1): pp. 19.

Wilmotte, A., Vandepeer, Y., Goris, A., Chapelle, S., Debaere, R., Nelissen, B., Neefs, J. M., Hennebert, G. L. & Dewachter, R. (1993) Evolutionary Relationships

Among Higher Fungi Inferred From Small Ribosomal-Subunit RNA Sequence-Analysis. *Systematic And Applied Microbiology,* 16(3): pp. 436-444.

Winnepenninckx, B. & Backeljau, T. (1996) 18S rRNA alignments derived from different secondary structure models can produce alternative phylogenies. *Journal of Zoological Systematics and Evolutionary Research,* 34(3): pp. 135-143.

Winnepenninckx, B., Backeljau, T. & Dewachter, R. (1994) Small Ribosomal-Subunit RNA and the Phylogeny of Mollusca. *Nautilus,* 108: pp. 98-110.

Woese, C. R., Gutell, R., Gupta, R. & Noller, H. F. (1983) Detailed Analysis of the Higher-Order Structure of 16s-Like Ribosomal Ribonucleic-Acids. *Microbiological Reviews,* 47(4): pp. 621-&.

Woese, C. R. & Gutell, R. R. (1989) Evidence for Several Higher-Order Structural Elements in Ribosomal-RNA. *Proceedings of the National Academy of Sciences of the United States of America,* 86(9): pp. 3119-3122.

Woollacott, R. M. & Zimmer, R. L. (1971) Attachment and metamorphosis of the Cheilo-ctenostome bryozoan Bugula neritina (Linne). *Journal of Morphology,* 134: pp. 351-382.

Wright, S. J., Schatten, H. & Simerly, C. (1990) Three-dimentional fluorescence imaging with the tandem scanning confocal microscope. IN: Herman, B. & Jacobson, K. (Eds.) *Optical microscopy for biology : proceedings of the International Conference on Video Microscopy held in Chapel Hill, North Carolina, June 4-7, 1989.* New York, Wiley-Liss. pp. 29-43.

Wuyts, J., De Rijk, P., Van de Peer, Y., Pison, G., Rousseeuw, P. & De Wachter, R. (2000) Comparative analysis of more than 3000 sequences reveals the existence of two pseudoknots in area V4 of eukaryotic small subunit ribosomal RNA. *Nucleic Acids Research,* 28(23): pp. 4698-4708.

Xia, F. S., Zhang, S. G. & Wang, Z. Z. (2007) The oldest bryozoans: New evidence from the late Tremadocian (early Ordovician) of east Yangtze Gorges in China. *Journal of Paleontology,* 81(6): pp. 1308-1326.

Xia, X. H., Xie, Z. & Kjer, K. M. (2003) 18S ribosomal RNA and tetrapod phylogeny. *Systematic Biology,* 52(3): pp. 283-295.

Xuejun, H., Zhichao, Z., Yongming, B., Qing, Y. & Lija, A. (2002) An expeditious method for constructing t-Vectors using Eam 1105 I Cassettes. *Plant Molecular Biology Reporter,* 20(June): pp. 189a-189e.

Yang, Z. & Rannala, B. (2005) Branch-length prior influences Bayesian posterior probability of phylogeny. *Systematic Biology,* 54(3): pp. 455-470.

Yang, Z. H. (1996) Among-site rate variation and its impact on phylogenetic analyses. *Trends In Ecology & Evolution,* 11(9): pp. 367-372.

Ying, X. M., Luo, H., Luo, J. C. & Li, W. J. (2004) RDfolder: a web server for prediction of RNA secondary structure. *Nucleic Acids Research,* 32: pp. W150-W153.

Zardoya, R. & Meyer, A. (1996) Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates. *Molecular Biology And Evolution,* 13(7): pp. 933-942.

Zeeman, C. (1992) Catastrophe theory applied to Darwinian evolution. *The Linnean,* 21(3): pp. 22-34.

Zhang, D. X. & Hewitt, G. M. (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology,* 12(3): pp. 563-584.

Zimmer, R. L. & Woollacott, R. M. (1977) Structure and classification of Gymnolaemate larvae. IN: Woolacott, R. & Zimmer, R. L. (Eds.) *Biology of Bryozoans.* New York, Academic Press. pp. 57-89.

Zimmer, R. L. & Woollacott, R. M. (1989) Intercoronal Cell Complex of Larvae of the Bryozoan Watersipora arcuata (Cheilostomata, Ascophora). *Journal of Morphology,* 199(2): pp. 151-164.

Zimmer, R. L. & Woollacott, R. M. (1989) Larval Morphology of the Bryozoan Watersipora arcuata (Cheilostomata, Ascophora). *Journal of Morphology,* 199(2): pp. 125-150.

Zimmer, R. L. & Woollacott, R. M. (1993) Anatomy of the Larva of Amathia vidovici (Bryozoa, Ctenostomata) and Phylogenetic Significance of the Vesicullariform Larva. *Journal of Morphology,* 215(1): pp. 1-29.

Zrzavy, J., Mihulka, S., Kepka, P., Bezdek, A. & Tietz, D. (1998) Phylogeny of the Metazoa based on morphological and 18S ribosomal DNA evidence. *Cladistics-the International Journal of the Willi Hennig Society,* 14(3): pp. 249-285.

Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research,* 31(13): pp. 3406-3415.

Zuker, M. & Stiegler, P. (1981) Optimal Computer Folding of Large RNA Sequences Using Thermodynamics and Auxiliary Information. *Nucleic Acids Research,* 9(1): pp. 133-148.

Клюге, Г. А. (1962) *Мшанки северных морей СССР,* Москва-Ленинград, АН СССР, Ленингр. отд-ние.