Phylogeny and Evolution of the Heterobranchia (Mollusca, Gastropoda)

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Biowissenschaften (FB15)
der Johann Wolfgang Goethe – Universität
in Frankfurt am Main

von
Angela Dinapoli
aus Mannheim

Frankfurt am Main 2009
(D30)

Vom Fachbereich Biowissenschaften der
Johann Wolfgang Goethe – Universität als Dissertation angenommen

Dekan: Prof. Dr. Volker Müller
Gutachter: Prof. Dr. Annette Klussmann-Kolb und Prof. Dr. Bruno Streit
Datum der Disputation: 13.07.09

Die Schnecke kann dir mehr
über den Weg erzählen
als der Hase.

Bernd Stromberg

Table of contents

List of figures

List of tables

List of appendices

List of abbreviations (in alphabetical order)

AIC	Akaike information criterion
AU	Approximately Unbiased
BLAST	Basic Local Alignment Search Tool
bp	Base pair
BSA	Bovine serum albumin
cDNA	complementary DNA
CI	Confidence interval
CNS	Central nervous system
COI	Cytochrome c oxidase I
DAMBE	Data analysis in molecular biology and evolution
dH_2O	Distillated water
DMSO	Dimethyl sulfoxide
DNA	Desoxyribonucleic acid
dNTP	Desoxyribonucleotide triphosphate
ESTs	Expressed sequence tags
fig.	Figure
GTR	General Time Reversible
HKY	Hasegawa-Kishino-Yano
HPD	Highest posterior density
ILD	Incongruence length difference
LSU	Large subunit
Ma	Million years
MC	Monte Carlo
MCMC	Markov chain Monte Carlo
MDFs	Mantle dermal formations
$MgCl_2$	Magnesium chloride
ML	Maximum likelihood
MP	Maximum parsimony
Muscle	Multiple sequence comparison by log-expectation
PAUP	Phylogenetic Analysis Using Parsimony
PCR	Polymerase chain reaction
Phase	Package for Phylogenetics and Sequence Evolution
RAxML	Randomized Axelerated Maximum Likelihood
rDNA	Ribosomal desoxyribonucleic acid
rRNA	Ribosomal ribonucleic acid
SAMS	Splits analysis methods
SRD	Scientific Research and Development
SSU	Small subunit
tab.	Table
TBE	Tris/Borate/EDTA
TMAC	Tetra-methyl-ammonium chloride

Zusammenfassung

Ziel dieser Dissertation war es, die Phylogenie und Evolution der Heterobranchia (Mollusca, Gastropoda) auf der Basis von Nukleotidsequenzen zu klären. Ein Hauptfokus lag dabei auf den basalen Heterobranchia, die in vorangegangenen molekularen Studien meist unberücksichtigt blieben. Das Konzept der Heterobranchia basiert auf morphologischen Studien von Haszprunar (1985a und 1988) und umfasst die paraphyletische Gruppe der basalen Heterobranchia sowie die monophyletische Gruppe der Euthyneura, zu der die Opisthobranchia und Pulmonata gehören. Eine Bestätigung dieses Konzeptes anhand molekular-systematischer Analysen blieb bislang aus.

Zusätzlich wurden im Rahmen dieser Dissertation unterschiedliche (meist neu entwickelte) Softwareprogramme auf ihre Anwendbarkeit bzw. Nutzen getestet, um Fragen, die zum einen Verwandtschaftsverhältnisse und zum anderen Evolutionsereignisse der Heterobranchia betreffen, besser beantworten zu können.

Zur Klärung der Monophylie bzw. der Verwandtschaftsverhältnisse innerhalb der Heterobranchia wurden molekulare Analysen sowohl mit einem Bayesianischen als auch einem Likelihood Ansatz durchgeführt. Die dafür verwendeten Daten wurden in intensiven Voranalysen auf ihre Qualität (phylogenetisches Signal) überprüft, um die geeignetsten Daten *a priori* zu identifizieren.

Ausgangssituation für die Voranalyse waren drei verschiedene Datensätze (Datensatz 0, I und II), bestehend aus Sequenzen der nukleären 18S rDNA und 28S rDNA sowie Sequenzen der mitochondrialen 16S rDNA und Cytochrom Oxidase I (COI). Mit Hilfe der Software Muscle wurden Alignments der einzelnen genetischen Marker für alle 3 Datensätze erstellt. Die Alignments von Datensatz 0 blieben im Anschluss unmodifiziert, d. h. es wurden keine Basenpositionen herausgenommen, wodurch die Alignments ihre Originallänge beibehielten. Datensatz I und II hatten die gleichen Ausgangsdaten wie Datensatz 0, allerdings wurden in Datensatz I und II diverse Alignmentbereiche nach zwei verschiedenen Konzepten *a priori* eliminiert. Der Ansatz bei Datensatz I war hierbei eine visuelle Durchsicht der einzelnen Alignments nach langen Inserts und hypervariablen Bereichen. Beides kann die Phylogenierekonstruktion negativ beeinflussen und wurde deshalb vorab aus den Alignments von Datensatz I entfernt. Die Entscheidung, welche Bereiche in

Datensatz II vor der Phylogenierekonstruktion eliminiert werden, wurde mittels der Software Aliscore eruiert. Es handelt sich dabei um ein neu entwickeltes Programm von Misof & Misof (in press), das verrauschte Nukleotidpositionen im Alignment erkennen und entfernen kann.
Mit diesen drei unterschiedlichen Datensätzen wurden verschiedene statistische Tests (wie Chi-Quadrat-Test oder Relative-Rate-Test) sowie Sättigungsanalysen durchgeführt. Zusätzlich wurden intensive Netzwerkanalysen durchgeführt, zum einen mit der Software SplitsTree und zum anderen mit der Software SAMS. Dies diente vor allem dazu, herauszufinden, welcher dieser drei Datensätze das beste phylogenetische Signal für die Phylogenie-Rekonstruktion der Heterobranchia enthält.
Nach Auswertung der einzelnen Tests zeigte sich, dass Datensatz I am besten geeignet schien, die Phylogenie des Taxons Heterobranchia zu rekonstruieren. Allerdings musste festgestellt werden, dass der festgelegte Datensatz aufgrund eines hohen Sättigungsgrades (der bei Großgruppenphylogenien selten ausbleibt) kritisch zu betrachten ist. Zusätzlich zeigte der Datensatz für bestimmte Gruppierungen ein konfliktreiches phylogenetisches Signal. Um Unsicherheiten, die z.B. auf eine hohe Ratenheterogenität oder abweichende Basenkompositionen zurückzuführen sind, auszugleichen, wurden für die Baumrekonstruktion Analysemethoden verwendet, die evolutionäre Modelle der Nukleotidsubstitutionen mit berücksichtigen.

Die sich anschließende Phylogenierekonstruktion stützt die Monophylie der Heterobranchia. Einige traditionelle, auf Basis morphologischer Untersuchungen beschriebene Taxa, konnten nicht bestätigt werden, z. B. gruppieren die Pyramidellidae und Glacidorboidea nicht an der Basis der Heterobranchia.
Die „basalen Heterobranchia“ sind paraphyletisch. Aufgrund einer unaufgelösten Baumtopologie an der Basis der Heterobranchia kann keine Aussage darüber getroffen werden, welches basale Taxon als erstes im Laufe der Erdgeschichte aufgetreten ist.
Die Murchisonellidae stehen in keinem Schwestergruppenverhältnis zu den Pyramidellidae, was bedeutet, dass die Pyramidelloidea polyphyletisch sind.
Die bereits im Vorfeld angenommene Heterobranchia-Verwandtschaft der Gattungen *Graphis* und *Larochella* konnte durch den Einschluss der beiden Taxa in die Heterobranchia bestätigt werden.
Valvata und *Cornirostra* clustern zusammen als Valvatoidea und bilden die Schwestergruppe zu einer Klade bestehend aus Architectonicoidea und Omalogyroidea. Die Orbitestellidae (deren Zugehörigkeit zu den Valvatoidea in früheren Studien diskutiert wurde) sowie die

Cimidae stehen in keinem Schwestergruppenverhältnis sondern bilden in der Topologie einzelne evolutionäre Linien.
Ein unerwartetes Schwestergruppenverhältnis, welches die Phylogeniehypothese wiederspiegelt, besteht zwischen den Rissoelloidea und den Acteonoidea.
Die Euthyneura sind aufgrund der abgeleiteten Stellung der Pyramidellidae und Glacidorboidea innerhalb der Euthyneura in dieser Studie paraphyletisch. Die Pulmonata sind ebenfalls paraphyletisch wohingegen die Opisthobranchia polyphyletischen Ursprungs sind.
Innerhalb der Euthyneura bzw. Opisthobranchia zweigen die Nudibranchia als erstes Taxon ab und stehen dabei im Schwestergruppenverhältnis zu den restlichen Euthyneura, wohingegen die ebenfalls zu den Opisthobranchia gehörenden Umbraculoidea, Cephalaspidea, Akeroidea und Pteropoda als gut gestützte Clade im Baum erscheinen. Über die Verwandtschaftsverhältnisse der Sacoglossa (Opisthobranchia) und Siphonarioidea (Pulmonata) läßt sich aufgrund einer unaufgelösten Baumtopologie wenig sagen.

Eine weitere Klade im Baum umfasst die zu den Pulmonaten gehörenden Taxa Hygrophila und Amphiboloidea, die basalen Gruppen Glacidorboidea und Pyramidellidae und die monophyletischen Eupulmonata (Stylommatophora, Onchidioidea, Ellobioidea und Otinoidea). Innerhalb der Eupulmonata zeigen die Stylommatophora ein Schwestergruppenverhältnis mit den restlichen Eupulmonaten. Die Onchidioidea sind die Schwestergruppe der Ellobioidea und Otinoidea wobei die Ellobioidea die Schwerstergruppe der Otinoidea sind. Eine Monophylie der Basommatophora (Siphonarioidea, Hygrophila und Amphiboloidea) konnte nicht bestätigt werden.

Die Ergebnisse der Phylogenierekonstruktion wurden im Anschluss an die Analyse genutzt, um verschiedene evolutionäre Szenarien zu entwickeln bzw. zu diskutieren. Es konnte dabei festgestellt werden, dass die basalen Gruppen, im Hinblick auf die Diversität auf Gattungs- und Artebene, weit weniger Taxa hervorgebracht haben als die Euthyneura, die allgemein als Königsgruppe der Gastropoda bezeichnet werden. Dies könnte verschiedene Gründe haben. Zum einen scheint die Nahrungsspezialisierung vor allem innerhalb der Opisthobranchia zu einer explosionsartigen adaptiven Radiation einzelner Opisthobranchia-Gruppen geführt zu haben. Zum anderen war die erfolgreiche Besiedlung nicht-mariner Habitate innerhalb der Pulmonata ebenfalls ausschlaggebend für eine enorme Diversifikation. Solche Großereignisse fanden innerhalb der basalen Gruppen, wenn überhaupt, nur mit mäßigem Erfolg statt.

Des Weiteren wurden durch den Einschluss neuer limnischer Arten, wie *Valvata* oder *Glacidorbis*, in die Phylogenierekonstruktion, neue Erkenntnisse über die Besiedlung des Süßwassers gewonnen. Eine Kolonialisierung des Süßwassers erfolgte innerhalb der Heterobranchia mehrmals unabhängig voneinander. Innerhalb der Pulmonaten erfolgte die Besiedlung mindestens zweimal, einmal durch die Hygrophila und ein anderes Mal durch die Glacidorboidea, deren Pulmonaten-Zugehörigkeit durch die phylogenetischen Analysen bestätigt wurde.

Aufgrund von unzureichenden Erkenntnissen über die Funktionen bestimmter neuronaler Strukturen im Nervensystem der basalen Gruppen bzw. der Euthyneura, kann keine Aussage darüber getroffen werden, ob neuronale Unterschiede für den unterschiedlichen evolutionären Erfolg verantwortlich sind.

Um erste Einblicke in die Evolution der Heterobranchia zu bekommen, wurde eine Fallstudie durchgeführt. Hierfür wurden in einer intensiven Literaturrecherche fossile Daten gesammelt, mit denen im Anschluss eine molekulare Uhr geeicht wurde, die wiederum helfen sollte, bestimmte Aufspaltungsereignisse im phylogenetischen Baum zeitlich einzuordnen. Als Werkzeug diente das Programm Beast, das eine so genannte *"relaxed" molecular clock* implementiert hat. Durch dieses neue Verfahren können Evolutionsraten verschiedener Organismengruppen innerhalb einer Analyse variieren. Um mögliche Korrelationsmuster zwischen einem Anstieg von Diversifikations- und Massenaussterbeereignissen zu finden, wurde zusätzlich ein „Lineage-through-time plot" mit den gewonnenen Daten erstellt.
Aufgrund von großen 95% Konfidenzintervallen an den Knoten der mit Beast rekonstruierten Baumtopologie, ist die zeitliche Einordnug bestimmter Aufspaltungsereignisse nur ungefähr möglich. Dieser Versuchsansatz soll deshalb als Arbeitshypothese verstanden werden, um erste Einblicke in den Ursprung und das Alter des Taxons Heterobranchia und seine Untergruppen zu geben.

Da einige Ergebnisse der in dieser Arbeit aufgestellten molekularen Phylogeniehypothese mit morphologischen Erkenntnissen nicht übereinstimmen, wurden nachträglich verschiedene Methoden angewandt, um die Plausibilität dieser Hypothesen zu überprüfen. Der durchgeführte AU-Test, mit dem die Wahrscheinlichkeit von anderen, erzwungenen Baumtopologien getestet werden kann, lieferte keine eindeutigen Ergebnisse. Zwar zeigte die nicht erzwungene Hypothese dieser Arbeit die besten Likelihood-Werte, es konnten jedoch

andere Hypothesen (wie monophyletische Euthyneura, Opisthobranchia und Pulmonata) aufgrund einer nicht signifikanten statistischen Unterstützung nicht ausgeschlossen werden.

Des Weiteren wurden Sekundärstrukturrekonstruktionen der 18S rRNA und 28S rRNA durchgeführt. Zum einen sollten auf diese Weise weitere Erkenntnisse bezüglich der Verwandtschaftsverhältnisse innerhalb der Heterobranchia gewonnen werden und zum anderen sollte dies helfen evolutionäre Modelle, die zur Baumrekonstruktion eingesetzt werden, weiter zu verbessern.

Innerhalb der rekonstruierten Sekundärstrukturen konnten tatsächlich synapomorphe Strukturen gefunden werden, die verschiedene Gruppen innerhalb der Heterobranchia stützen. Außerdem zeigte diese Studie auch spezifische Strukturen, die vor allem die Vetigastropoda von den restlichen Gruppen trennt. Daraus lässt sich schließen, dass 18S rRNA und 28S rRNA Sekundärstrukturen potentiell geeignet sind, um Verwandtschaftsverhältnisse innerhalb höherer taxonomischer Einheiten wie Gastropoda oder Mollusca aufzuklären.
Leider konnte keine Verbesserung des phylogenetischen Signals durch den Einsatz von spezifischen rDNA Evolutionsmodellen (wie sie in dem Programm Phase implementiert sind) sowie der Berücksichtigung von gepaarten und ungepaarten Basenpaaren in der Phylogenierekonstruktion beobachtet werden. Dies lag möglicherweise daran, dass aufgrund von fehlenden Übereinstimmungen im Taxonsampling nur Einzelanalysen der 18S und 28S rDNA Sequenzen und keine Kombinationsanalysen durchgeführt werden konnten und in den einzelnen Marker nicht genügend phylogenetisches Signal vorhanden war.
Es konnte jedoch gezeigt werden, dass es sich bei der neu entwickelten Software RNAsalsa um ein geeignetes Werkzeug handelt, schnell und zuverlässig Sekundärstrukturen der 18S rRNA und 28S rRNA zu rekonstruieren.

Zusammenfassend ist zu sagen, dass die Ergebnisse dieser Arbeit zahlreiche neue Einblicke bzw. Erkenntnisse über die Phylogenie und Evolution der Heterobranchia liefern und als Basis für weiterführende Analysen verwendet werden können.
Außerdem sollen die Erfahrungen, die aus zum Teil neu entwickelten und hier getesteten Programmen gewonnen werden konnten, anderen Wissenschaftlern helfen, eigene Fragestellungen besser beantworten zu können.

Abstract

Many questions regarding gastropod phylogeny have not yet been answered like the molecular confirmation of the Heterobranchia concept based on morphological studies from Haszprunar (1985a; 1988). This taxon contains the "Lower Heterobranchia" (with several "primitive" or "basal" members) and the Euthyneura (with the Opisthobranchia and Pulmonata).

Phylogenetic relationships of subgroups within the Heterobranchia have not been satisfactorily resolved and monophyly of some taxa within the Heterobranchia (e.g. Opisthobranchia) is questionable. Moreover, most of the "Lower Heterobranchia" have not been included in former molecular studies.

In order to resolve phylogenetic relationships within the Heterobranchia, I pursued a molecular systematic approach by sequencing and analysing a variety of genetic markers (including nuclear 28S rDNA + 18S rDNA and mitochondrial 16S rDNA + COI sequences). Maximum likelihood as well as Bayesian inference methods were used for phylogenetic reconstruction.

The data were investigated *a priori* to tree reconstruction in order to find the most appropriate dataset for reconstructing heterobranch phylogeny. A variety of statistical tests (like Chi-Square-Test or Relative-Rate-Test) were applied and the substitution saturation was measured. The Relative-Rate-Test revealed the highest evolution rates within the "Lower Heterobranchia" (*Omalogyra* sp., *Omalogyra fusca*, *Murchisonella* sp., *Ebala* sp. and *Architectonica perspectiva*) and Opisthobranchia (*Hyalocylis striata*). Furthermore, many of the nucleotide positions show a high degree of substitution saturation. Additionally, bipartitions (splits) in the alignment were examined and visualized by split network analyses to estimate data quality. A high level of conflict indicated by many parallel edges of the same lengths could be observed in the neighbournet graphs. Moreover, several taxa with long terminal branches could be identified in all three datasets belonging to the Vetigastropoda, Caenogastropoda, "Lower Heterobranchia" or Opisthobranchia (Nudipleura).

All phylogenetic analyses revealed a monophyletic Heterobranchia. Within the Heterobranchia several well supported clades could be resolved. However, the traditional

classification based on morphological data could not be confirmed due to paraphyletic Euthyneura (because of the inclusion of the Pyramidellidae and Glacidorboidea) as well as paraphyletic Pulmonata and polyphyletic Opisthobranchia.
Based on the phylogenetic inferred evolutionary trends regarding habitat colonisation or character complexes could be deduced.

A case study was conducted in order to estimate divergence ages using a "relaxed" molecular clock approach with fossils as minimum age constraints. However, due to large 95% confidence intervals a precise dating of the nodes was not possible. Hence, the results are considered as preliminary.

To test the plausibility of the newly obtained hypotheses, the results were evaluated *a posteriori* using a hypothesis test and secondary structures of the complete 18S rRNA and 28S rRNA. Secondary structure motifs were found within domain 43 and E23 2 &5 of the 18S rRNA as well as within domain E11 and G5_1 of the 28S rRNA, which contain phylogenetic signals to support various groups within the Heterobranchia. In addition, taxon specific motifs were found separating the Vetigastropoda from the Caenogastropoda and Heterobranchia, indicating a possible application of the secondary structure of 18S rRNA and 28S rRNA to reveal phylogenetic relationships at higher taxonomic levels such as Gastropoda or even Mollusca.

The utility of the newly invented software RNAsalsa for the reconstruction of secondary structures was tested. The obtained structures were used to adjust evolutionary models specific to rRNA stem (paired basepairs) and loop (unpaired basepairs) regions with the intention of improving phylogenetic results. This approach proved unsuccessful.

This molecular phylogenetic investigation provides the most comprehensive molecular study of Heterobranchia relationships to date. Substantial insights into the evolution and phylogeny of this enigmatic taxon have been gained.

1. General introduction

The phylum Mollusca is extremely diverse, enabling a great variety of functional body plans to evolve. The Gastropoda comprise the largest class of the eight living classes representing about 80% of the extant Mollusca (Haszprunar et al. 2008). They are defined by the following apomorphic characters in relation to their sister taxa: torsion, larval operculum and the shape of the larval shell (Ponder & Lindberg 1997). Many gastropod taxa have become important model organisms in various biological fields like ecology, evolutionary biology or neurobiology.

The current classification of the Gastropoda is a consensus of phylogenetic hypotheses proposed by several authors during the last two decades e.g. Haszprunar (1985a; 1988), Bieler (1992), Salvini-Plaven & Steiner (1996), Ponder & Lindberg (1997), Colgan et al. (2000; 2003; 2006), Dayrat et al. (2001), Dayrat & Tillier (2002), Grande et al. (2004a; 2008), Klussmann-Kolb et al. (2008). Currently, the Gastropoda are divided into six major groups: Patellogastropoda, Neritopsina, Cocculiniformia, Vetigastropoda, Caenogastropoda and Heterobranchia (Ponder & Lindberg 1997, Grande et al. 2008).

Gastropoda have a rich fossil record dating back to the Cambrian (Fryda et al. 2008). The oldest known Heterobranchia occurred in the Middle Paleozoic (Bandel 1994, Bandel & Heidelberger 2002, Fryda et al. 2008) but are more abundant in the Late Paleozoic (Bandel 2002, Fryda et al. 2008). Up to date no Opisthobranchia or Pulmonata are known from the Paleozoic (Fryda et al. 2008). The oldest Opisthobranchia appeared in the Triassic and the Pulmonata in the Jurassic as proposed by Bandel (1994; 2002).

The most heterogeneous Gastropoda are the Heterobranchia which were classified by Haszprunar in 1985 and 1988. They comprise the paraphyletic “Lower Heterobranchia” and the Euthyneura (including Opisthobranchia and Pulmonata). The monophyly of the Heterobranchia is well supported based on morphological characters like a sinistral larval shell produced by a planktotrophic veliger, a distinctive sperm ultrastructure, a medial position of the eyes in many taxa, a lack of a true ctenidium, a simple oesophagus and a pigmented mantel organ (which is reduced in more derived taxa) (Haszprunar 1985a, Ponder & Lindberg 1997).

The molecular confirmation of the Heterobranchia concept including representatives of most of the major groups is lacking to date. The inclusion of lower heterobranch taxa (e.g. Architectonicoidea, Glacidorboidea, Omalogyroidea, Pyramidelloidea, Rissoelloidea, Valvatoidea) has been particularly neglected in most of the former molecular studies. Moreover, phylogenetic relationships of subgroups of Heterobranchia have not been resolved satisfactorily and monophyly of some taxa within Heterobranchia is questionable.

A long evolutionary history, often rapid radiations, and the adaptation to many habitats by members of the same evolutionary line as well as to the same habitat by distantly related forms, results in a multitude of convergences. These convergences render the reconstruction of gastropod phylogeny difficult (Bieler 1992).

There is a high degree of homoplasy in many morphological gastropod characters leading to difficulties in obtaining significant results from phylogenetic analyses based on morphology. The reduction and loss of plesiomorphic structures, rather than their structural modification is responsible for much of the homoplasy in gastropods (Ponder & Lindberg 1997). Moreover, parallel trends, such as the evolution of various body forms (e.g. limpets, slugs), habits or dietary specialisations and the resulting homoplasy are major problems of the phylogenetic reconstruction (Ponder & Lindberg 1997). This is particularly true for the Opisthobranchia (Gosliner 1985; 1991, Gosliner & Ghiselin 1984, Ponder & Lindberg 1997, Dayrat & Tillier 2002) and partly for the Pulmonata (Tillier 1989, Ponder & Lindberg 1997, Dayrat & Tillier 2002).

Phylogenetic inferences based on molecular data are known to also have problems with homoplasy. Substitution saturation caused by multiple-hits is responsible for homoplastic changes (Grande et al. 2004a). High rates of homoplasy cause a loss of phylogenetic signal. Moreover, convergent evolutionary changes could be misinterpreted to support nonexisting relationships (Boore and Brown 1998).

Therefore, when working with molecular data one must answer different questions (as already proposed by Wägele & Mayer 2007) before conducting phylogenetic analyses like “How informative is the data set?”, “Is it possible to discern signal and noise?”, “How likely are specific alternative tree topologies?” or “Is the substitution model adequate?” to enable that the best possible results and the most plausible hypotheses, respectively, are obtained.

Structure and aim of the present study

The aim of this comprehensive study is the evaluation of the Heterobranchia concept based on morphological studies from Haszprunar (1985a, 1988) with molecular methods. The main focus lies on the "Lower Heterobranchia", which were neglected in former molecular studies. Moreover, the implementation of novel methodological approaches will be tested, which include the detection of ambiguously aligned positions in sequence alignments, reconstruction of rRNA[1] secondary structures and the application of specific rDNA[2] substitution models.

This thesis is divided into seven chapters. The following chapter 2 deals with the *a priori* evaluation of data quality in order to determine whether the data are suitable for phylogenetic reconstruction in the case of the Heterobranchia. As aforementioned, molecular data of Gastropoda could show a high degree of homoplasy. Therefore, it is important to improve the information value of molecular data using tools which are independent from tree reconstruction. In this light, the first aim of chapter 2 is to identify ambiguous nucleotide sites in the alignment using the newly developed software Aliscore. The second aim is to verify the most appropriate data to infer a highly probable phylogenetic hypothesis of the Heterobranchia. To reach this aim, a variety of statistical tests (like the Chi-Square-Test or Relative-Rate-Test) are conducted and substitution saturation is measured. In addition, bipartitions (splits) in the alignment are examined and visualized by split network analyses to estimate data quality.

Chapter 3 provides a new phylogenetic hypothesis based on a multigene approach using nuclear (18S rDNA and 28S rDNA) as well as mitochondrial (16S rDNA and COI) sequences. The dataset with the highest phylogenetic signal (as estimated with the methods described in chapter 2) is used for phylogenetic inference. This is the first time a large number of representatives of "Lower Heterobranchia" is included along with taxa of most of the major Euthyneura groups.
The aim of chapter 3 is to reconstruct the phylogeny of the Heterobranchia by means of Maximum likelihood and Bayesian inference methods. Moreover, based on the phylogenetic hypothesis proposed here, various evolutionary scenarios are discussed in order to give new insights into evolutionary trends within Heterobranchia.

[1] Regarding the genes
[2] Regarding the sequences

Chapter 4 gives first insights into the evolution of the Heterobranchia by using fossil data and molecular clock approaches in order to calibrate the phylogenetic tree and to estimate divergence ages.
The aim of chapter 4 is to estimate divergence times of groups belonging to the Heterobranchia with the newly developed software Beast which is a relaxed-clock Bayesian dating approach. Moreover, in order to place the phylogeny into a temporal framework and to recover possible correlation patterns between accelerated lineage splitting and mass extinction events, the lineages through time diversification patterns are analysed. These results are discussed in an evolutionary context.

Chapter 5 deals with the *a posteriori* evaluation of data quality using various approaches to verify the phylogenetic hypotheses proposed in chapter 3.
The aim of chapter 5 is to prove or reject the plausibility of tree reconstruction. The AU (Approximately Unbiased) Test is performed to evaluate how likely alternative hypotheses are. Furthermore, to verify the phylogenetic hypothesis of chapter 3, secondary structures of an almost complete 18S rRNA and a reduced 28S rRNA dataset are reconstructed. The secondary structures are treated as morphological characters and are parsimoniously mapped onto the phylogenetic tree in order to search for potential synapomorphies for members of certain clades. For this purpose, a recently developed software for secondary structure reconstruction called RNAsalsa is tested. The obtained consensus structures are used to determine evolutionary models specific to rRNA stem (paired basepairs) and loop (unpaired basepairs) regions with the intention to improve phylogenetic results.

Chapter 6 provides a review of the newly obtained results regarding heterobranch phylogeny. Furthermore, a general discussion of the employed methods regarding the results of the present study is given.

Chapter 7 gives a prospect for future projects while underscoring the inclusion of additional "Lower Heterobranchia" taxa such as Mathildoidea (Architectonicoidea), *Amathina* (Pyramidelloidea), Hyalogyrinidae (Valvatoidea?) and Xylodisculidae (Valvatoidea). The utility of using new phylogenetic tools (e.g. 3D reconstruction) and markers (e.g. gene arrangement, ESTs) is also discussed.

2. *A priori* evaluation of data quality

2.1 Introduction

Molecular phylogenies are usually based upon data whose quality has not been investigated *a priori* to tree inference. This could lead to incorrect results because phylogenetic trees obtained with traditional methods conceal conflicting evidence. To assess the reliability of an analysis conventional methods compare the fit between results and data (e.g. bootstrapping). Therefore, statistical support values may be high even if there is an ambiguous phylogenetic signal (Wägele & Mayer 2007).

Hence, any phylogenetic analysis should begin with an investigative evaluation of the quality of the dataset.

Several tools have been published that allow an *a priori* examination of data quality so far (Wu & Li 1985, Lyons-Weiler et al. 1996, Wilkinson 1998, Wägele & Rödding 1998, Holland et al. 2002, Xia et al. 2003, Mayer & Wägele 2005, Huson & Bryant 2006). Nevertheless, only a few scientists use them to test whether their data are suitable for a phylogenetic analysis or not (Wägele & Mayer 2007).

However, *a priori* analysis of data quality is a little explored field, and only a few tools that are independent of tree reconstruction are existing.

The first *a priori* analysis of data quality starts with the alignment. Very often a reliable alignment of divergent regions is hopeless because positional homology cannot be detected unambiguously. Especially hyper variable regions, nested within conserved, slowly evolving sections of ribosomal RNA sequences make the aligning procedure difficult and can have an impact on phylogenetic analyses. Thus, some authors proposed to search for ambiguous alignment positions and exclude them before tree reconstruction (Kjer 1995). Nevertheless, the removal of problematic alignment regions could have a strong influence on the tree reconstruction. Therefore, scientists should protocol and justify the exclusion of data (Gatesy et al. 1993).

Generally, ambiguous nucleotide sites are excluded from the alignment prior to a phylogenetic analysis by visual judgement which depends of course on the intuition or experience of the scientist and is rarely impartial.
A few approaches are available, applying objective algorithms to identify ambiguous alignment positions. Programs like Comalign (Bucka-Lassen et al. 1999), T-Coffee (Notredame et al. 2000), Gblocks (Castresana 2000), Soap (Loytynoja & Milinkovitch 2001), Altavist (Morgenstern et al. 2003) and Mumsa (Lassmann & Sonnhammer 2005) compare different alignments of similar sequences to test positional homology hypotheses and consistency of the alignments.

A new algorithm implemented in the software Aliscore (Misof & Misof, in press) is available and able to detect random similar sites (including ambiguously aligned positions and non-signal sections) which might have negative effects on tree reconstruction and exclusion of the identified characters is recommended.

The reliability of results from molecular phylogenetics also depends on how well the analysis deals with the problem whether some or all sequences in the data set have already lost phylogenetic information due to substitution saturation (Lopez et al. 1999, Philippe & Forterre 1999). Moreover, substitution saturation decreases phylogenetic information contained in the sequences and interferes phylogenetic analysis aiming to resolve deep nodes (Xia et al. 2003). In the worst case, sequences have experienced full substitution saturation and the similarity between the sequences which depends entirely on the similarity in nucleotide frequencies does not reflect phylogenetic relationships (Xia 2000).
There are currently two main approaches to test the degree of substitution saturation *a priori* in the aligned nucleotide sequences: The first approach plots patristic distances against distances obtained with different models of sequence evolution. The second approach developed by Xia et al. (2003) has been implemented in the software DAMBE (Xia 2000, Xia & Xie 2001) and is a new entropy-based index of substitution saturation.

Various other statistical tests exist for evaluating the data quality *a priori*, e.g. estimating the base composition to check whether there is a variation in GC content among the investigated species which can influence tree reconstructing or conducting a Chi-Square-Test to test for homogeneity of base frequencies across taxa.

The most promising *a priori* approach to evaluate data quality is the examination of bipartitions (splits) that are present in an alignment, to compare their support by nucleotide patterns, and to check the compatibility of these patterns (Wägele & Mayer 2007). To visualize these splits two different methods can be used: split decomposition (networks) and split support spectra. Networks or spectra of supporting positions can be generated without reference to a tree topology or a model of sequence evolution and are therefore ideal tools for *a priori* estimation of data quality (Wägele & Rödding 1998). Most notably is the possibility of networks to visualize various possible evolutionary scenarios and not only one evolutionary pathway like tree topologies do (Huson & Bryant 2006).
The first efficient tool to visualize split support present in an alignment was spectral analysis developed by Hendy & Penny (1993). Other methods followed like Rasa (Relative Apparent Synapomorphy Analysis) (Lyons-Weiler et al. 1996), Splits Randomization Tests (Wilkinson 1998), Physid (Wägele 1996 and Wägele & Rödding 1998) and δ Plots (Holland et al. 2002). Due to the large computing time increasing exponentially with the number of sequences Wägele & Mayer (2007) developed a simpler method (implemented in the software SAMS 1.4 beta). This method searches only for those splits that are represented in the data and visualises them as split support spectra. Additionally, Huson & Bryant (2006) provided a new program called SplitsTree4, an interactive and comprehensive tool for inferring different types of phylogenetic networks from sequences, distances and trees.

This cheapter deals with *a priori* evaluation of the molecular data and aims at recovering the most informing dataset of the three available concatenated datasets. Moreover, it will be tested whether the data are suitable for phylogenetic analysis and contain enough phylogenetic signal to infer a highly probable phylogenetic hypothesis.

2.2 Material and methods

Taxon sampling

A total of 52 gastropod species have been investigated (2 Vetigastropoda, 4 Caenogastropoda, 18 "Lower Heterobranchia", 14 Opisthobranchia, 12 Pulmonata and 2 taxa not assigned to the Heterobranchia yet). For details about the taxonomy and collecting locations of the sampled taxa as well as Genbank accession numbers see tab. A1 in the appendix.

The animals were collected from the field by hand, snorkelling or scuba diving and stored in 70-100% ethanol. Most of the "Lower Heterobranchia" were collected intertidally by collecting algae or substrata where they are living on. The material was washed and sieved and the animals were picked alive under the binocular.

DNA extraction, amplification and sequencing

For details on used chemicals and kits see also tab. A2 in the appendix.

Until further processing specimens were stored in 70–100% ethanol at -20 °C. DNA was isolated from foot tissue or the entire animal using the DNeasy Tissue Kit (Qiagen, Hilden, Germany) according to manufacturer's instructions.

The amount of obtained DNA was evaluated by electrophoresis with the molecular weight marker Lamda-Hind-III-Ladder in a 1,4% agarose gel in 10x TBE buffer. The DNA was visualised with ethidium bromide and documented with the camera Canon Power Shot G9 and the software PS Remote 1.5.7.

Sequences of the complete nuclear 18S rDNA, partial nuclear 28S rDNA and partial mitochondrial 16S rDNA and one protein coding gene fragment (Cytochrome C Oxidase subunit I – COI) were amplified.

18S rRNA and 28S rRNA are slowly evolving genes and are known to be more conservative than 16S rRNA and COI, hence they were used to infer deep phylogenetic nodes (e.g. order and family level). 16S rRNA as well as COI are fast evolving genes and were therefore used to reconstruct terminal nodes (e.g. genus and species level).

The PCR technique was used to amplify defined gene fragments (primer designs see tab. A3 in the appendix). PCRs were generally performed using a standard protocol (see tab. 2.1) for 18S rDNA, 16S rDNA and COI and a slightly modified protocol for 28S rDNA. To check for contaminations negative controls (dH_2O) were included in each reaction array.

Tab. 2.1: PCR protocol for a total reaction volume of 25µl

Components	Concentration	Volumes for the standard protocol	Volumes for the modified protocol
DNA	1 ng	5,00 µl	5,00 µl
Taq polymerase	1 Unit	0,20 µl	0,20 µl
Buffer	10x	2,50 µl	2,00 µl
$MgCl_2$	50mM	2,00 µl	1,00 µl
dNTP	25mM	0,20 µl	0,20 µl
Primer	10nmol	1,00 µl	0,80 µl
Primer	10nmol	1,00 µl	0,80 µl
BSA	10mg/ml	1,50 µl	1,00 µl
TMAC	0,5M	0,25 µl	-
DMSO	0,5M	-	1,25 µl
dH_2O	-	11,35 µl	12,75 µl

Thermal cycling was performed with a Primus 96 AdvancedGradient Thermal Cycler (Peqlab, Erlangen, Germany) using the following programs:

a) 18S (annealing temperature 52,5 °C), 16S and COI (annealing temperature 52 °C)

Denaturation	95 °C	01:00 min	
Denaturation	95 °C	00:30 min	
Annealing	52-52,5 °C	00:30 min	30x
Extension	72 °C	00:30 min	
Extension	72 °C	03:00 min	
Store	08 °C	Forever	

b) 28S

Denaturation	95 °C	04:00 min	
Denaturation	94 °C	00:30 min	
Annealing	52,5 °C	00:30 min	38X
Extension	72 °C	02:50 min	
Extension	72 °C	10:00 min	
Store	08° C	Forever	

The success of the PCR was verified by electrophoresis with the molecular weight marker 100-bp-DNA-Leiter-extended in a 1,4% agarose gel in 10x TBE buffer. The DNA was visualised with ethidium bromide and documented with the camera Canon Power Shot G9 and the software PS Remote 1.5.7.

Amplification products were purified by cutting out corresponding bands from a 1,4% agarose gel. DNA was isolated from the gel using a QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany) following the manual instructions. Both sense and antisense strands were sequenced directly either on the CEQ 2000 Beckmann Coulter capillary sequencer at the

Institute for Ecology, Evolution and Diversity, Frankfurt/Main or on the ABN 3130 XL Applied Biosystems capillary sequencer at the SRD GmbH, Bad Homburg.

Sequence editing and alignment

To check if the correct genes have been amplified BLAST searches (Altschul et al. 1990) were performed to compare amplified sequences with all sequences stored in the Genbank database (www.ncbi.nlm.nih.gov/Genbank/index.html).

Sequence chromatograms of each amplified fragment were displayed with the software Chromas lite 2.0.1 (www.technelysium.com.au/chromas_lite.html) and browsed for reading mistakes of the sequencer by eye.

Sequences were aligned using the default parameters of Muscle 3.6 (Edgar 2004) and checked manually with BioEdit 7.0.5.3 (Hall 1999). Regions which could not be unambiguously aligned and long inserts were excluded by eye or based on the analysis of the software Aliscore 0.2 (see tab. A4 in the appendix).

The following datasets were composed:

Dataset 0 = combination of complete 18S rDNA, partial 28S rDNA, partial 16S rDNA and COI sequences; no alignment positions were excluded (see tab. A5 in the appendix).

Dataset I = combination of complete 18S rDNA, partial 28S rDNA, partial 16S rDNA and COI sequences; long inserts and ambiguous alignment positions were excluded by visual judgement (see tab. A4 and A5 in the appendix).

Dataset II = combination of complete 18S rDNA, partial 28S rDNA, partial 16S rDNA and COI sequences; ambiguous alignment positions were determined with the software Aliscore 0.2 and excluded from further analyses (see tab. A4 and A5 in the appendix).

Aliscore

Random similarity within multiple sequence alignments were identified with the software Aliscore 0.2 (see tab. A4 in the appendix) which has been newly invented by Misof & Misof (in press) and Bernhard Misofs former working group at the Forschungsmuseum König in

Bonn. This method is based on Monte Carlo (MC) resampling within a sliding window. The MC resampling compares the score of the originally aligned sequences in a given window position with scores of randomly drawn sequences of similar character composition. Sequences are assumed unrelated if the observed score is not better then 95 % of scores of random sequences of similar window size and character composition.

Substitution saturation

Gene sequences can become saturated when the visible genetic distance of the sequences may not increase at the same rate as the evolutionary distances. This could be due to multiple substitutions when comparing the same gene fragment in different taxa if these taxa have been separated by long divergence times. This could lead to a loss of phylogenetic information within the sequences.

The substitution saturation was tested in two different ways:

1. with the test by Xia et al. (2003) implemented in the software DAMBE 4.5.47 (Xia & Xie 2001). This method is based on the notion of entropy in information theory. One derives the critical values of the index based on computer simulation with different sequence lengths, different number of taxon units and different topologies. A quick evaluation whether a set of aligned sequences is useful for phylogenetic studies is possible.
2. by plotting patristic distances against distances obtained with different models of sequence evolution (see tab. A5 in the appendix). Transition and transversion data were calculated with the program PAUP 4.0 beta 10 (Swofford 2002) and examined separately.

Base composition and Chi-Square-Test

Base compositions were estimated using the software PAUP 4.0 beta 10 (Swofford 2002) and the software SAMS 1.4 beta (Mayer & Wägele 2005) to check whether there is a variation in GC content among the investigated species. This variation can influence tree reconstructing because unrelated species with similar GC content are often grouped together.

A Chi-Square-Test was conducted using the program PAUP 4.0 beta 10 (Swofford 2002) to test for homogeneity of base frequencies across taxa.

Relative-Rate-Test

The relative rate test is used to check whether two species evolve at the same rate by testing whether their distances to an outgroup are equal (Philippe & Laurent 1998). In this study, the relative rate test of Wu and Li (1985) as implemented in the program K2WuLi (Jermiin 1997) was performed and *Littorina littorea* (Caenogastropoda) was used as outgroup.

Incongruence length difference test

The incongruence length difference (ILD) test was performed to verify whether the single 18S, 28S, 16S and COI data sets contain the same phylogenetic signal and therefore could be analysed as a single concatenated dataset (taxa for which a gene region was unavailable were excluded from the test). The test, described by Farris et al. (1994) measures the significance of incongruence among data sets. The ILD test is also known as the partition-homogeneity test, which is implemented in the software PAUP 4.0b10 (Swofford 2002). Using the maximum parsimony criterion heuristic searches with 100 replicates were conducted.

Network analyses

To visualize variations in signal distinctness, network analyses were used based on split decomposition (applied with SplitsTree 4.10 (Huson & Bryant 2006)) and split support spectra (applied with Sams 1.4 beta (Mayer & Wägele 2005)). Both tools allow an *a priori* examination of data quality.

SplitsTree 4.10 was used to calculate phylogenetic networks. The compared network structures were based on the Neighbournet algorithm.

The phylogenetic signal present in the data that supports or contradicts putative splits were estimated with the program SAMS 1.4 beta using the default parameters and visualized with the diagram-assistant implemented in the program Microsoft Excel 2002.
SAMS is an analysing software for molecular data and implements several features which estimate the phylogenetic signal present in the data that supports or contradicts putative splits. With this information it is possible to visualize the information content of the data set and the signal to noise relationship.

2.3 Results

2.3.1 *A priori* evaluation of data quality by the identification of random similarity within sequence alignments using Aliscore and by visual judgement

Aliscore assigns every position in an alignment a positive or negative score; a positive value indicates non-random similarity a negative value random similarity. Alignment positions with negative scores are phylogenetically uninformative and are therefore advised to be excluded prior to phylogenetic analyses.

The single alignments comprised the following bp: 18S rDNA (complete) 2716 bp, 28S rDNA (partial) 1980 bp, 16S rDNA (partial) 722 bp and COI 579 bp (each codon position with 193 bp) (see also tab. A4 in the appendix).

Aliscore detected as putative randomly similar nucleotide positions within 18S rDNA 80 bp (2,95%), within 28S rDNA 171 bp (8,64%), within 16S rDNA 153 bp (21,19%), within COI first codon position 17 bp (8,81%), within second codon position none and within third codon position 175 bp (90,67%) (see fig. 2.1 and 2.2).

Due to the visual judgment 941 bp of 18S rDNA (34,66%), 1150 bp of 28S rDNA (58,08%), 444 bp of 16S rDNA (61,49%), none of COI first and second codon position and all of COI third codon positions (100%) were identified as inserts or ambiguous alignments.

Summing up the visual judgment yielded more ambiguous positions than Aliscore did but both methods identified the third codon position of COI as the one with the most and the 16S rDNA alignment as the one with the second most critical positions. Aliscore as well as visual judgment excluded no positions of the second codon position of COI and identified 18S rDNA as the alignment with the fewest critical positions.

In the following it is tested which of the three datasets (dataset 0 – all positions, dataset I – alignment positions were excluded by visual judgement and dataset II – alignment positions were excluded by Aliscore) (see also tab. A4 and A5 in the appendix) is the most informative one for phylogenetic reconstruction by *a priori* evaluation of the data.

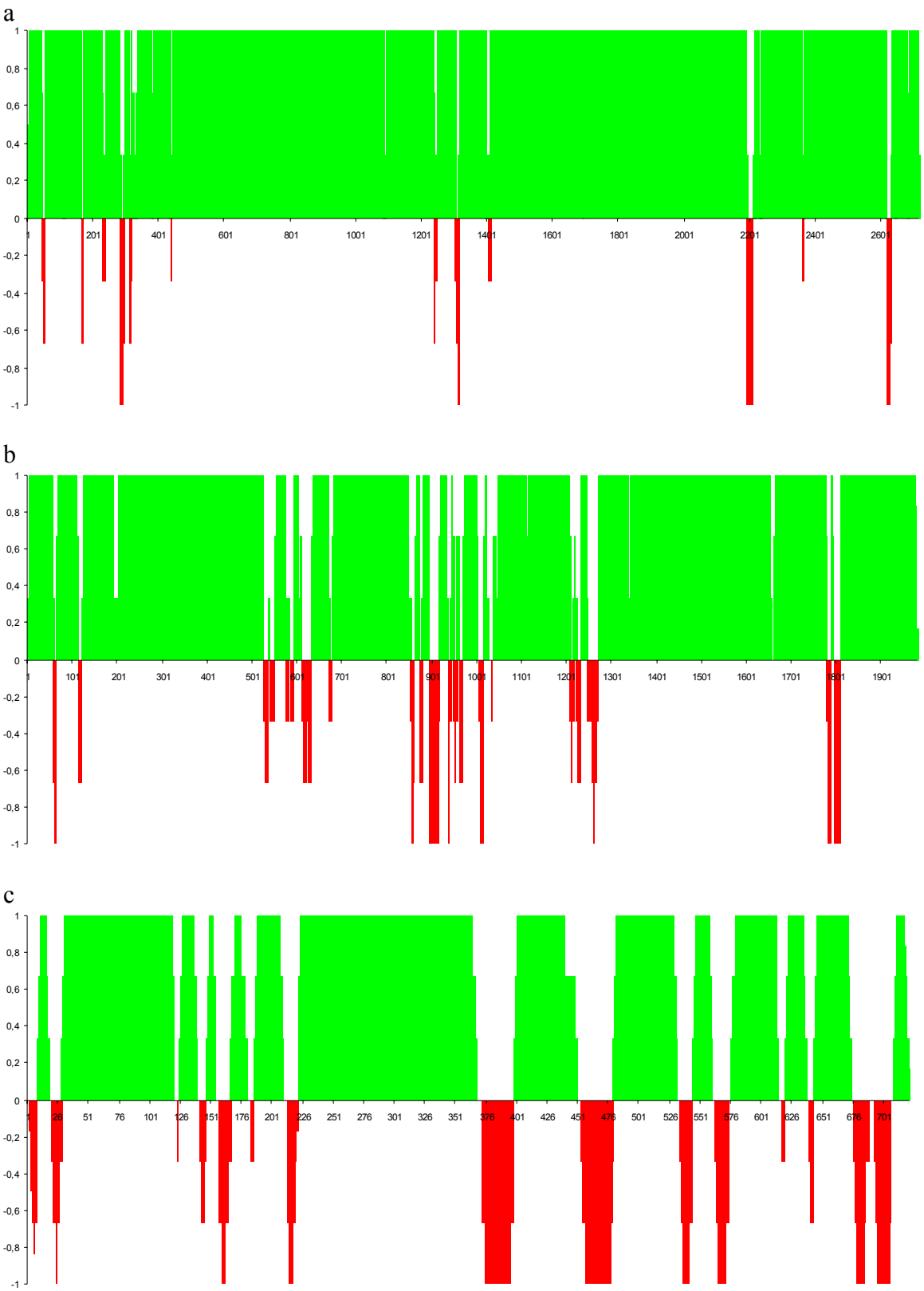


Fig. 2.1: Consensus profile of the Aliscore check for random similar characters of a: 18S rDNA, b: 28S rDNA and c: 16S rDNA; x-axis = alignment positions, y-axis = scores, green = positive scores, red = negative scores, positions with negative scores should been excluded from further investigations.

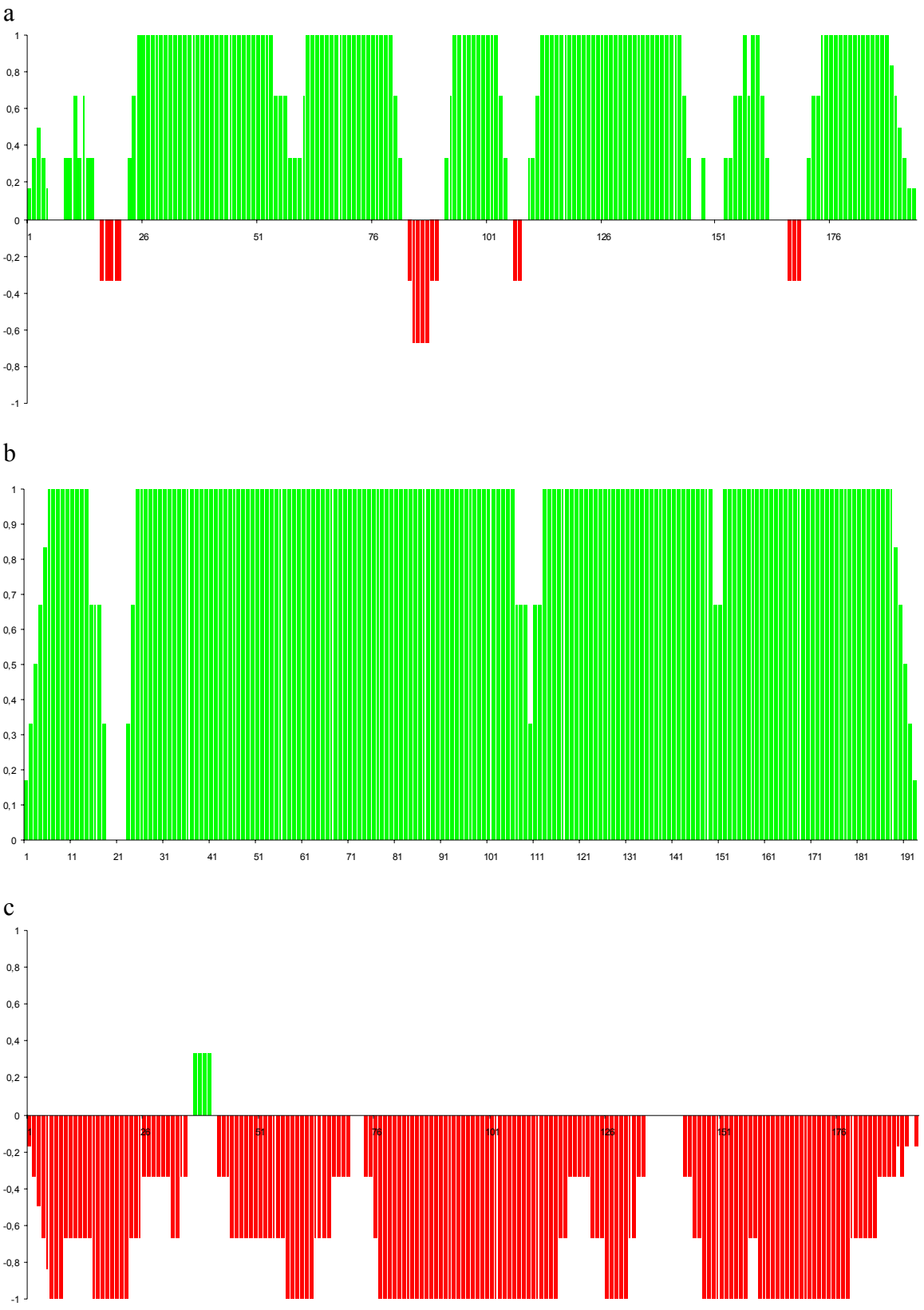


Fig. 2.2: Consensus profile of the Aliscore check for random similar characters of a: COI first codon position, b: COI second codon position and c: COI third codon position; x-axis = alignment positions, y-axis = scores, green = positive scores, red = negative scores, positions with negative scores should been excluded from further investigations.

2.3.2 *A priori* evaluation of data quality by the measurement of substitution saturation in the aligned nucleotide sequences

Here I tested the degree of substitution saturation in two different ways: by an index to measure substitution saturation developed by Xia et al. (2003) and by plotting patristic distances against distances obtained with different models of sequence evolution (graphically).

2.3.2.1 Index to measure substitution saturation (by Xia et al. 2003)

Genetic sequences will fail to recover the true phylogeny long before the full substitution saturation is reached indicated by the index of substitution saturation (Iss). For this reason, one needs to find the critical index of substitution saturation (Iss.c) at which the sequences will begin to fail to recover the true tree. According to Xia et al. (2003) the results of the test should be interpreted in the following way: Iss < Iss.c indicating little or no saturation while Iss > Iss.c indicating phylogenetic uninformative sequences.

Tab. 2.2: Substitution saturation meassured by Xia et al. (2003)

Dataset	Substitution saturation
Dataset 0	
18S rDNA	Iss 1,398 > Iss.c 0,371
28S rDNA	Iss 1,795 > Iss.c 0,455
16S rDNA	Iss 1,847 > Iss.c 0,398
COI position 1	Iss 0,369 > Iss.c 0,306
COI position 2	Iss 0,127 < Iss.c 0,306
COI position 3	Iss 0,793 > Iss.c 0,305
Dataset I	
18S rDNA	Iss 0,682 > Iss.c 0,336
28S rDNA	Iss 0,592 > Iss.c 0,357
16S rDNA	Iss 0,719 > Iss.c 0,332
COI position 1	Same as dataset 0
COI position 2	Same as dataset 0
COI position 3	No data
Dataset II	
18S rDNA	Iss 1,324 > Iss.c 0,367
28S rDNA	Iss 2,209 > Iss.c 0,347
16S rDNA	Iss 2,052 > Iss.c 0,381
COI position 1	Iss 0,326 > Iss.c 0,304
COI position 2	Iss 0,124 < Iss.c 0,306
COI position 3	Iss 0,705 > Iss.c 0,283

Comparing the substitution saturation data of the three datasets with each other (see tab. 2.2) it becomes evident that in all datasets only the second codon position of COI was not saturated. All other markers showed a high degree of saturation.
Comparing the Iss-Data with each other, dataset I showed the lowest Iss values in most of the markers.

2.3.2.2 Plotting patristic distances against distances obtained with different models of sequence evolution (graphically)

A sequence is saturated when the visible genetic distances (p-distances) of a sequence is not increasing at the same rate as the evolutionary distances (d-distances) because of multiple substitutions. Saturation can be detected with plots and a bisecting line indicating a linear increase of p- and d-distances. When the p-distances increase faster than the d-distances a sloping curve is the result falling below the bisecting line.

The graphs of all genes in all three datasets showed a high degree of saturation while transitions showed a higher saturation than transversions. The 18S rDNA curve shape in all three datasets was quite similar showing a first saturation effect at a value of 0.04 (fig. 2.3). This applied also to the 28S rDNA curve shape in all three datasets (fig. 2.4) showing a first saturation effect at the same value as 18S rDNA but with a more scattered curve shape for the transitions. All positions of 16S rDNA of dataset 0 were saturated (fig. 2.5). With the exception of a few positions at the beginning of the graph (up to a value of 0.04) of 16S rDNA of dataset I and II all positions showed saturation, too. Transition and transversion curve shapes were scattered. The COI curve shape of first and second codon position in all three datasets (dataset 0 and I are the same) was quite similar showing a first saturation effect at a value of 0.04 and a scattered curve shape (figs. 2.6, 2.7). A graphical display of the third codon position of COI of all three datasets was not possible because PAUP was not able to calculate the genetic distances. The software stopped at a d-distance of 4.664.742.279 (transversion) and 115.871 (transition) in dataset 0 and I and at a d-distance of 402.808.285 (transversion) and 2.183.206.940 (transition) in dataset II. All four values indicated a genetic distance higher than by chance.

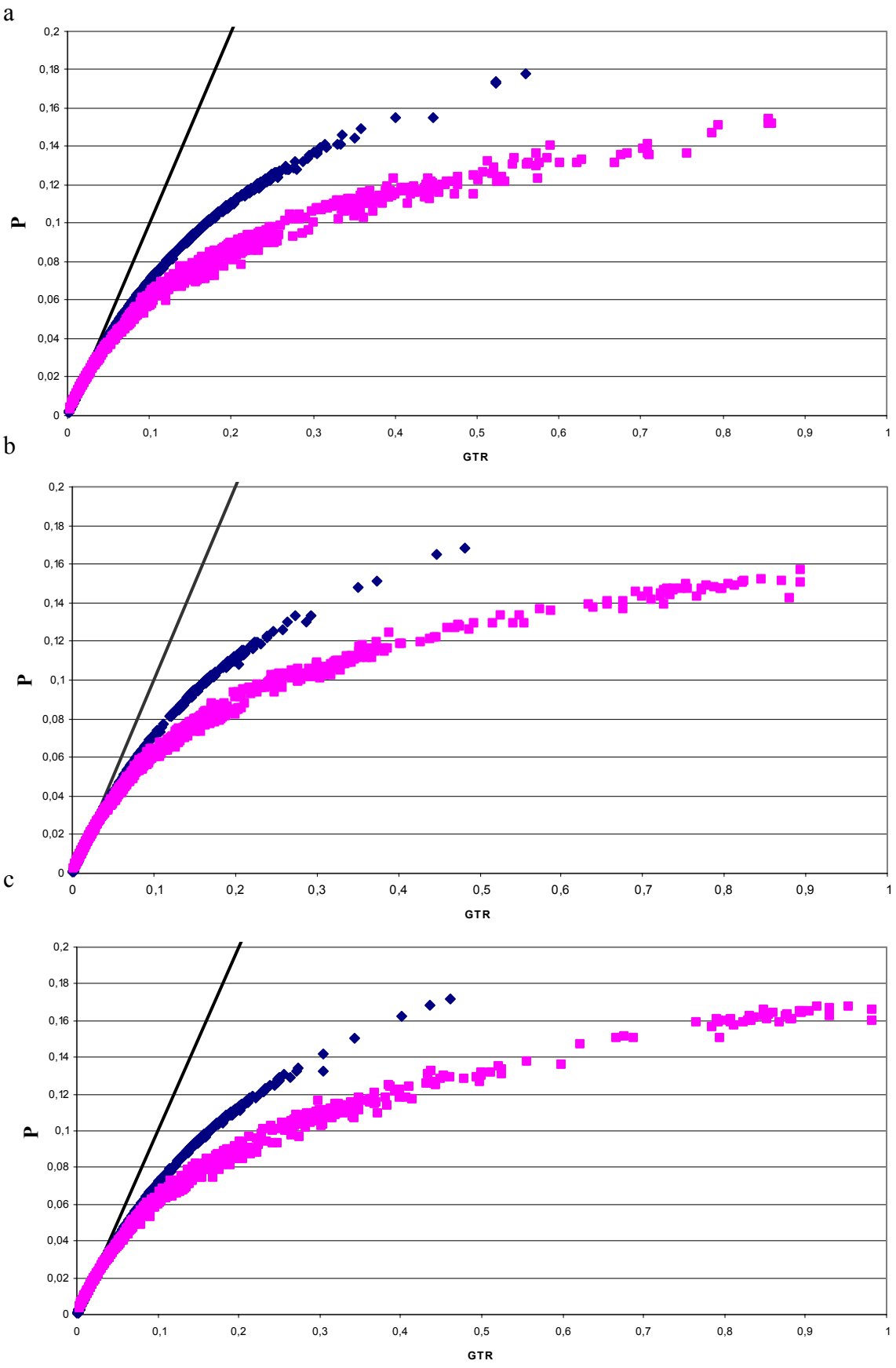


Fig. 2.3: Saturation of substitution of 18S rDNA. Distances are calculated as patristic distances (y-axis) against d-distances calculated by applying the GTR model (x-axis), blue = transversion, pink = transition; a: dataset 0; b: dataset I; c: dataset II.

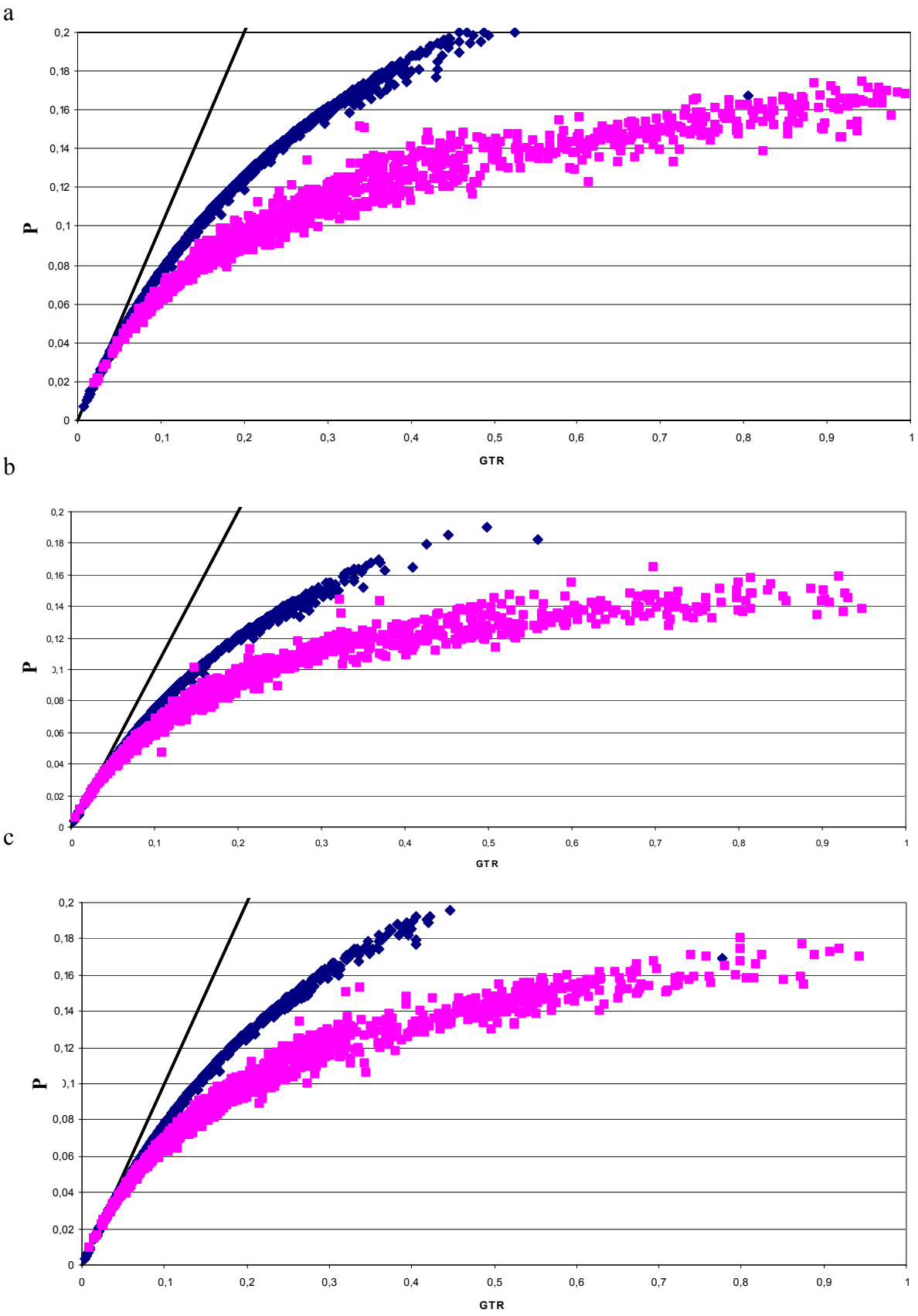


Fig. 2.4: Saturation of substitution of 28S rDNA. Distances are calculated as patristic distances (y-axis) against d-distances calculated by applying the GTR model (x-axis), blue = transversion, pink = transition; a: dataset 0; b: dataset I; c: dataset II.

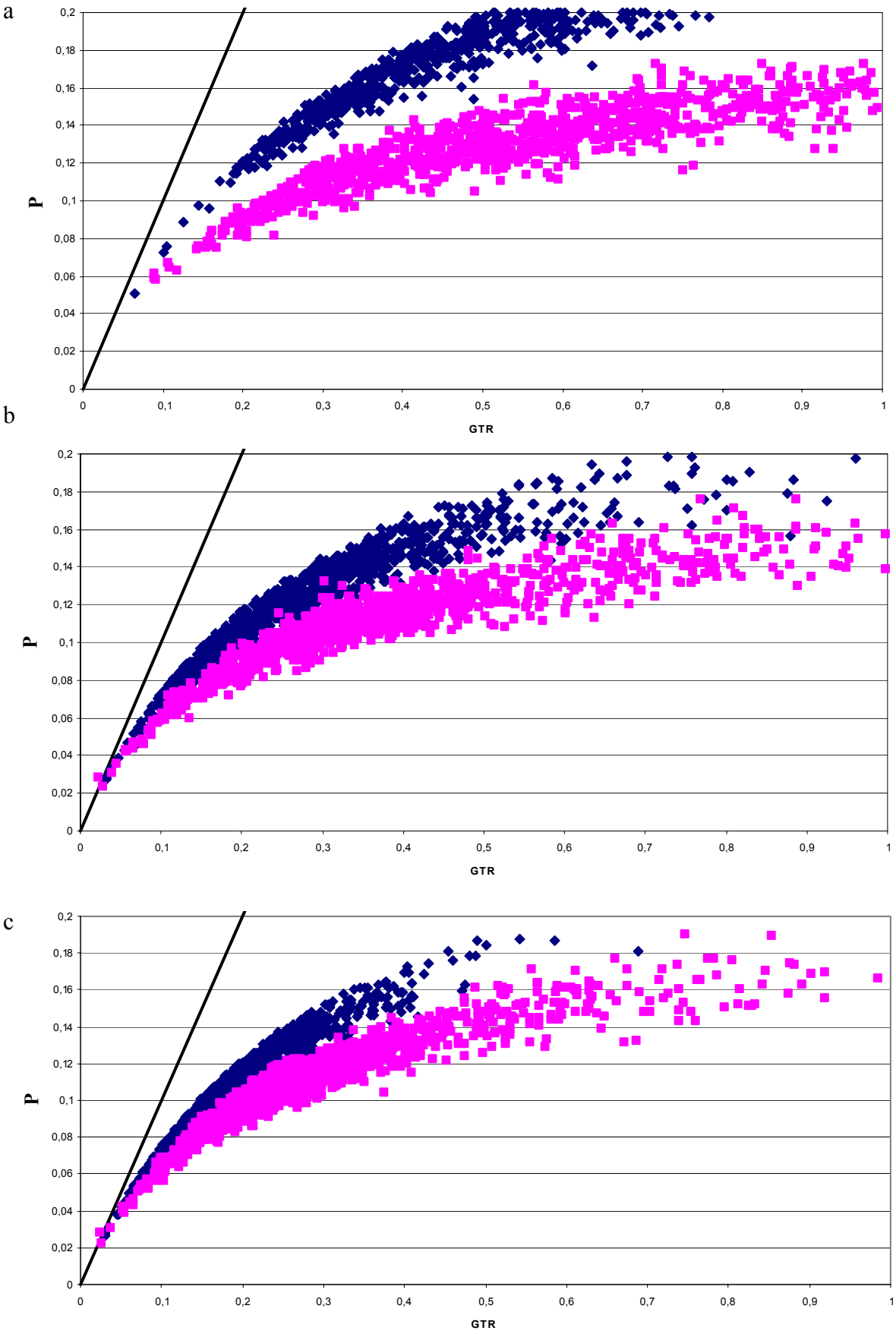


Fig. 2.5: Saturation of substitution of 16S rDNA. Distances are calculated as patristic distances (y-axis) against d-distances calculated by applying the GTR model (x-axis), blue = transversion, pink = transition; a: dataset 0; b: dataset I; c: dataset II.

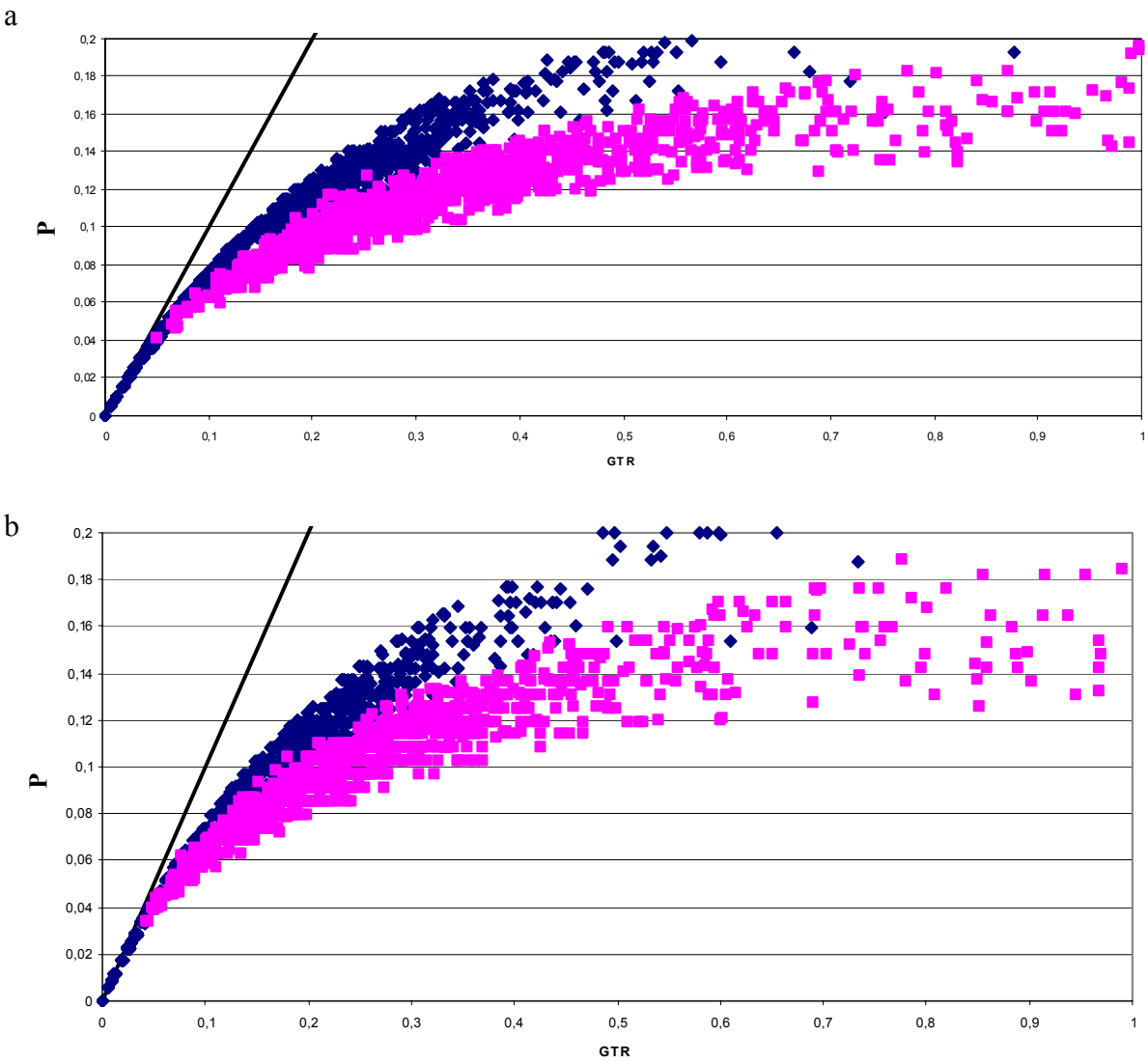


Fig. 2.6 Saturation of substitution of COI position 1. Distances are calculated as patristic distances (y-axis) against d-distances calculated by applying the GTR model (x-axis), blue = transversion, pink = transition; a: dataset 0 and I; b: dataset II.

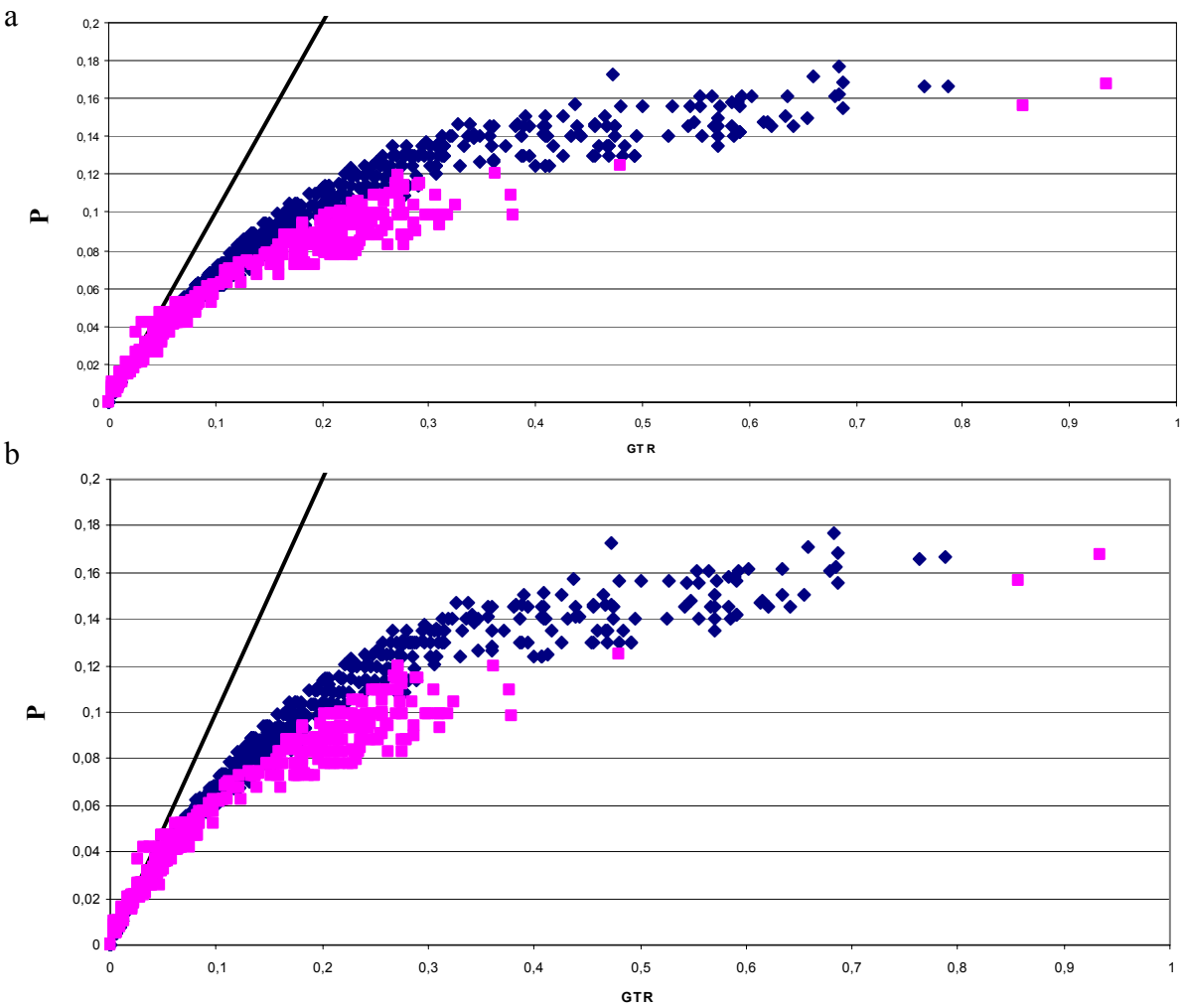


Fig. 2.7: Saturation of substitution of COI position 2. Distances are calculated as patristic distances (y-axis) against d-distances calculated by applying the GTR model (x-axis), blue = transversion, pink = transition; a: dataset 0 and I; b: dataset II.

2.3.3 *A priori* evaluation of data quality by a variety of statistical tests

2.3.3.1 Base composition

The mean base frequencies of all markers show the same distribution in all datasets (see tab. 2.3). The G+C content is higher than the A+T content in the 18S rDNA and 28S rDNA and lower in the 16S rDNA and all positions of COI.

Tab. 2.3 Mean base frequencies

Genes	A+T	G+C
Dataset 0		
18S rDNA	0.47477	0.52523
28S rDNA	0.35938	0.64062
16S rDNA	0.63891	0.36109
COI position 1	0.55414	0.44586
COI position 2	0.58160	0.41840
COI position 3	0.78410	0.21590
Dataset I		
18S rDNA	0.49403	0.50597
28S rDNA	0.38121	0.61879
16S rDNA	0.59006	0.40994
COI position 1	same as dataset 0	same as dataset 0
COI position 2	same as dataset 0	same as dataset 0
COI position 3	no data	no data
Dataset II		
18S rDNA	0.47776	0.52224
28S rDNA	0.36875	0.63124
16S rDNA	0.59482	0.40518
COI position 1	0.55560	0.44441
COI position 2	0.58160	0.41840
COI position 3	0.86309	0.13691

Omalogyra fusca showed the largest deviation of the mean G+C content in the 18S rDNA in all three datasets (0 = 0.586079, I = 0.5723319, II = 0.585003) while *Ebala sp.* showed the largest deviation of the mean A+T content in the 18S rDNA in all 3 datasets (0 = 0.650894, I = 0.638568, II = 0.646868).

Chromodoris krohni showed the largest deviation of the mean G+C content in the 28S rDNA in dataset 0 (0.713998) and I (0.688950) while *Architectonica perspectiva* showed the largest deviation of the mean G+C content in the 28S rDNA in dataset II (0.693784). *Murchisonella* sp. showed the largest deviation of the mean A+T content in the 18S rDNA in all three datasets (0 = 0.437560, I = 0.433113, II = 0.447939).

Diodora graeca showed the largest deviation of the mean G+C content in the 16S rDNA in all three datasets (0 = 0.454357, I = 0.484496, II = 0.467066) while *Ebala sp.* showed the largest

deviation of the mean A+T content in the 16S rDNA in dataset 0 (0.732620), *Murchisonella* sp. showed the largest deviation of the mean A+T content in the 16S rDNA in dataset I (0.681452) and *Omalogyra fusca* showed the largest deviation of the mean A+T content in the 16S rDNA in dataset II (0.676157).
Bathymargarites symplector showed the largest deviation of the mean G+C content in the first codon position of COI in all three datasets (0 = 569948, I = same as dataset 0, II = 0.556818) while *Omalogyra* sp. showed the largest deviation of the mean A+T content in the first codon position of COI in all three datasets (0 = 661458, I = same as dataset 0, II = 0.685714).
None of the taxa showed a mentionable deviation of the mean G+C content as well as the A+T content in the second codon position of COI.
Diodora graeca showed the largest deviation of the mean G+C content in the third codon position of COI in dataset 0 (0.450777) while *Murchisonella* sp. showed the largest deviation of the mean A+T content in the third codon position of COI in dataset 0 (0.927461). The results of dataset II for the third codon position was not representative because of the small number of base positions (18).

To sum up little deviation of the mean G+C and A+T content was observed in all markers of all datasets. If a deviation was observed then within taxa belonging to the Vetigastropoda or "Lower Heterobranchia" (and in one case to the Opisthobranchia).

2.3.3.2 Chi-Square-Test

When $P \leq 0{,}05$ the base composition indicates a significant heterogeneity and when $P \geq 0{,}05$ the base composition indicates a significant homogeneity.

Within dataset 0 only the first and second codon position of COI showed homogeneity of base frequencies while 18S rDNA, 28S rDNA, 16S rDNA and third codon position of COI showed a heterogeneity (see tab. 2.4). Within dataset I all sequences with the exception of 18S rDNA showed homogeneity of base frequencies. Within dataset II the sequences of 16S rDNA and all codon positions of COI showed homogeneity while 18S rDNA and 28S rDNA showed heterogeneity of base frequencies.

Tab. 2.4: Chi-Square-Test of homogeneity of base frequencies across taxa

Sequences	Chi-Square	df	P
Dataset 0			
18S rDNA	351.024.747	153	0.00000000
28S rDNA	312.690.985	153	0.00000000
16S rDNA	248.124.449	138	0.00000003
COI position 1	132.394.187	150	0.84621041
COI position 2	26.833.931	150	1.00000000
COI position 3	756.200.271	150	0.00000000
Dataset I			
18S rDNA	222.934.498	153	0.00019294
28S rDNA	121.079.129	153	0.97325718
16S rDNA	105.994.410	138	0.98023221
COI position 1	same as dataset 0	same as dataset 0	same as dataset 0
COI position 2	same as dataset 0	same as dataset 0	same as dataset 0
COI position 3	no data	no data	no data
Dataset II			
18S rDNA	323.553.686	153	0.00000000
28S rDNA	240.363.336	153	0.00000821
16S rDNA	113.982.072	141	0.95396238
COI position 1	111.101.936	150	0.99258445
COI position 2	26.833.931	150	1.00000000
COI position 3	163.771.109	150	0.20891673

2.3.3.3 Relative-Rate-Test

For a better estimation of the relative substitution rates within a dataset it is important to choose a closely related reference taxon. The closer the outgroup is related to the ingroup the higher is the probability to estimate differences in relative rates correctly. Therefore a Caenogastropoda (*Littorina littorea*) and not a Vetigastropoda was defined as outgroup for this test.

The Relative-Rate-Test revealed no significant difference in evolutionary rates between dataset 0, I and II but between the investigated taxa and genetic markers (tab. 2.5).
One could observe the highest evolution rates within the „Lower Heterobranchia“ (*Omalogyra* sp., *Omalogyra fusca*, *Murchisonella* sp., *Ebala* sp. and *Architectonica perspectiva*) and Opisthobranchia (*Hyalocylis striata*) (tab. 2.5). 18S rDNA and 28S rDNA were the markers with the highest z-scores while 16S and the first and second codon position of COI were the markers with the lowest z-scores.
The program K2WuLi was not able to estimate the relative rates of the third codon position of COI in all three datasets.

Tab. 2.5: Maximum z-scores (only scores above 5.0 are shown - with the exception of 16S rDNA)

Sequences	Species	z-scores (max.)
Dataset 0		
18S rDNA	***Omalogyra* sp.** vs. *Orbitestella vera*	13.379992
	Omalogyra fusca vs. *Orbitestella vera*	13.323153
	***Murchisonella* sp.** vs. *Orbitestella vera*	12.860143
	***Ebala* sp.** vs. *Orbitestella vera*	11.155443
	Architectonica perspectiva vs. *Orbitestella vera*	10.658446
28S rDNA	***Omalogyra* sp.** vs. *Orbitestella vera*	9.138800
	***Murchisonella* sp.** vs. *Orbitestella vera*	8.586671
	***Ebala* sp.** vs. *Orbitestella vera*	8.108975
	Omalogyra fusca vs. *Orbitestella vera*	7.724675
	Architectonica perspectiva vs. *Otina ovata*	7.688050
16S rDNA	***Architectonica perspectiva*** vs. *Cornirostra pellucida*	4.256653
COI position 1	***Omalogyra* sp.** vs *Aperostoma pelermi*	5.273325
	Omalogyra fusca vs *Aperostoma pelermi*	5.118954
	Hyalocylis striata vs. *Cornirostra pellucida*	4.248102
COI position 2	***Architectonica perspectiva*** vs *Valvata piscinalis*	4.726909
COI position 3		no data
Dataset I		
18S rDNA	***Omalogyra fusca*** vs *Orbitestella* sp.	13.722511
	***Omalogyra* sp.** vs *Orbitestella* sp.	13.472862
	***Murchisonella* sp.** vs *Orbitestella vera*	12.782731
	***Ebala* sp.** vs *Orbitestella* sp.	11.007350
	Architectonica perspectiva vs *Orbitestella vera*	10.742792
	Larochella alta vs *Orbitestella vera*	8.970545
28S rDNA	***Omalogyra* sp**. vs *Orbitestella vera*	8.040198
	Architectonica perspectiva vs *Orbitestella vera*	7.946169
	***Ebala* sp**. vs *Orbitestella vera*	7.815249
	Omalogyra fusca vs *Orbitestella vera*	7.723543
	***Murchisonella* sp.** vs *Orbitestella vera*	7.766580
16S rDNA	***Architectonica perspectiva*** vs *Umbraculum umbraculum*	4.192090
COI position 1		same as dataset 0
COI position 2		same as dataset 0
COI position 3		no data
Dataset II		
18S rDNA	***Omalogyra* sp.** vs. *Orbitestella vera*	13.618592
	Omalogyra fusca vs. *Orbitestella vera*	13.507419
	***Murchisonella* sp.** vs. *Orbitestella vera*	13.099827
	***Ebala* sp.** vs. *Orbitestella vera*	11.350836
	Architectonica perspectiva vs. *Orbitestella vera*	11.133709
28S rDNA	***Omalogyra* sp.** vs. *Orbitestella vera*	9.348291
	***Murchisonella* sp**. vs. *Orbitestella vera*	8.691861
	***Ebala* sp.** vs. *Orbitestella vera*	8.433427
	Architectonica perspectiva vs. *Orbitestella vera*	8.131231
	Omalogyra fusca vs. *Orbitestella vera*	8.037855
16S rDNA	***Architectonica perspectiva*** vs. *Cornirostra pellucida*	4.817522
COI position 1	***Omalogyra* sp.** vs. *Cornirostra pellucida*	5.625573
	Architectonica perspectiva vs. *Cornirostra pellucida*	5.609777
	Omalogyra fusca vs. *Cornirostra pellucida*	5.531493
COI position 2	***Hyalocylis striata*** vs. *Valvata piscinalis*	4.738736
COI position 3		no data

2.3.3.4 ILD Test

Investigation of differences in incongruence length between 18S rDNA, 28S rDNA, 16S rDNA and COI revealed that combination of the partitions improves phylogenetic signal with a p-value of 0.01 (10.000 replicates) in all three datasets.
According to this result the next chapters investigated the combined datasets only.

2.3.4 *A priori* evaluation of phylogenetic signal by split network analyses

To visualise variations in signal distinctness we used network analyses based on split decomposition (applied with SplitsTree 4.10) and split support spectra (applied with SAMS 1.4 beta).

2.3.4.1 SplitsTree

In a network a split between two groups is indicated by parallel edges of the same length while edge length is proportional to the weight of the associated split (Huson & Bryant 2006).

Taking a first look at the network of all three datasets one was able to see a high level of conflict indicated by many parallel edges of the same lengths (forming a netlike structure). Several taxa with long terminal branches could be identified in all three datasets belonging to the Veti- and Caenogastropoda, “Lower Heterobranchia” or Opisthobranchia (Nudipleura) (see figs. 2.8 – 2.10).

Investigating the deep splits (regarding higher taxonomic levels) different split support within the three datasets was evident (see figs. 2.8 – 2.10 – deep splits are marked with a dotted line). The Vetigastropoda and Caenogastropoda alone were supported by splits in all three datasets while the Caenogastropoda and Vetigastropoda together as outgroup differentiated to the Heterobranchia (I-XXV) were only supported within dataset 0 and I. Neither the Pulmonata nor the Opisthobranchia were supported by any split in any of the datasets. However, the Nudipleura (III in dataset 0, VII in dataset I and XIX in dataset II) had very good split support in all three datasets.

There were several deep splits only occurring within one of the three datasets. Dataset 0 showed a split separating the Veti- and Caenogastropoda together with the „Lower Heterobranchia“ (with exception of Cimidae I, Rissoelloidea II, and Acteonoidea XIX) from the remaining taxa (see fig. 2.8). Within dataset I a deep split separated the Pyramidellidae (XII), Glacidorboidea (XIV), Pulmonata (with exception of Amphiboloidea XXI), Umbraculoidea (VIII), Akeroidea (X) and Pteropoda (IX+XIII) from the remaining taxa (see fig. 2.9). Within dataset II a deep split separated the Architectonicoidea (XXIV), Omalogyroidea (XXV) and Murchisonellidae (XXIII) from the remaining taxa. An additional split partitioned the Pyramidellidae (X), Glacidorboidea (XI), Pulmonata, Opisthobranchia (with the exception of Nudipleura XIX), Acteonoidea (III) and Rissoelloidea (II) from the remaining sequences (see fig. 2.10).

According to Bouchet & Rocroi (2005) (see classification in tab. A1 in the appendix) most of the major heterobranch subgroups were supported by splits in all three combined datasets (see figs. 2.8 – 2.10). Only Pyramidelloidea (as well as Pyramidellidae) and Siphonaroidea were not supported by any split. Stylommatophora, Onchidioidea and Nudipleura were supported by splits in all three datasets while Hygrophila were only supported in dataset I and II. A sister group relationship between Omalogyroidea and Architectonicoidea was well supported by splits with very long edges in all three datasets.

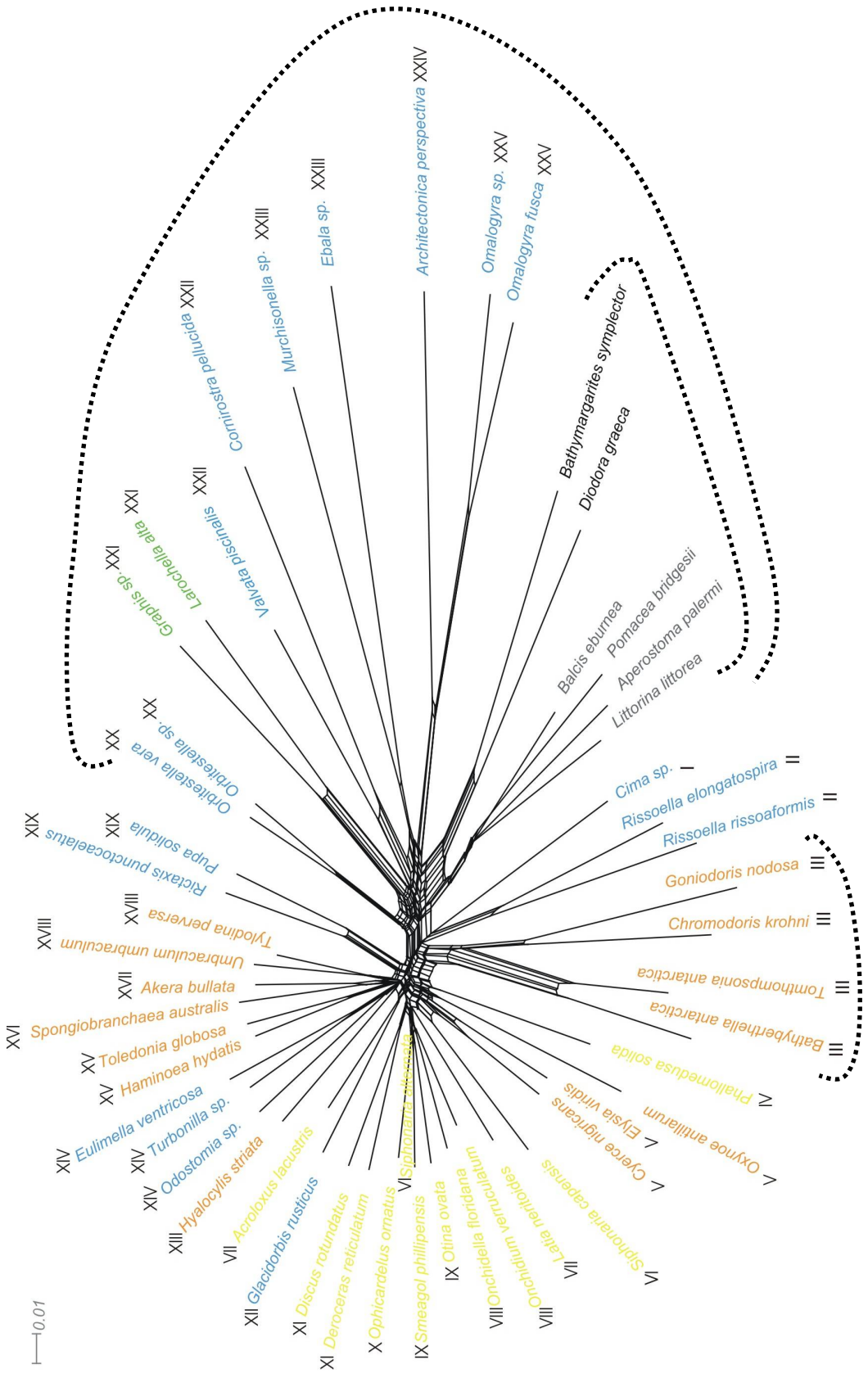


Fig. 2.8: Neighbournet graph of dataset 0; Heterobranchia are numbered in a clockwise direction from I to XXV; deep splits are marked with a dotted line; I: Cimidae, II: Rissoelloidea, III: Nudipleura, IV: Amphiboloidea, V: Sacoglossa, VI: Siphonarioidea, VII: Hygrophila, VIII: Onchidioidea, IX: Otinoidea, X: Ellobioidea, XI: Stylommatophora, XII: Glacidorboidea, XIII+XVI: Pteropoda , XIV: Pyramidellidae, XV: Cephalaspidea, XVII: Akeroidea, XVIII: Umbraculoidea, XIX: Acteonoidea, XX: Orbitestellidae, XXI: Aclididae, XXII: Valvatoidea, XXIII: Murchisonellidae, XXIV: Architectonicoidea, XXV: Omalogyroidea.

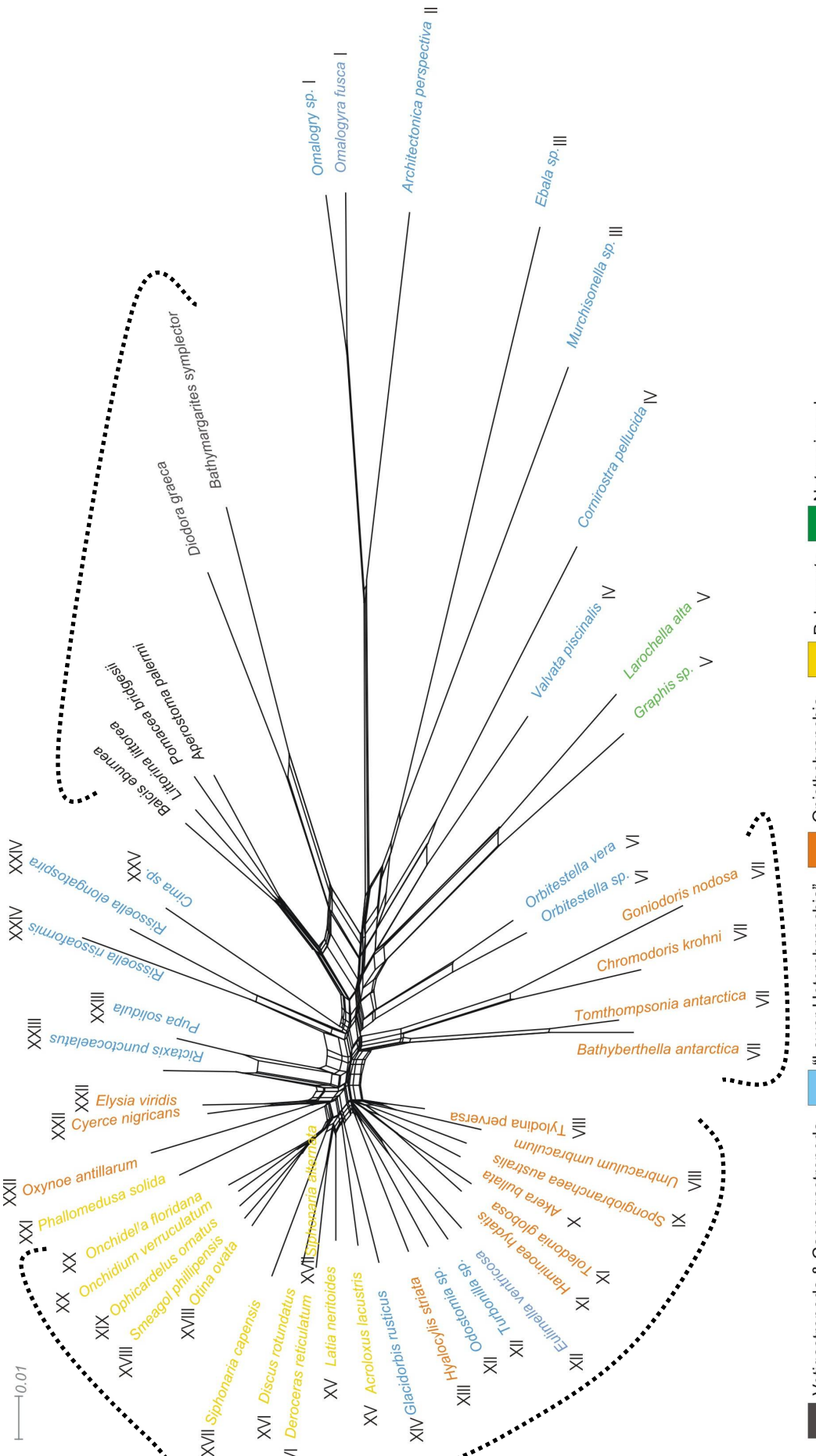


Fig. 2.9: Neighbournet graph of dataset I; Heterobranchia are numbered in a clockwise direction from I to XXV; deep splits are marked with a dotted line; I: Omalogyroidea, II: Architectonicoidea; III: Murchisonellidae; IV: Valvatoidea; V: Aclididae; VI: Orbitestellidae; VII: Nudipleura; VIII: Umbraculoidea; IX+XIII: Pteropoda; X: Akeroidea; XI: Cephalaspidea; XII: Pyramidellidae; XIV: Glacidorboidea; XV: Hygrophila; XVI: Stylommatophora; XVII: Siphonarioidea; XVIII: Otinoidea; XIX: Ellobioidea; XX: Onchidioidea; XXI: Amphiboloidea; XXII: Sacoglossa; XXIII: Acteonoidea, XXIV: Rissoelloidea; XXV: Cimidae .

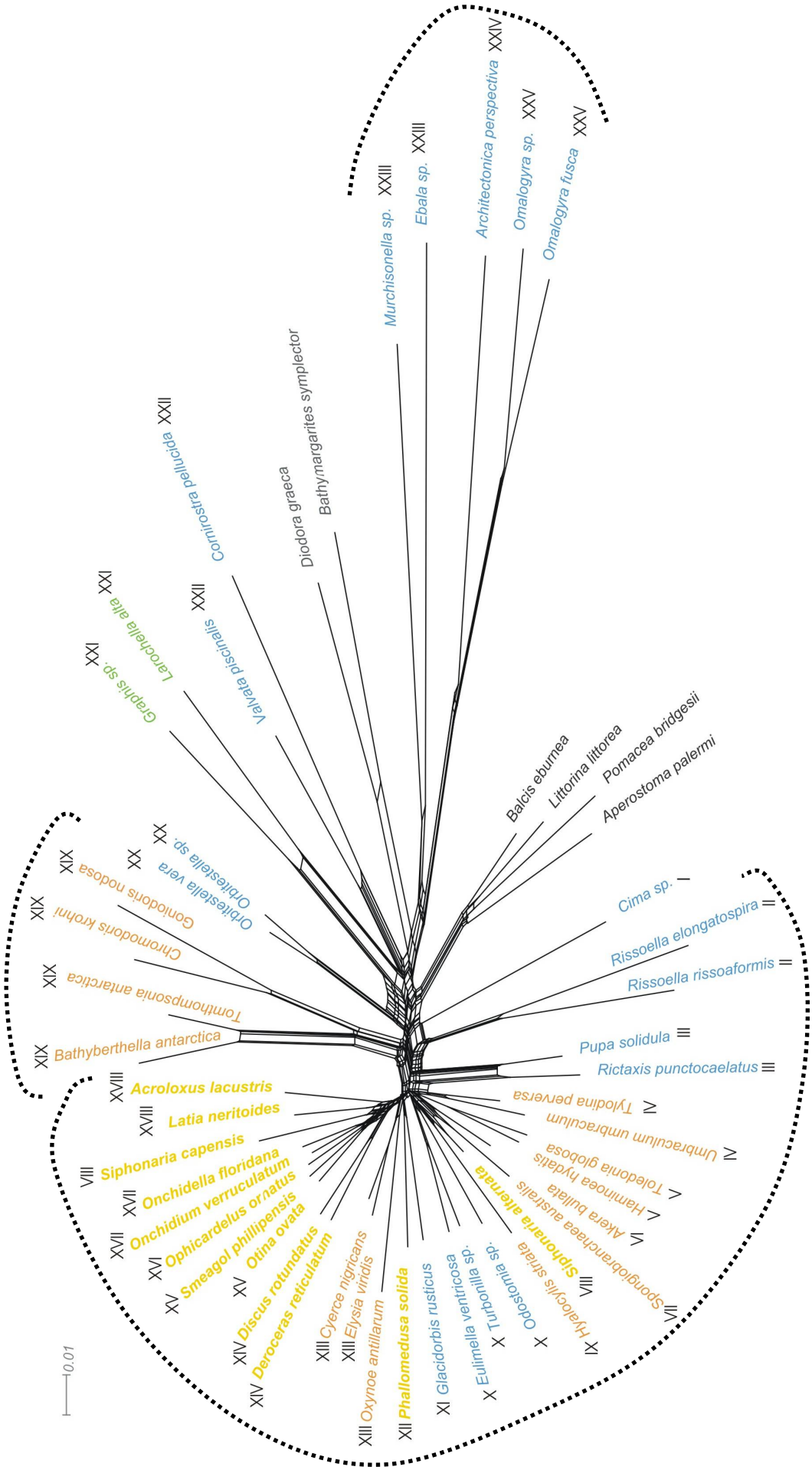


Fig. 2.10: Neighbournet graph of dataset II; Heterobranchia are numbered in a clockwise direction from I to XXV; deep splits are marked with a dotted line; I: Cimidae, II: Rissoelloidea, III: Acteonoidea, IV: Umbraculoidea, V: Cephalaspidea, VI:Akeroidea, VII+IX: Pteropoda, VIII: Siphonarioidea, X: Pyramidellidae, XI: Glacidorboidea, XII: Amphiboloidea, XIII: Saccoglossa, XIV: Stylommatophora, XV: Otinoidea, XVI: Ellobioidea, XVII: Onchidioidea, XVIII: Hygrophila, XIX: Nudipleura, XX: Orbitestellidae, XXI: Aclididae, XXII: Valvatoidea, XXIII: Murchsionellidae, XXIV: Architectonicoidea, XXV: Omalogvroidea.

2.3.4.2 SAMS

In spectral analyses a split is represented by a column in the spectrum graph. In addition, following the concept of Wägele & Rödding (1998) and Wägele & Mayer (2007) in each column three different types of positions can be discerned: binary support (red) = with only two character states (symmetric positions), noisy outgroup support (green) = one partition of the split with only one character state, the other with more than one state (asymmetric positions) and noisy in- and outgroup support (yellow) = more than one state in each partition (noisy positions).

The complete spectrum of the combined dataset 0 contained 5890 splits, of the combined dataset I 3808 splits and of the combined dataset II 3102 splits. There was little binary support (red), some noisy outgroup support (green) and many noisy in- and outgroup support (yellow) (see fig. 2.11) in all three datasets.

The most prominent split (supporting Omalogyroidea) was present in all three combined datasets while the second (supporting Architectonicoidea, Omalogyroidea and *Balcis eburnean* - Caenogastropoda) and third (supporting Architectonicoidea and Omalogyroidea) most prominent splits only occurred in datasets 0 and II (see tabs. 2.6 – 2.8). Within the 60 best splits one was able to find signals for partitions concerning mainly "Lower Heterobranchia" as well as Veti- and Caenogastropoda and some Opisthobranchia taxa. None of the deeper splits were found among the 60 best splits. However, only the first couple of splits of all three datasets were distinctly stronger, while the remaining spectral signals were not higher than background noise.

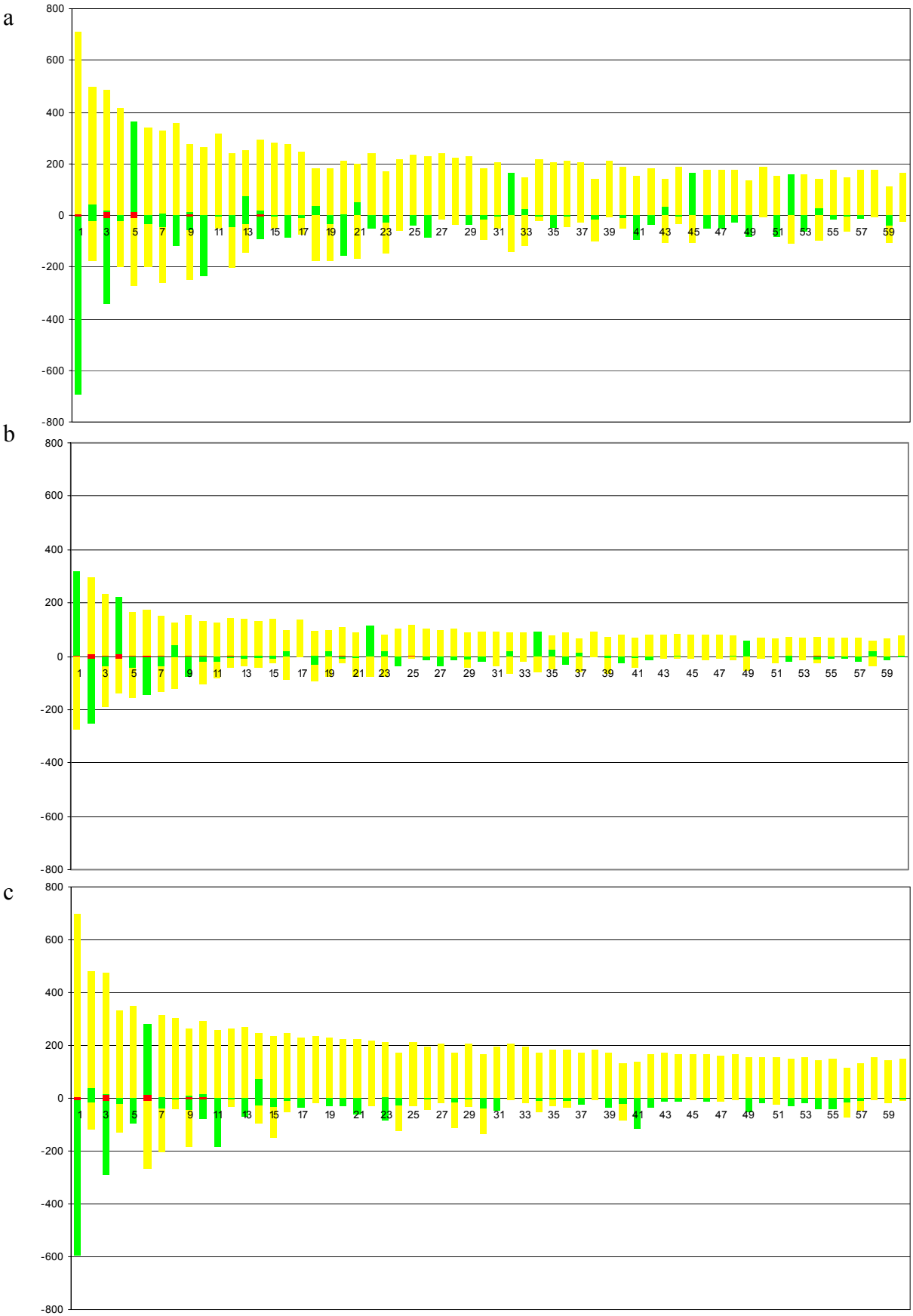


Fig. 2.11: Split support spectrum of the 60 most frequent partitions, x-axis = splits, y-axis = number of sequence positions, above x-axis = outgroup, below x-axis = ingroup, red = binary support, green = noisy outgroup support, yellow = noisy in- and outgroup support; a: dataset 0; b: dataset I; c: dataset II

Tab. 2.6: Split support values of dataset 0; Taxa names were coded in a four letter name using the first two letters of the genus and species name (for abbreviation see tab. A1 in the appendix).

Split No.	Split Taxa	Outgroup support			Ingroup support		
		binary support	noisy outgroup support	noisy in- and outgroup support	binary support	noisy outgroup support	noisy in- and outgroup support
1	(Omfu,Omsp)	3	0	708	0	694	0
2	(Arpe,Omfu,Omsp,Baeb)	0	38	460	0	23	150
3	(Arpe,Omfu,Omsp)	11	9	464	11	332	0
4	(Arpe,Omfu,Omsp,Pobr)	0	0	415	0	21	176
5	(Laal,Grsp)	10	353	0	10	0	266
6	(Arpe,Omfu,Omsp,Grsp)	0	2	339	0	36	162
7	(Arpe,Omfu,Omsp,Ebsp)	1	4	320	1	46	217
8	(Arpe,Omfu)	0	0	357	0	114	0
9	(Arpe,Omfu,Omsp,Musp)	3	8	263	3	55	192
10	(Basy,Digr)	1	0	262	1	233	0
11	(Cope,Omfu,Omsp,Grsp)	0	0	315	0	5	49
12	(Cope,Arpe,Omfu,Omsp)	1	0	241	1	48	158
13	(Arpe,Omfu,Omsp,Gono)	0	75	178	0	37	110
14	(Riel,Grsp)	4	16	271	4	92	0
15	(Omfu,Omsp,Musp,Baeb)	0	1	278	0	8	47
16	(Riel,Baan)	0	0	275	0	88	0
17	(Riel,Riri,Omfu,Omsp)	0	0	247	0	11	63
18	(Arpe,Riri,Omfu,Omsp)	1	32	148	1	0	177
19	(Arpe,Omfu,Omsp,Digr)	0	0	181	0	33	145
20	(Ebsp,Musp)	1	2	206	1	155	0
21	(Arpe,Omfu,Omsp,Basy)	0	53	144	0	0	168
22	(Arpe,Omsp)	0	0	238	0	50	0
23	(Arpe,Omfu,Omsp,Appa)	0	0	168	0	30	117
24	(Riel,Omfu,Omsp,Baan)	0	0	218	0	2	54
25	(Omsp,Grsp)	0	0	235	0	38	0
26	(Grsp,Baan)	0	0	225	0	87	0
27	(Arpe,Riel,Omfu,Baan)	0	0	240	0	1	14
28	(Cope,Omfu,Omsp,Ebsp)	0	1	222	0	2	35
29	(Omfu,Grsp)	0	0	229	0	33	0
30	(Ebsp,Musp,Laal,Grsp)	0	0	181	0	17	77
31	(Ebsp,Musp,Basy,Digr)	0	1	201	0	7	46
32	(Chkr,Gono)	0	162	0	0	0	139
33	(Omfu,Omsp,Ebsp,Musp)	0	24	122	0	0	119
34	(Cope,Omfu,Omsp,Baan)	0	0	215	0	3	22
35	(Omfu,Omsp,Grsp)	0	0	203	0	46	0
36	(Omfu,Omsp,Grsp,Baan)	0	0	208	0	5	42
37	(Cope,Riel,Laal,Grsp)	0	2	202	0	1	26
38	(Omfu,Omsp,Laal,Grsp)	0	0	139	0	20	81
39	(Riel,Omfu,Baan,Pobr)	0	0	211	0	1	7
40	(Arpe,Omfu,Omsp,Basy,Digr)	0	1	184	0	13	40
41	(Musp,Basy)	0	0	150	0	92	0
42	(Omsp,Musp)	0	0	180	0	33	0
43	(Arpe,Omfu,Omsp,Toan)	0	36	105	0	0	104
44	(Riel,Omfu,Grsp,Baan)	0	0	189	0	6	27
45	(Riel,Riri)	0	165	0	0	0	108
46	(Arpe,Musp)	0	0	173	0	51	0
47	(Riel,Omfu)	0	0	178	0	50	0
48	(Omfu,Musp)	0	0	173	0	27	0
49	(Musp,Digr)	0	0	135	0	83	0
50	(Arpe,Omfu,Omsp,Grsp,Baan)	0	0	187	0	1	6
51	(Riel,Oxan)	0	0	151	0	81	0
52	(Toan,Baan)	0	156	0	0	0	110
53	(Omfu,Omsp,Musp)	0	0	159	0	65	0
54	(Arpe,Omfu,Omsp,Chkr)	0	28	110	0	1	97
55	(Omfu,Baan)	0	0	173	0	19	0
56	(Riel,Omfu,Omsp,Musp)	0	0	146	0	8	56
57	(Riel,Grsp,Baan)	0	0	173	0	14	0
58	(Cope,Omfu,Omsp,Grsp,Baan)	0	0	173	0	1	7
59	(Toan,Baan,Chkr,Gono)	1	0	110	1	40	67
60	(Cope,Omfu,Omsp,Ebsp,Musp)	0	1	162	0	2	19

Tab. 2.7: Split support values of dataset I; Taxa names were coded in a four letter name using the first two letters of the genus and species name (for abbreviation see tab. A1 in the appendix).

Split No.	Taxa	Outgroup support			Ingroup support		
		binary support	noisy outgroup support	noisy in- and outgroup support	binary support	noisy outgroup support	noisy in- and outgroup support
1	(Omfu,Omsp),	3	318	0	3	0	269
2	(Arpe,Omfu,Omsp)	8	1	285	8	243	0
3	(Arpe,Omfu,Omsp,Ebsp)	1	4	228	1	37	153
4	(Laal,Grsp)	10	214	0	10	0	131
5	(Arpe,Omfu,Omsp,Musp)	1	1	164	1	42	110
6	(Basy,Digr)	1	0	173	1	142	0
7	(Cope,Arpe,Omfu,Omsp)	1	0	151	1	36	96
8	(Arpe,Omfu,Omsp,Basy)	0	42	83	0	0	124
9	(Ebsp,Musp)	1	1	152	1	74	0
10	(Arpe,Riri,Omfu,Omsp)	1	0	130	1	17	89
11	(Omfu,Omsp,Ebsp,Musp)	0	0	125	0	18	64
12	(Arpe,Omfu,Omsp,Ebsp,Musp)	1	2	139	1	7	37
13	(Ebsp,Musp,Laal,Grsp)	0	2	139	0	8	31
14	(Ebsp,Musp,Basy,Digr)	0	1	134	0	7	38
15	(Arpe,Omfu,Omsp,Basy,Digr)	0	1	140	0	6	20
16	(Arpe,Omfu,Omsp,Grsp)	0	21	78	0	0	89
17	(Omfu,Omsp,Ebsp,Basy,Digr)	0	0	136	0	1	4
18	(Arpe,Omfu,Omsp,Gono)	0	1	95	0	29	65
19	(Arpe,Omfu,Omsp,Sial)	0	20	81	0	1	75
20	(Arpe,Omfu,Omsp,Ebsp,Basy)	1	0	109	1	5	21
21	(Arpe,Omfu,Omsp,Euve)	0	0	86	0	10	66
22	(Toan,Baan)	0	114	0	0	0	74
23	(Arpe,Omfu,Omsp,Appa)	0	17	65	0	0	78
24	(Musp,Basy)	0	0	107	0	38	0
25	(Arpe,Omfu,Omsp,Ebsp,Basy,Digr)	1	2	114	1	2	6
26	(Omsp,Musp)	0	0	105	0	14	0
27	(Arpe,Omfu)	0	0	98	0	36	0
28	(Arpe,Musp)	0	0	106	0	16	0
29	(Omfu,Omsp,Laal,Grsp)	0	0	85	0	12	33
30	(Arpe,Omsp)	0	0	92	0	20	0
31	(Omfu,Omsp,Basy,Digr)	0	0	92	0	5	33
32	(Arpe,Omfu,Omsp,Chkr)	0	20	65	0	1	64
33	(Ebsp,Baeb,Basy,Digr)	0	0	89	0	1	18
34	(Riel,Riri)	0	91	0	0	0	60
35	(Vapi,Arpe,Omfu,Omsp)	0	23	55	0	0	50
36	(Omfu,Omsp,Musp)	0	0	89	0	32	0
37	(Arpe,Omfu,Omsp,Hyst)	0	15	48	0	0	59
38	(Cope,Omfu,Omsp,Musp,Digr)	0	0	94	0	1	1
39	(Arpe,Omfu,Omsp,Ripu)	0	5	64	0	6	60
40	(Ebsp,Laal)	0	0	84	0	28	0
41	(Riel,Riri,Omfu,Omsp)	0	0	71	0	7	36
42	(Omsp,Ebsp)	0	0	80	0	15	0
43	(Arpe,Omfu,Laal,Grsp)	0	0	83	0	1	8
44	(Ebsp,Laal,Grsp,Basy)	0	1	79	0	1	9
45	(Arpe,Omsp,Musp,Digr)	0	0	81	0	1	6
46	(Musp,Laal,Grsp,Digr)	0	0	84	0	1	15
47	(Arpe,Omfu,Omsp,Musp,Grsp)	0	0	80	0	1	10
48	(Arpe,Omfu,Omsp,Laal,Grsp)	0	2	74	0	5	10
49	(Chkr,Gono)	0	62	0	0	0	52
50	(Vapi,Omfu,Omsp,Ebsp,Musp)	0	0	73	0	1	6
51	(Omfu,Omsp,Musp,Grsp)	0	0	66	0	4	19
52	(Ebsp,Basy)	0	1	69	0	17	0
53	(Omfu,Omsp,Ebsp,Grsp)	0	0	70	0	2	14
54	(Toan,Baan,Chkr,Gono)	1	0	67	1	14	13
55	(Omsp,Basy)	0	0	70	0	11	0
56	(Omfu,Ebsp)	0	0	71	0	6	0
57	(Omfu,Omsp,Ebsp)	0	0	72	0	19	0
58	(Arpe,Omfu,Omsp,Lili)	0	19	43	0	0	38
59	(Omfu,Basy)	0	0	66	0	13	0
60	(Arpe,Omfu,Omsp,Ebsp,Musp,Basy,Digr)	0	2	73	0	0	1

Tab. 2.8: Split support values of dataset II; Taxa names were coded in a four letter name using the first two letters of the genus and species name (for abbreviation see tab. A1 in the appendix).

Split No.	Split Taxa	Outgroup support			Ingroup support		
		binary support	noisy outgroup support	noisy in- and outgroup support	binary support	noisy outgroup support	noisy in- and outgroup support
1	(Omfu,Omsp)	3	0	697	3	593	0
2	(Arpe,Omfu,Omsp,Baeb)	0	38	443	0	17	105
3	(Arpe,Omfu,Omsp)	11	9	453	11	282	0
4	(Arpe,Omfu,Omsp,Grsp)	0	2	328	0	25	107
5	(Arpe,Omfu)	0	0	349	0	96	0
6	(Laal,Grsp)	9	271	0	9	0	258
7	(Arpe,Omfu,Omsp,Ebsp)	1	4	309	1	37	165
8	(Cope,Omfu,Omsp,Grsp)	0	0	304	0	4	38
9	(Arpe,Omfu,Omsp,Musp)	3	8	253	3	42	138
10	(Riel,Grsp)	4	16	269	4	74	0
11	(Basy,Digr)	1	0	258	1	183	0
12	(Omfu,Omsp,Musp,Baeb)	0	1	264	0	5	31
13	(Riel,Baan)	0	0	270	0	67	0
14	(Arpe,Omfu,Omsp,Gono)	0	75	168	0	27	68
15	(Cope,Arpe,Omfu,Omsp)	1	0	235	1	34	114
16	(Riel,Riri,Omfu,Omsp)	0	0	243	0	9	44
17	(Arpe,Omsp)	0	0	231	0	36	0
18	(Arpe,Riel,Omfu,Baan)	0	0	234	0	1	15
19	(Omsp,Grsp)	0	0	229	0	30	0
20	(Omfu,Grsp)	0	0	222	0	30	0
21	(Grsp,Baan)	0	0	224	0	60	0
22	(Cope,Omfu,Omsp,Ebsp)	0	1	214	0	2	28
23	(Ebsp,Musp)	1	2	206	1	84	0
24	(Arpe,Omfu,Omsp,Digr)	0	0	174	0	27	96
25	(Riel,Omfu,Omsp,Baan)	0	0	211	0	2	29
26	(Ebsp,Musp,Basy,Digr)	0	1	196	0	7	39
27	(Cope,Omfu,Omsp,Baan)	0	0	207	0	2	16
28	(Arpe,Riri,Omfu,Omsp)	1	0	172	1	15	98
29	(Omfu,Omsp,Grsp,Baan)	0	0	204	0	4	33
30	(Arpe,Omfu,Omsp,Basy)	0	0	163	0	39	99
31	(Omfu,Omsp,Grsp)	0	0	195	0	43	0
32	(Riel,Omfu,Baan,Pobr)	0	0	203	0	1	5
33	(Cope,Riel,Laal,Grsp)	0	2	191	0	1	14
34	(Ebsp,Musp,Laal,Grsp)	0	0	174	0	14	39
35	(Riel,Omfu,Grsp,Baan)	0	0	184	0	6	24
36	(Arpe,Omfu,Omsp,Basy,Digr)	0	1	180	0	9	28
37	(Omsp,Musp)	0	0	174	0	23	0
38	(Arpe,Omfu,Omsp,Grsp,Baan)	0	0	182	0	1	5
39	(Riel,Omfu)	0	0	173	0	37	0
40	(Arpe,Omfu,Omsp,Ebsp,Musp)	1	0	128	1	19	63
41	(Chkr,Gono)	0	0	136	0	114	0
42	(Arpe,Musp)	0	0	166	0	34	0
43	(Riel,Grsp,Baan)	0	0	171	0	14	0
44	(Omfu,Musp)	0	0	168	0	10	0
45	(Cope,Omfu,Omsp,Grsp,Baan)	0	0	168	0	1	6
46	(Omfu,Baan)	0	0	168	0	14	0
47	(Cope,Omfu,Omsp,Ebsp,Musp)	0	1	161	0	1	12
48	(Omfu,Omsp,Ebsp,Basy,Digr)	0	0	165	0	1	4
49	(Omfu,Omsp,Musp)	0	0	153	0	49	0
50	(Riel,Omsp)	0	0	157	0	17	0
51	(Riel,Omsp,Grsp,Baan)	0	0	153	0	2	23
52	(Musp,Laal)	0	0	149	0	29	0
53	(Omsp,Baan)	0	0	153	0	15	0
54	(Musp,Basy)	0	0	145	0	42	0
55	(Riel,Oxan)	0	0	148	0	41	0
56	(Omfu,Omsp,Ebsp,Musp)	0	0	116	0	18	58
57	(Omfu,Omsp,Laal,Grsp)	0	0	132	0	14	31
58	(Riel,Omfu,Omsp,Baan,Pobr)	0	0	152	0	1	5
59	(Arpe,Omfu,Laal,Grsp)	0	0	144	0	2	13
60	(Riel,Riri,Chkr,Gono)	0	0	149	0	3	11

2.4 Discussion

The first aim of this chapter was to identify ambiguous nucleotide sites by visual judgement and with the software Aliscore while the second aim was to decide which of the three available datasets (see tab. A5 in the appendix) is the most appropriate one for the phylogenetic reconstruction of the Heterobranchia and to discuss the data quality of the elected dataset *a priori*.

The measurement of substitution saturation of the aligned nucleotide sequences with the method developed by Xia et al. (2003) showed that all markers with the exception of the second codon position of COI were saturated in all datasets. When comparing the Iss.data, dataset I showed the lowest Iss values indicating the least saturation (see tab. 2.2). However, when plotting patristic distances against distances obtained with different models of sequence evolution, no essential differences could be observed within all three datasets (see figs. 2.3 – 2.7). The inability of calculating the genetic distances of the third codon position of COI implied how improper this codon position is to reconstruct phylogenetic relationships of the Heterobranchia due to the large genetic distances which could also be attributed to chance.
The base composition was of little value to decide which dataset is the most appropriate one for phylogenetic reconstruction of the Heterobranchia because only little deviation of the mean G+C and A+T content was observed in all markers of all datasets (see tab. 2.3).
dataset 0 showed little homogeneity of base frequencies compared to datasets I and II while Dataset I showed the highest p-values in most of the markers (see tab. 2.4).
The relative rate test was also of little help to decide which dataset is the most promising one because the test revealed no significant differences in evolutionary rates between datasets 0, I and II (see tab. 2.5). The inability of the software K2WuLi to estimate the relative rates of the third codon position of COI in datasets 0 and II again implied how improper this codon position is for phylogenetic reconstruction at the taxonomic level investigated in the current study.
The SplitsTree analysis of all three datasets showed quite similar results regarding the superfamily level (see figs. 2.8 – 2.10). The supported splits of the combined datasets I and II are the same while the combined dataset 0 showed less supporting splits for the major heterobranch subgroups (according to the taxonomic classification by Bouchet & Rocroi 2005). Comparing datasets I and II with each other regarding the deep splits it was evident that only in dataset I the Vetigastropoda and Caenogastropoda are clearly distinguished from

the remaining Heterobranchia by various splits. Therefore, dataset I seems more eligible to reconstruct the phylogeny of the Heterobranchia (see figs. 2.9 and 2.10).
Because of very similar results in all three datasets the split network analyses with SAMS was little helpful for the decision which dataset should be used for phylogenetic reconstruction.

Taking all results of the *a priori* evaluation of chapter 2 into account I decided to use dataset I for further phylogenetic analyses of the Heterobranchia because it seemed to be the most informative one. According to the results of the statistical tests dataset I and II were more informative than dataset 0. Dataset I and II (except of little differences) were quite similar but taking also the split network analyses into account dataset I was more promising. Nevertheless, I think that Aliscore is a suitable tool to detect random similarity within sequence alignments and should be tested for other taxon samplings besides Heterobranchia.

In the following the quality of the elected dataset (dataset I) will be discussed.

As already mentioned, according to the test developed by Xia et al. (2003) only the second codon position of COI was not saturated. 18S rDNA, 28S rDNA, 16S rDNA and the first codon position of COI showed little saturation (see tab. 2.2). Taking also the graphical display of the saturation into account, all markers showed saturation at a value of 0.04 (see figs. 2.3 – 2.7 – dataset I). To exclude all saturated markers for further investigations would mean to exclude nearly all available data gathered in the current study for the phylogenetic reconstruction of the Heterobranchia. I therefore decided to use the complete dataset I to avoid the loss of phylogenetic signal at all taxonomic levels. Other authors, e.g., Thollesson (1999) demonstrated that even in a "fast" gene like 16S rRNA one can find a useful amount of variation for "higher-level" phylogenies despite the noise due to multiple substitutions. Yang (1998) even went a step further and proposed that the problem of saturation may have been exaggerated. He examined the effect of the evolutionary rate of a gene on the accuracy of phylogeny reconstruction by computer simulation. Yang found out that saturation occurs only at a much higher level of sequence divergence than was previously suggested and phylogenetic methods appear quite tolerant of multiple substitutions at the same site.

Generally, the proportion of A+T in a genome is rarely equal to the G+C proportion and different organisms exhibit different patterns of base composition variation, especially mitochondrial genomes are GC poor (Mooers & Holmes 2000). The current study supported

this assumption (see tab. 2.3 – dataset I) because the G+C and A+T content in the 18S rDNA was almost equal while the G+C content in the 28S rDNA was much higher than the A+T content. Within the mitochondrial 16S rDNA and COI one was able to observe a high A+T content compared to the G+C content.
Unequal base frequencies among species indicate that the substitution process is not homogeneous among lineages, as is commonly assumed in most phylogeny reconstruction methods. In such cases, tree reconstruction methods tend to group species with similar base content instead of common ancestries (Steel et al. 1993). Especially the heterogeneity of the G+C content of rDNA (35 – 70%) brought the first criticisms because it can lead to tree reconstruction artefacts (Hasegawa & Hashimoto 1993). In this study mentionable deviation of the mean G+C and A+T content was observed within taxa belonging to the Vetigastropoda (*Diodora graeca* and *Bathymargarites symplector*) and "Lower Heterobranchia" (*Omalogyra* sp., *Omalogyra fusca*, *Ebala* sp. and *Murchisonella* sp.) as well as Opisthobranchia (only *Chromodoris krohni*). These taxa should be viewed carefully when interpreting the phylogenetic hypotheses proposed in the current thesis.

The Chi-Square-Test revealed homogeneity of base frequencies of 28S rDNA, 16S rDNA and first and second codon position of COI while only 18S rDNA showed heterogeneity of base frequencies (see tab. 2.4 – dataset I).

The Relative-Rate-Test showed a significant difference in evolutionary rates between the genetic markers and investigated taxa (see tab. 2.5 – dataset I). The highest z-scores were found within 18S rDNA, the second highest within 28S rDNA, the third highest within COI and the lowest z-scores within 16S rDNA. The Omalogyroidea, Murchisonellidae and Architectonicoidea as well as *Larochella alta* (Aclididae) and *Hyalocylis striata* (Pteropoda) evolved faster than the remaining taxa indicated by high z-scores. These taxa should be considered carefully.

Investigating the deep splits calculated with SplitsTree one was able to see much conflict in the dataset (see fig. 2.9). There was a good split support separating the Veti- and Caenogastropoda from the Heterobranchia and also a good support for the Nudipleura (VII). This taxon was introduced by Wägele & Willan (2000) and comprises the Pleurobranchoidea and the Nudibranchia. Indicated by very long edges there was a split supporting the Omalogyroidea (I) and Architectonicoidea (II) as sister groups. In the literature one can find

evidence based on morphology data (e.g Haszprunar 1988) for a sister relationship of these groups. Another deep split could be observed within the neighbournet graph separating the Pyramidellidae (XII), Glacidorboidea (XIV), Pulmonata (with exception of Amphiboloidea XXI), Umbraculoidea (VIII), Akeroidea (X) and Pteropoda (IX+XIII) from the remaining taxa. No characters from the literature support this split. Anyway, there is little split support regarding the relationships of taxa within this deep split. This applies especially to the Pulmonata.

The network graph also showed that most of the major heterobranch subgroups were supported by splits and these groupings are in accordance with the latest taxonomic classification by Bouchet & Rocroi (2005) (see fig. 2.9 and tab. A1 in the appendix). This result indicated a good phylogenetic signal at superfamily level. Only Pyramidelloidea (III+XIV) (as well as Pyramidellidae XIV) and Siphonarioidea (XVII) were not supported by any split while *Siphonaria alternata* showed a conspicuous short terminal branch compared to the other taxa. Glacidorboidea (XIV) as well as Pyramidellidae (XIV) were separated from the remaining "Lower Heterobranchia" and nested somewhere between the Pulmonata and Opisthobranchia. This could be a hint for a relationship with the Euthyneura because the phylogenetic positions of both taxa were not clarified yet.

Larochella alta and *Graphis* sp. which have not been assigned to the Heterobranchia yet clustered within the "Lower Heterobranchia" suggesting a close relationship. Also noticeable is the position of the Sacoglossa (XXII). They were separated from the remaining Opisthobranchia and nested between the Pulmonata and „Lower Heterobranchia".

The spectrum graph evaluated with SAMS showed much conflicts because of the 60 best splits only 13 splits have a binary support and 28 splits a noisy outgroup support while most of the remaining splits showed a high noisy in- and outgroup support only (see fig. 2.11b – dataset I). Only the first couple of splits were distinctly stronger than the visible high background noise.

None of the deeper splits (above superfamily level) were found among the 60 best splits, only signals for groupings between taxa belonging to "Lower Heterobranchia", Veti- and Caenogastropoda and some Opisthobranchia were detected.

Conclusion

The following phylogenetic reconstruction based on the here *a priori* investigated dataset should be handled with care. Substitution saturation was observed in most of the alignment positions and the relative rate test revealed taxa with high evolutionary rates. Both split network analyses (with SplitsTree and SAMS) showed many conflicts, not within the terminal branches but the deep splits. Especially within the Pulmonata little signal could be determined. As a consequence of these results no Maximum parsimony analyses will be used for reconstructing heterobranch phylogeny but model based Maximum likelihood and Bayesian approaches. The evolutionary models should help to compensate the expected problems discussed above.

3. Phylogeny of the Heterobranchia

3.1 Introduction

The Gastropoda are the largest and most diverse class of the phylum Mollusca and exhibit the highest diversity in morphology. Many questions regarding gastropod phylogeny have not yet been answered. One major question is the molecular confirmation of the Heterobranchia concept based upon morphological studies by Haszprunar (1985a, 1988). This diverse taxon comprises the Pentaganglionata, also known as Euthyneura (with the Opisthobranchia and Pulmonata), and several less known "basal" groups such as Valvatoidea, Architectonicoidea, Omalogyroidea, Rissoelloidea and Pyramidelloidea. These lesser known "basal" groups supposedly present a step-by-step evolution towards the euthyneuran level of organisation (Haszprunar 1988). The systematic position and much disputed taxonomic history of these "Lower Heterobranchia" has been discussed in detail by Haszprunar (1985a, 1988), Bieler (1992) and Huber (1993).

The heterobranch clade is supported by numerous autapomorphies, including a pigmented mantel organ (which is reduced in more derived taxa), a medial position of the eyes in many taxa, a lack of a true ctenidium, a simple oesophagus, a distinctive sperm ultrastructure and most importantly a sinistral larval shell produced by a planktotrophic veliger (Haszprunar 1985a, Ponder & Linderberg 1997).

The monophyly of the Heterobranchia based upon morphological characters is confirmed by many authors (Haszprunar 1985a, 1988, Ponder & Lindberg 1997) (see fig. 3.1 a-b) while the monophyly of several currently recognised groupings within the Heterobranchia is equivocal.

The "Lower Heterobranchia" are clearly paraphyletic for which Haszprunar (1985a) was the first investigator to include many former "prosobranch-like" taxa in this informal group. Other authors followed, such as Ponder & Lindberg (1997), Dayrat & Tillier (2002) (see fig. 3.1 c) or Healy (1988; 1993).

The taxon Euthyneura includes the Pulmonata and Opisthobranchia, whose members secondarily reduce or revert the effects of torsion on the nervous system and other organ systems (Bieler 1992). Euthyneury itself, is convergently originated by detorsion, nerve concentration, or a combination of both (Haszprunar 1985a). Nevertheless, the monophyly of the Euthyneura based upon morphological data is generally accepted (Haszprunar 1988, Ponder & Lindberg 1997, Dayrat & Tillier 2002) and characterised by the presence of two additional (so-called parietal) ganglia on the visceral loop (Haszprunar 1985a; 1990). Within

the Euthyneura, the Opisthobranchia which share few if any obvious synapomorphies, may be paraphyletic (Haszprunar 1985b, Dayrat & Tillier 2002 and Wägele & Klussmann-Kolb 2005). Pulmonata have also been analysed morphologically and have mostly been recovered monophyletic (Tillier 1984, Haszprunar 1988, Haszprunar & Huber 1990, Nordsieck 1992, Dayrat & Tillier 2002).

Molecular analyses of heterobranch relationships demonstrate varying degrees of success in identifying the placement and monophyly of various groups within the Heterobranchia. This is largely due to the absence of molecular work including an adequate taxon sampling of all major heterobranch groups. Former investigations of the Gastropoda included only a few heterobranch taxa. They were rarely monophyletic in the analyses of Colgan et al. (2003) (including 9 heterobranch taxa) due to the variable position of the architectonicoid *Philippea*. Grande et al. (2008) (including 11 heterobranch taxa) (see fig. 3.1 d) found them to be monophyletic.

The lower heterobranchs have been neglected amongst the Heterobranchia because only few were included in phylogenetic analyses based upon molecular data. Until now, there is no comprehensive investigation concerning more than only a few representative taxa (e.g. Valvatoidea – *Cornirostra pellucida*, Architectonicoidea – *Philippea lutea*, Pyramidelloidea – *Pyramidella dolabrata*) (Colgan et al. 2000, Grande et al. 2004a; 2004b).

The monophyly of the Euthyneura has not yet been clarified via molecular studies. In some studies they are recovered monophyletic (Colgan et al. 2000; 2003, Knudsen et al. 2006) while in others, there are not (Thollesson 1999, Klussmann-Kolb et al. 2008). The molecular confirmation regarding the monophyly of the Opisthobranchia (Vonnemann et al. 2005, Dayrat et al. 2001, Grande et al. 2004a; 2008, Klussmann-Kolb et al. 2008 – see fig. 3.1 e, Wollscheid & Wägele, 1999) and the Pulmonata (Tillier et al. 1996, Winnepenninckx et al. 1998, Wade & Mordan 2000, Dayrat et al. 2001, Grande et al. 2004a; 2008, Klussmann-Kolb et al. 2008) is also still a matter of debate.

There appears good evidence in the literature that other minute snails such as the genera *Graphis* Jeffreys, 1867 and *Larochella* Powell, 1927 previously assigned to the Caenogastropoda, should perhaps also be integrated into the Heterobranchia (Ponder 1991). Little morphological and no molecular investigations have been undertaken within these groups.

In the present study a phylogenetic hypothesis of the Heterobranchia was inferred by using a multigene dataset including nuclear (28S rDNA + 18S rDNA) and mitochondrial (16S rDNA + COI) sequences. Phylogenetic trees were reconstructed using Maximum likelihood and Bayesian methodologies. *Bathymargarites symplector* (Vetigastropda) was defined as outgroup.

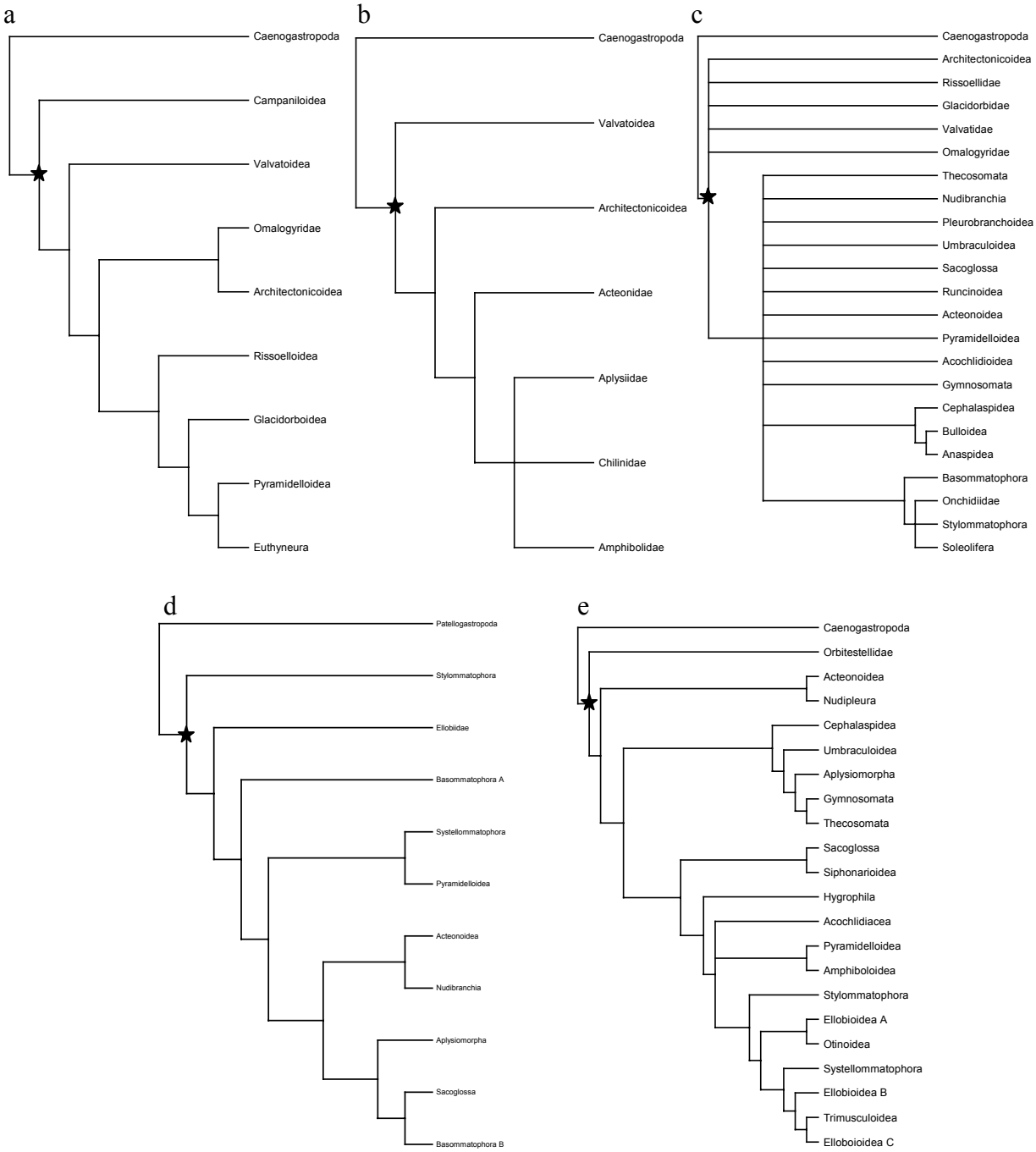


Fig. 3.1: Phylogenetic relationships among gastropods (Heterobranchia are marked by an asterisk): a) morphological data (Haszprunar 1988); b) morphological data (Ponder & Lindberg 1997); c) morphological data (Dayrat & Tillier 2002); d) molecular data (Grande et al. 2008); e) molecular data (Klussmann-Kolb et al. 2008).

The outstanding taxon sampling includes representatives of several groups, which have been poorly represented in earlier morphological studies and have never been included in molecular investigations (e.g. *Ebala, Murchisonella, Larochella*, *Graphis, Glacidorbis, Smeagol*).

3.2 Material and methods

Taxon sampling

(according to chapter 2.2)

DNA extraction, amplification and sequencing

(according to chapter 2.2)

Sequence editing and alignment

(according to chapter 2.2)

Maximum likelihood and Bayesian inference phylogenetic analyses were performed using dataset I (see tab. A5 in the appendix):

Dataset I = combination of complete 18S rDNA, partial 28S rDNA, partial 16S rDNA and COI sequences; long inserts and ambiguous alignment positions were excluded by visual judgement (see tab. A4 in the appendix).

Phylogenetic analyses

Maximum Likelihood (ML) methods seek to identify the most likely tree given the available data. An evolutionary model needs to be identified which estimates the probability of each possible individual nucleotide change.

ML analyses of sequences were carried out using the program RAxML 7.0.3. (Stamatakis 2006) adapting the program parameters to the alignment manually as recommended in the manual (the “hard & slow” way). RAxML is able to handle large datasets. It only implements GTR-based models of nucleotide substitution arguing with the idea that GTR is the most common and general model for DNA analysis. Hence, the GTRmixed model was used and 200 multiple inferences were executed on the original alignment. Bootstrapping was performed for 1.000 replicates.

Bayesian inference of phylogeny is based upon a quantity called the posterior probability distribution of trees, which is the probability of a tree conditioned on the observations. The core algorithm implemented in software packages like MrBayes (Huelsenbeck & Ronquist 2001) or Phase (http://www.bioinf.manchester.ac.uk/resources/phase/) is the Metropolis-Hastings Markov Chain Monte Carlo (MCMC). MCMC is a stochastic algorithm that produces sample-based estimates of a target distribution of choice.

Bayesian analyses were conducted with the program MrBayes 3.1.2 (Huelsenbeck & Ronquist 2001) for which 4 simultaneous Markov chains were run twice. Likelihood parameters were estimated separately for each gene (and each codon position within coding sequences) using a character partition. The analysis for each data set was run for 2.000.000 generations, with a sample frequency of 10. The first 20.000 generations were discarded as burnin. If likelihoods had not reached a plateau the burnin was increased to 40.000 generations.
Support for nodes is expressed as posterior probabilities.
The best-fit models of nucleotide substitution were selected with MrModeltest 2.2 (Nylander 2004) while choosing the Akaike information criterion (AIC) (see tab. A5 in the appendix).

3.3 Results

The Maximum likelihood and Bayesian analyses yielded similar results regarding phylogenetic relationships of subgroups within Heterobranchia however, with different statistical support which will be displayed in the following way: (posterior probability/bootstrap support). Only posterior probabilities of ≥ 0.95 and bootstrap support values of ≥ 75 respectively are statistically significant. Hence, support values below this significance level will not be discussed.

The reconstructed Bayesian 50% majority rule consensus tree based upon dataset I is shown in fig. 3.2. The Maximum likelihood tree is not shown, only the resultant bootstrap support values are plottet on the 50% majority rule consensus tree of the Bayesian analyses.

In general, it can be stated that both trees show a good resolution with high statistical support at the terminal branches while support for the deep nodes is sometimes nonexistent, particularly regarding bootstrap supports of the Maximum likelihood analyses.

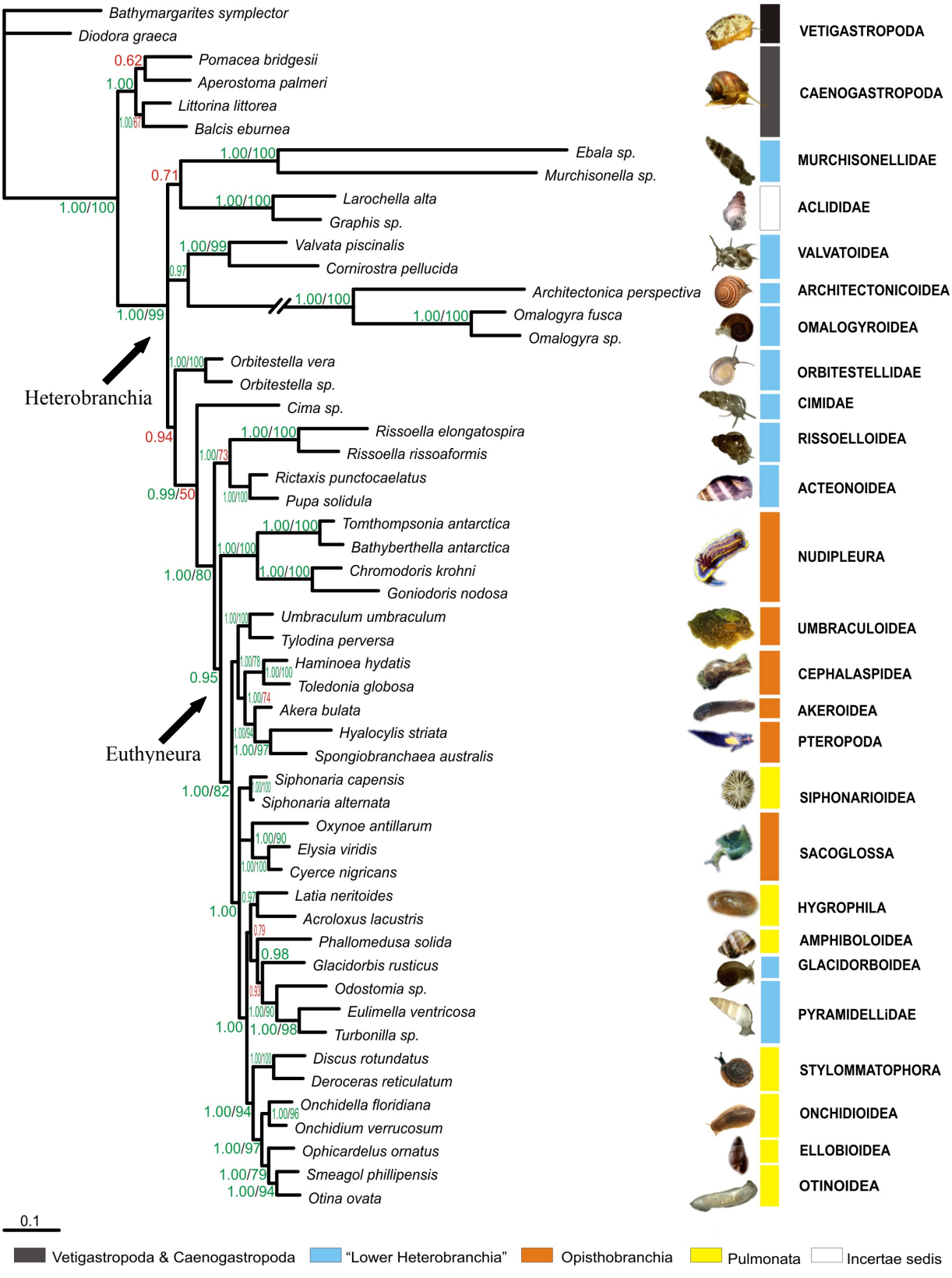


Fig. 3.2: Bayesian inference phylogram of the 50% majority rule consensus tree of dataset I; posterior probabilities as well as bootstrap support are provided at the branches (only supports above 0.5/50 are shown; green: statistically significant, red: statistically insignificant); taxonomic classification follows Bouchet & Rocroi (2005); the branch leading to the Architectonicoidea and Omalogyroidea was shortened due to a better presentability of the tree topology; for references of taxa images see tab. A6 in the appendix.

Extremely long terminal branches can be observed within the Murchisonellidae, Architectonicoidea and Omalogyroidea.

In both analyses, the Heterobranchia are monophyletic with good statistical support (1.00/99). Most of the "Lower Heterobranchia" are nested at the base of the Heterobranchia with the exception of the Glacidorboidea and Pyramidellidae. However, there is no resolution at the base of the Heterobranchia, thus there is no evidence which group/clade is the earliest offshoot.

Murchisonellidae is not the sister group of the Pyramidellidae which renders the Pyramidelloidea polyphyletic.

The two aclidids (*Larochella alta* and *Graphis* sp.) cluster within the Heterobranchia.

The next clade comprises the Valvatoidea (with *Valvata* and *Cornirostra*), Architectonicoidea and Omalogyroidea with a high posterior probability but no significant statistical bootstrap support (0.97/27). The Architectonicoidea are the sister group to the Omalogyroidea with a very good statistical support (1.00/100). The Orbitestellidae as well as the Cimidae appear as single offshoots and are in no sister group relationship to any other single taxon. The position of the Orbitestellidae within the system of the Heterobranchia has no statistical significant support (0.94/34) while the position of the Cimidae is supported by a high posterior probability but low bootstrap support (0.99/50).

The Rissoelloidea appear as sister to the Acteonoidea (1.00/73).

The Euthyneura are paraphyletic due to the inclusion of the Pyramidellidae and Glacidorboidea. The Pulmonata are also paraphyletic while the Opisthobranchia are polyphyletic. Within Euthyneura the Nudipleura (0.95/33) appear as the first single offshoot of the Opisthobranchia taxa.

The following Opisthobranchia clade comprises the Umbraculoidea, Cephalaspidea, Akeroidea and Pteropoda with a high statistical support (1.00/82). The phylogenetic position of the Siphonarioidea (Pulmonata) and the Sacoglossa (Opisthobranchia) remains unclear because of an unresolved tree topology. Another large clade comprises the pulmonate taxa Hygrophila and Amphiboloidea, the Glacidorboidea and Pyramidellidae as well as the monophyletic Eupulmonata (Stylommatophora, Onchidioidea, Ellobioidea and Otinoidea) with a high posterior probability and no significant statistical bootstrap support (1.00/41). The monophyly of the Basommatophora (Pulmonata) (comprising the Siphonarioidea, Hygrophila and Amphiboloidea) is rejected.

3.4 Discussion

Phylogenetic analyses

The current study represents the first molecular phylogeny of the Heterobranchia including representatives of most of the major taxonomic groups. A phylogenetic hypothesis was inferred by using a multigene dataset comprising nuclear (18S rDNA + 28S rDNA) and mitochondrial (16S rDNA + COI) sequences of the Heterobranchia. Because the phylogenetic software MrBayes allows the simultaneous application of more than one evolutionary model the analysis of this study was applied with different models for each genetic marker (18S, 28S and 16S) as well as for each codon position of COI (see tab. A5 in the appendix). Applying unique models to different regions of DNA takes the heterogeneous nature of DNA evolution into account while reducing systematic errors (Brandley et al. 2005), a practice which becomes more and more common in molecular phylogeny (e.g. Klussmann-Kolb & Dinapoli 2006, Grande et al. 2008, Klussmann-Kolb et al. 2008, Voigt et al. 2008). Unfortunately, RAxML does not have the feature to apply different evolutionary models for different gene partitions. Therefore, only one model (GTRmixed model) was used for the entire dataset I. It is obvious that in the present study, posterior probabilities (which were estimated with MrBayes) are generally higher than the bootstrap support (which was estimated with RAxML). At this point, it is difficult to infer if the better posterior probabilities are due to a sometimes overoptimistic estimation (which is a well known problem of Bayesian approaches) or due to a more realistic modelling of substitution rates by application of different models to the different partitions of the dataset.

Two Vetigastropoda and four Caenogastropoda were chosen as outgroup to infer phylogenetic relationships of the Heterobranchia. Caenogastropoda and Heterobranchia are considered to be sister taxa (Apogastropoda) based on morphological as well as molecular data (Ponder & Lindberg 1997, Colgan et al. 2003). In the recently published paper of Grande et al. (2008), a close relationship between the Veti- and Caenogastropoda was recovered in all analyses. Moreover, the Patellogastropoda were identified as sister to the Heterobranchia while rejecting the validity of the derived clade Apogastropoda. These are surely unexpected results and further investigations including morphological as well as molecular data are needed. Nevertheless, no Patellogastropoda was included in the present study because this would extend beyond the scope of this work but, should be taken into account for future studies.

Haszprunar introduced the concept of the Heterobranchia in 1985 and 1988 and included the Euthyneura as well as many "prosobranch-like" taxa which he grouped as Triganglionata Haszprunar, 1985 or Allogastropoda Haszprunar, 1985. The same paraphyletic grouping was referred to as Heterostropha by Golikov & Starobogatov (1975) and Ponder & Warén (1988). Bouchet & Rocroi (2005) introduced the informal group "Lower Heterobranchia" as a synonym to Allogastropoda.

The monophyly of Heterobranchia is widely accepted based on morphological characters (Haszprunar 1985a; 1988, Ponder & Lindberg 1997 and Dayrat & Tillier 2002) as well as molecular data (Grande et al. 2008). My results also reveal monophyletic Heterobranchia in all analyses and were always supported by very high statistical values (1.00/99). In the molecular analyses of Colgan et al. (2003) Heterobranchia were rarely monophyletic due to the variable position of *Philippea* (Architectonicoidea). However, blasting the 28S rDNA as well as the COI sequences of *Philippea* used by Colgan et al. (2003) (and deposited in Genbank) revealed a high affinity of these sequences to Arthropoda, Annelida and Bivalvia rather than Gastropoda. So, the reason for the position of *Philippea* outside of the Heterobranchia was probably due to contamination and not because of ancestral relationship.

Two Murchisonellidae (*Ebala* and *Murchsisonella*), also known as Ebalidae (Warén, 1994) or Anisocyclidae (Aartsen, 1995), as well as three Pyramidellidae (*Odostomia*, *Eulimella* and *Turbonilla*) were included in the present analyses. Traditionally, the murchisonellid taxa have been classified to the Pyramidellidae based upon shell characters only. However, Warén (1994) demonstrated that unlike the Pyramidellidae, the Murchisonellidae possess a complex jaw apparatus instead of a diagnostic buccal stylet (Wise 1996). These and other differences in their respective nervous systems (Huber 1993) and sperm morphology (Healy 1993) further support the separation of Murchisonellidae and Pyramidellidae. Therefore this family is considered to be a sister taxon of the Pyramidellidae of which bose comprise the Pyramidelloidea. Interestingly, in the present study the Murchisonellidae grouped at the base of the Heterobranchia, outside of the Euthyneura not forming a sister group relationship with the Pyramidellidae which are nested within the Euthyneura. This renders the Pyramidelloidea polyphyletic and supports the idea of a non basal position of the Pyramidellidae. As already mentioned, the relationship between the Murchisonellidae and the Pyramidellidae previously has been based mainly upon shell morphology. Huber (1993) investigated the cerebral nervous system of marine Heterobranchia and also included pyramidellid taxa (like *Odostomia*, *Boonea*, *Turbonilla* and *Pyramidella*) as well as a murchisonellid taxon (*Ebala*)

in his studies. He discovered significant differences in the nervous system of *Ebala* and the remaining pyramidellids (e.g. no rhinophores and lateral nerves are present in *Ebala*). Huber (1993) as well as the author of the present study call for further investigations of this group. In particular anatomical as well as histological studies are needed to shed light on the taxonomic position of the Murchisonellidae. Without any statistical support, they show an affinity to *Larochella* and *Graphis* in this study.

As already mentioned, the systematic position of many gastropod taxa has not yet been clarified. Groups traditionally assigned to the "Prosobranchia" are now considered to be "primitive" Heterobranchia such as Valvatoidea or Rissoelloidea (Haszprunar 1988, Healy 1993). Two genera, which had an unsteady taxonomic position in the past, are *Graphis* and *Larochella*. Fretter & Graham (1982) noted that the taxonomic position of *Graphis* remains unclear because there were reasons for doubting the placement within the family Aclididae. These doubts were mainly based on shell morphology and the presence of a pallial tentacle which separates *Graphis* from *Aclis*. Ponder (1984) remarked that the examination of the shell confirmed that *Larochella* (together with *Graphis*) should be placed in a new group near the Aclididae, concluding (1991) that these two genera have a pigmented mantle gland and belong to the heterobranch gastropods. Based on the molecular data of the present study, I support the opinion of Fretter and Ponder and suggest the inclusion of these two genera in the informal group "Lower Heterobranchia".

The next clade comprises the Valvatoidea, Architectonicoidea and Omalogyroidea. The taxon Valvatoidea is of special interest because some taxa occur in freshwater (e.g. Valvatidae) as well as marine habitats (e.g. Cornirostridae, Hyalogyrinidae). In addition, their early offshoot within the heterobranch clade probably provides information on what the early heterobranchs were like (e.g. morphology, life-style) (Ponder 1991). The taxonomic classification of the Valvatoidea is still under review (Ponder 1990a). According to the taxonomic classification of Bouchet & Rocroi (2005) the taxon consists of at least three recent families (Valvatidae, Cornirostridae and Hyalogyrinidae). Since they have several features in common (e.g. similar nervous system, similar sperm morphology), a close relationship of the Orbitestellidae with the Valvatidae was suggested by Ponder & Warén (1988), Ponder (1990b) and Healy (1993). On the other hand, Bieler at al. (1998) surveyed distinguishing characters of different Valvatoidea families and found Cornirostridae significantly different from Orbitestellidae. So, according to Ponder (1990b) the question arose whether the Orbitestellidae and Valvatidae are

similar because both groups are primitive, or whether they are part of the same monophyletic group at the base of the Heterobranchia. To answer this question I included Valvatidae, Cornirostridae and Orbitestellidae in my molecular analyses. Fig. 3.2 shows a sister group relationship between *Valvata* (Valvatidae) and *Cornirostra* (Cornirostridae). Together, they are sister to a clade comprising the Architectonicoidea and Omalogyroidea. The Orbitestellidae appear not to be related to the Valvatoidea, supporting the latest taxonomic classification by Bouchet & Rocroi (2005), where the Orbitestellidae are not yet assigned to any superfamily. Nevertheless, the position of the Orbitestellidae in the system of the Heterobranchia remains unclear because of a non-significant statistical support at the respective node (fig. 3.2) in the present study.

The Omalogyroidea as well as the Architectonicoidea are considered to be basal taxa of the Heterobranchia (Haszprunar 1985a; 1988). However, their exact systematic affinities, in particular the supposed close relationship between these two taxa is still unclear. Healy (1988) mentioned a possible affinity between Omalogyroidea and Architectonicoidea based on sperm morphology. Haszprunar (1985a; 1988) proposed the same relationship based on similarities in the mantle cavity and in the genital system. The results of the present study, based on molecular data, support these assumptions and propose a close relationship between Architectonicoidea and Omalogyroidea with a high statistical support (1.00/100) in all analyses. Nevertheless, recently Bäumler et al. (2008) studied the anatomy of the taxon *Omalogyra atomus* using 3D reconstructions and postulated that a closer relationship with the Arcitectonicoidea is not likely due to an erroneously stated (Haszprunar 1988) left-side position rather than a right-side position of the ciliary stripes (which replace the ctenidial function).

Another enigmatic taxon with a formerly unclear taxonomic position is *Cima*. Fretter & Graham (1982) concluded that *Cima* belongs to the Aclididae (Caenogastropoda). However, morphological characters such as shellmorphology, the ptenoglossate radula, absence of penis and the anterior edge of the foot being obtusely rounded in the Aclididae are different from other Aclididae. Warén (1993) assigned *Cima* to the “Lower Heterobranchia” based on the following characters: presence of a pigmented mantle organ, anteriorly deeply bifurcated foot, brownish digestive gland with darker granulae, and eyes situated centrally at the bases of the cephalic tentacles. He introduced a new family Cimidae. At this time, it was not possible for Warén (1993) to recognize any described family of the Heterobranchia to which the Cimidae

show a clear affinity. He therefore did not place this taxon in any superfamily. My results support the inclusion of the Cimidae in the Heterobranchia as well as the isolated position as sister group to Rissoelloidea + Acteonoidea + Euthyneura (incl. Pyramidelloidea).

Relative to the "Lower Heterobranchia" the most striking result of this new phylogeny is the sister group relationship between the Rissoelloidea and Acteonoidea, a sister group relationship which has never been proposed before. This could be due to the fact that previous studies on Heterobranchia (Gastropoda) phylogeny did not include both taxa in the same analyses. Nevertheless, in the current study bootstrap support for this sister group relationship was low (73) and synapomorphies (e.g. from morphology) to support this relationship are missing. Further testing of this phylogenetic hypothesis with additional morphological, ultrastructural and molecular data is urgently needed.
While the basal position of the Rissoelloidea within the Heterobranchia is supported by various authors (e.g. Haszprunar 1988, Healy 1993), the position of Acteonoidea is still unresolved. Traditionally, the latter have been regarded as opisthobranchs mostly with a basal position (Ponder & Lindberg 1997, Dayrat et al. 2001, Grande et al. 2004a and Klussmann-Kolb et al. 2008) or have even been excluded from Opisthobranchia (Mikkelsen 1996; 2002, Thollesson 1999, Bouchet & Rocroi 2005 and Wägele & Klussmann-Kolb 2005) (for a short review about the phylogenetic position of Acteonoidea see Mikkelsen 2002). My results support the latter authors because of the aforementioned sister group relationship of Rissoelloidea and Acteonoidea and support a taxonomic position of the Acteonoidea within the Heterobranchia but outside the Euthyneura.

As already mentioned Opisthobranchia and Pulmonata together comprise the Euthyneura, which have been accepted as monophyletic in most phylogenetic investigations (Nordsieck 1992, Tillier et al. 1994, Salvini-Plawen & Steiner 1996, Ponder & Lindberg 1997, Wade & Mordan 2000, Yoon & Kim 2000, Dayrat & Tillier 2002 and Knudsen et al. 2006). Nevertheless, the Euthyneura have been recovered paraphyletic in the current study because of the inclusion of two "Lower Heterobranchia" taxa (Pyramidellidae and Glacidorboidea). Other authors obtained similar results when including the Pyramidellidae in their molecular analyses. Grande et al. (2004a) and (2008) found Pyramidellidae nested deeply within Pulmonata whereas Klussmann-Kolb et al. (2008) recovered them as sister group to the Amphiboloidea (Pulmonata). Unlike these studies, which included only one Pyramidellidae in the analyses, three pyramidellid taxa were included in the current investigation (*Odostomia*,

Eulimella and *Turbonilla*). The Pyramidellidae nested within Pulmonata in all my analyses forming a clade with the Amphiboloidea and Glacidorboidea.

The Opisthobranchia show only few apomorphies, like the bifurcation of the Nervus labiotentacularis (N2) (Salvini-Plawen & Steiner 1996) but current studies of Staubach (2008) detected a homologous cerebral nerve (Nervus tentacularis) in *Achatina fulica* (Pulmonata) which is also bifurcated. Hence, the apomorphic character of this nerv for the Opisthobranchia is questionable. Nevertheless, the Opisthobranchia have also been recovered paraphyletic regardless of whether the analyses were based upon morphological or molecular data (Haszprunar 1988, Ponder & Lindberg 1997, Thollesson 1999, Wägele et al. 2003, Grande et al. 2004a; 2008, Vonnemann et al. 2005, Wägele & Klussmann-Kolb 2005 and Klussmann-Kolb et al. 2008). In all current analyses Opisthobranchia was never recovered monophyletic.

In the current phylogeny the first offshoot within Euthyneura respectively Opisthobranchia are the Nudipleura (see fig. 3.2). This is remarkable because the Nudipleura are usually regarded as highly derived (Wägele & Willan 2000, Wägele et al. 2003 and Wägele & Klussmann-Kolb 2005). This taxon was introduced by Wägele & Willan (2000), which comprises the Pleurobranchoidea and the Nudibranchia, both characterised by the possession of a blood gland, an androdiaulic reproductive system and the loss of the osphradium. Grande et al. (2004a), Vonnemann et al. (2005) and Klussmann-Kolb et al. (2008) found Acteonoidea to be the sister group of Nudipleura. It is noticeable that, Klussmann-Kolb et al. (2008) observed deviant base composition and rate heterogeneity in Nudipleura which could consequently lead to an artificial basal position in a molecular tree. Hence, evolutionary models which have the ability to compensate rate heterogeneity are urgently needed. A PhD student (Karen Meusemann) from the Forschungsmuseum Alexander Koenig in Bonn, who is working on this problem, has most likely found a possibility to solve the problem of high rate heterogeneity by modifying models of the software package called Phase. However, these models are currently not publicly available.
Nevertheless, a basal position of the highly derived Nudipleura seems to be unlikely, underscoring the fact that further investigations are indeed necessary. Currently, in the Department of Phylogeny and Systematics, a PhD study is being conducted, which is investigating molecular as well as sperm morphology of the Nudipleura. This investigation intends to shed more light on the systematic position of this enigmatic group.

The next clade comprised the Umbraculoidea, Cephalaspidea, Akeroidea (Aplysiomorpha) and Pteropoda with a good statistical support (1.00/78). The same clade was revealed by the study of Klussmann-Kolb et al. (2008) with a high posterior probability (1.00) as well as bootstrap support (99). Klussmann-Kolb et al. (2008) only included one umbraculoid taxon and considered taking more taxa into account for further phylogenetic studies in order to clarify the position of this group within the heterobranch system. Additionally, I included *Tylodina* in my analyses, which did not change the topology of the clade comprising the Umbraculoidea, Cephalaspidea, Akeroidea and Pteropoda. Dayrat et al. (2001) and Klussmann-Kolb & Dinapoli (2006) found a strongly supported monophyly of a group comprising the Aplysiomorpha (including Akeroidea) and Pteropoda. The latter authors also discussed the stomach caecum with a typhlosolis and specialized glandular epithelium as a possible synapomorphy of these two taxa. Due to the adaptation to a pelagic life style, the relationship of the pteropods with other opisthobranchs has been difficult to reveal but Klussmann-Kolb & Dinapoli (2006) considered this caecum to be homologous in Thecosomata and Akeroidea (Aplysiomorpha). Their molecular analyses also showed a support for the hypothesis of Cephalaspidea being the sister group of Aplysiomorpha + Pteropoda but with rather low bootstrap support. However, histological investigations of gizzard plates revealed that these plates are built by a very similar epithelium as seen in the representatives of all three taxa (Klussmann-Kolb & Dinapoli 2006). Another feature probably uniting the Pteropoda, Aplysiomorpha and Cephalaspidea are the parapodia. However, the homology of these structures in the three taxa has not yet been clarified, rendering a discussion at this point inappropriate.

The systematic position of the Sacoglossa as well as the Siphonarioidea is still a matter of debate (Grande et al. 2004a; 2004b, Vonnemann et al. 2005 and Klussmann-Kolb et al. 2008). Unfortunately, according to my results, any statement about the phylogeny of both taxa is impossible because my analyses reveal an unresolved tree topology at the position of the Siphonarioidea and Sacoglossa. This topology presents both taxa outside a clade comprising the remaining Pulmonata (Hygrophila, Amphiboloidea, Stylomatophora, Onchidioidea, Ellobioidea, Otinoidea) + Glacidorboidea + Pyramidellidae.

Different molecular analyses assign the Sacoglossa and Siphonarioidea equivocally to different clades within Euthyneura. Dayrat et al. (2001) found Sacoglossa to be basal within the Euthyneura. According to Grande et al. (2004b), they are basal but sister to *Siphonaria* and the remaining Opisthobranchia. The molecular data of Klussmann-Kolb et al. (2008)

suggested a close affinity of the Sacoglossa to primitive Pulmonata, especially Siphonarioidea. In Remigio & Hebert (2003) the marine basommatophoran *Siphonaria* did not group with other members of this order or even with pulmonates but appears as an offshoot within the Opisthobranchia. All analyses of Grande et al. (2008) also support *Siphonaria* as an opisthobranch.

Because of the variable position of the Siphonarioidea and Sacoglossa within the system of the Heterobranchia respectively Euthyneura a re-evaluation of morphological as well as molecular data is urgently needed.

In contrast to the Opisthobranchia, the monophyly of Pulmonata is widely accepted based on morphological characters (Tillier 1984, Haszprunar 1985a; 1990, Nordsieck 1992 and Dayrat & Tillier 2002). However, paraphyly or even polyphyly of Pulmonata was recovered using molecular data (Tillier et al. 1996, Grande et al. 2004a; 2008, Knudsen et al. 2006 and Klussmann-Kolb et al. 2008). The molecular results of the present study also support the idea of a paraphyletic Pulmonata.

The phylogeny of the Pulmonata has been discussed controversially over the years (Tillier 1984, Haszprunar & Huber 1990, Nordsieck 1992, Salvini-Plawen & Steiner 1996, Barker 2001, Dayrat et al. 2001, Dayrat & Tillier 2002, Grande et al. 2004a; 2008, Wade et al. 2006 and Klussmann-Kolb et al. 2008).

Within the present study, the monophyly of the Hygrophila is confirmed. The hypothesis of a common origin of the freshwater taxa belonging to the Hygrophila was supported earlier by morphological (Salvini-Plawen 1990, Nordsieck 1992, Barker 2001) and molecular studies (Dayrat et al. 2001, Albrecht 2005, Klussmann-Kolb et al. 2008). However, only two Hygrophila taxa were included in the current study. Hence, the result is of little significance.

The monophyly of the Eupulmonata is also confirmed whereas, neither the monophyly of the Basommatophora (Siphonarioidea, Hygrophila and Amphiboloidea) nor the monophyly of the Thalassophila (= Amphiboloidea + Siphonarioidea) is supported. Recent molecular studies also failed to recover Basommatophora as a monophyletic group (Tillier et al. 1996, Yoon & Kim 2000, Grande et al. 2004a; 2008 and Klussmann-Kolb et al. 2008).

The superfamily Glacidorboidea and the family Glacidorbidae were introduced by Ponder (1986) and placed within the Basommatophora based on several morphological characters (e.g. similarity of the genital system, dorsal and ventral jaw elements, sperm morphology, euthyneurous nervous system). Ponder (1986) also proposed a possible relationship between the freshwater genus *Glacidorbis* and the Amphiboloidea. A pulmonate relationship was also

accepted by Visser (1988) and Healy (1996) based on sperm ultrastructural data. In contrast to this view, Haszprunar (1988) argued that Glacidorbidae are not related to Pulmonata and should be placed within the "Lower Heterobranchia" due to the lack of a typical pentaganglionate nervous system, a pneumostome, a procerebrum and dorsal bodies. Barker (2001) and Dayrat & Tillier (2002) followed this opinion. In 2000 Ponder & Avern considered that a pulmonate relationship is still possible because *Glacidoris* is highly paedomorphic, which would explain the absence of many of the typical pulmonate characters. Based on my molecular data I follow the opinion of Ponder (1986) and Ponder & Avern (2000). We agree upon the pulmonate relationship suggesting a closer relationship to the Amphiboloidea.

The systematic position of the Pyramidellidae within the gastropod system has been discussed controversially for over 130 years. This controversy is caused in part by the lack of information about this taxon but also due to changing views about gastropod phylogeny (Wise 1996). Based on morphological characters, older studies placed them in the "Prosobranchia" because of a spirally coiled calcareous shell into which the entire body is retractable, a foot with an operculum, a long proboscis and an anteriorly oriented mantle cavity (e.g. Golikov & Starobogatav 1975). Younger studies placed them in the Opisthobranchia because of a pallial kidney, subepithelial eyes on the median side of the tentacles, an ovotestis and a heterostrophic protoconch (e.g. Salvini-Plawen 1980). At present, many scientists assign them to the "Lower Heterobranchia" (e.g. Haszprunar 1985a; 1988, Ponder & Warén 1988) (for a short review about the current state of pyramidellid phylogeny see also Wise 1996). Haszprunar (1990) discussed synapomorphies of high significance, such as giant nerve cells, a rhinophoral and a lateral nerve and characters of sperm morphology possibly shared by Euthyneura and Pyramidelloidea. However, he placed the Pyramidelloidea closest to the Euthyneura but still outside of the latter. Huber (1993) investigated the cerebral nervous system of marine Heterobranchia and included also a remarkable number of "Lower Heterobranchia". According to Huber (1993), rhinophoral and lateral nerves are present in the Pyramidellidae and the Opisthobranchia but absent in the Caenogastropoda, Architectonicoidea, Omalogyroidea and Rissoelloidea. Moreover, he observed giant cerebral nerve cells in *Amathina tricarinata* (Amathinidae, Pyramidelloidea) but not in small Pyramidellidae like *Pyramidella* or *Odostomia*. Therefore, he favoured the idea that the Pyramidellidae have an intermediate position between the "Lower Heterobranchia" and the Opisthobranchia. According to the results of the present study, I propose a different scenario.

The Pyramidellidae should be included within the Euthyneura based on the newly acquired molecular data (see fig. 3.2). This assumption is supported by morphological data of the nervous system. Rhinophoral and lateral nerves are present in the Pyramidellidae (Huber 1993) while the additional pair of ganglia (which is an autapomorphy for the Euthyneura) have possibly been lost due to small body size or the parasitic life style of the Pyramidellidae. These aspects are often associated with enormous morphological changes. Another hint is the presence of giant nerve cells in *Amathina tricarinata* (Amathinidae, Pyramidelloidea). These particular cells occurred only within Euthyneura and are linked to body size and therefore possibly not present in the minute Pyramidellidae. Anyway, the inclusion of Amatinidae taxa in further phylogenetic analyses could be the key to the answer of the phylogenetic position of the Pyramidellidae. If *Amathina* clustered with the Pyramidellidae within the Euthyneura, then one could assume that the giant cells are secondarily lost in the minute Pyramidellidae. However, if *Amathina* clustered with the Murchisonellidae outside the Euthyneura, then an earlier occurrence of giant cells in the evolution would have to be proposed.
According to the aforementioned molecular data, the Pyramidellidae show Pulmonata affinity. Morphological characters which support these results are lacking to date. Thollesson (1999) investigated the phylogeny of the Euthyneura based on molecular data and found an apomorphic deletion (gap of ca. 20 bp) in the helix G16 of the 16S rRNA molecule. This gap was also found in the Pyramidellidae, supporting not a close relationship with the Pulmonata but rather a possible synapomorphy of Pyramidellidae and Euthyneura. A recent diploma thesis in our working group, took additional pyramidellid taxa into account. The results support this assumption. The Pyramidellidae nested within the Euthyneura in all analyses but neither a closer relationship with the Pulmonata nor with the Opisthobranchia could be concluded in this diploma study (Zinßmeister 2008).

According to the latest review by Bouchet & Rocroi (2005), the clade Eupulmonata comprises the Stylommatophora + Onchidioidea + Ellobioidea + Otinoidea + Trimusculoidea. Although, there are no morphological apomorphies known to date, this taxon receives good support in molecular studies (Wade & Mordan 2000, Klussmann-Kolb et al. 2008, current study). Monophyly of Stylommatophora within the Euthyneura is strongly supported by Nordsieck (1992), Tillier et al. (1996), Wade & Mordan (2000), Dayrat et al. (2001), Dayrat & Tillier (2002), Grande et al. (2004a; 2008), Wade et al. (2006), Klussmann-Kolb et al. (2008) and the present study. As sister to the Stylommatophora, the present study shows a clade comprising the Onchidioidea, Ellobioidea and Otinoidea with high statistical support (1.00/97). These

results are congruent to former molecular studies of Klussmann-Kolb et al. (2008) who discussed the phylogeny of the Eupulmonata in detail.

The taxon *Smeagol* (Smeagolida) was introduced by Climo (1980) who assigned them to the Gymnomorpha (syn. Systellommatophora). Tillier (1984) argued that *Smeagol* is related to the Otininidae and has undergone modification due to "limicization" (e.g. loss of shell). Tillier & Ponder (1992) reinvestigated *Smeagol* and came to the same conclusion that (based upon the synapomorphic occurrence of an ocular ridge and arrangement of the heart and kidney, together with the probable symplesiomorphic foot morphology) *Otina* and *Smeagol* form a monophyletic group. Barker (2001) followed this opinion, whereas Haszprunar & Huber (1990) assumed a close relationship between *Smeagol* and the Onchidioidea based upon the nervous system. Nordsieck (1992) followed the opinion of the latter one. To prove or reject both phylogenetic hypotheses, based on morphological data, *Otina* as well as *Smeagol*, were included in the present molecular study. According to the results (fig. 3.2), the author favours the hypothesis that *Otina* and *Smeagol* form a monophyletic group since these two taxa appear as sister taxa in all analyses with a high statistical support (1.00/94).

Conclusion

Although, the *a priori* analyses have revealed much conflict in the dataset, the here presented molecular hypotheses show new insights into heterobranch phylogeny mainly due to the outstanding taxon sampling of the "Lower Heterobranchia". Moreover, this is the first analysis comprising a multigene dataset of two molecular and two mitochondrial genes (about 4000 bp) in representatives of all major lineages of Heterobranchia. Species like *Ebala*, *Murchisonella, Glacidorbis* or *Smeagol* were included for the first time in a molecular analysis of the Gastropoda. The phylogeny also shows that taxonomy never ends due to the inclusion of two aclidids within the Heterobranchia.

The monophyly of the Heterobranchia was confirmed while the monophyly of Euthyneura as well as Opisthobranchia and Pulmonata was rejected. This supports the request of re-evaluation of morphological characters as well as molecular data that have been used to analyse relationships within gastropods due to a missing phylogenetic signal.

Except for Glacidorboidea and Pyramidellidae all "Lower Heterobranchia" are nested at the base of the Heterobranchia. Good evidence (morphologically as well as molecular) suggests that Glacidorbidae and Pyramidellidae are more derived than originally hypothesized.

The Rissoelloidea were recovered as sister group to the Acteonoidea supporting a basal position of the latter taxon.
Nevertheless, many questions remain unanswered due to unresolved nodes in the tree. These include the clarification of the most basal heterobranch or the taxonomic position of Siphonarioidea and Sacoglossa.
Despite the integration of more taxa and molecular data in the present study, some aspects of heterobranch phylogeny remain equivocal. There were incongruencies between morphological trees and the molecular trees of this study, as well as between this and other molecular trees. Additional data such as gene order data or more refined morphological data will be required to resolve some of these problems. Moreover, fossil data are needed due to the fact that the origin of many major gastropod groups remains unclear.

Evolutionary scenarios

Other than classification, the main goal of phylogenetic studies is to give insights into the evolutionary history of characters and the evolution of taxa. This means the phylogenetic tree is more a preliminary step than a final goal (Dayrat & Tillier 2003).

When investigating the evolution of the Heterobranchia, one comes to the conclusion that the exception proves not the rule but rather, the exception is the rule. Regardless of which evolutionary event one might examine, it seems that all processes evolved convergent rather than synapomorphic. Even within related groups, homologies are often uncertain and independent origins of some structures have been proposed.

With the new and comprehensive phylogenetic framework obtained in this study it is now possible to propose evolutionary scenarios that have lead to vast diversification within the Heterobranchia.

Upon comparing the paraphyletic “Lower Heterobranchia” with the Euthyneura, a distinguishing feature within species abundance is evident. The “Lower Heterobranchia” represent a step by step evolution with a marginal richness in species. Most of the families include only one or two genera (e.g. Rissoellidae, Orbitestellidae, Cimidae). In contrast, the Euthyneura are considered to be the crown group of the Gastropoda because they show an amazing species richness and ecological diversity. Reasons for their evolutionary success were probably due to several newly acquired features.

The central nervous system (CNS) has played an increasingly important role in our understanding of gastropod relationships. Torsion is probably the most distinguishing characteristic of the Gastropoda. All modern Gastropoda undergo torsion during some stage in their development. Streptoneury (twisted/crossed visceral loop) is the result. Contrasting to the streptoneury is secondary euthyneury (uncrossed visceral loop). The change from streptoneury to euthyneury is obtained by detorsion. Euthyneury distinguishes the Opisthobranchia and the Pulmonata from the remaining Gastropoda (formerly "Prosobranchia"). Many primitive features, including torsion, have been secondarily lost through evolution in opisthobranchs and pulmonates.

Little is known about the nervous system of the "Lower Heterobranchia" but as far as one knows they are more aligned with the Opisthobranchia and the Pulmonata than with the remaining Gastropoda, mainly because they have also an uncrossed (euthyneural) nervous system (Chase 2002).

Anyway, euthyneury of the Heterobranchia is a result of multiple convergences (Haszprunar 1985a, Haszprunar & Huber 1990, Bieler 1992). In contrast to euthyneury itself, there is another character of the central nervous system which is diagnostic for the Euthyneura. This is the presence of an additional pair of ganglia resulting in a so-called pentaganglionate visceral loop (Haszprunar 1988). Haszprunar (1985a) therefore, introduced the taxon Pentaganglionata Haszprunar, 1985 as a synonym of Euthyneura. Ponder & Lindberg (1997) noted that these ganglia are often absent, especially in pulmonates. However, Haszprunar (1985b) argues that this is through fusion with other ganglia. Nevertheless, Dayrat & Tillier (2000) conclude that the occurrence of five visceral ganglia is not ascertained for all euthyneuran taxa. Therefore it cannot be accepted as general character of Euthyneura.

Aside from the discussion regarding the pentaganglionate condition and the consideration as an apomorphic character of the Euthyneura, the function of this additional pair of ganglia is not yet entirely clarified. Hence, discovering the function of the additional pair of ganglia will possibly also answer the question as to what enables Euthyneura more successful than "Lower Heterobranchia".

Maybe the condition of neuronal giantism in the Pulmonata and Opisthobranchia has been a significant selective advantage. Caenogastropoda have no giant neurons while most species of the Opisthobranchia and Pulmonata show 10-20 neurons in the category "giant" (Chase 2002). The significance of giant neurons to the animals in which they are found has not been satisfactorily understood (Gillette 1991). Nevertheless, because of their size they are very

popular in neuroscience, e.g. Nudibranchia have relatively simple nervous systems, with large identifiable neurons and clusters of neurons, making them amenable to neural circuit analysis (Newcomb et al. 2006).

Gillette (1991) introduced an interesting speculation beginning with the observation that the earliest molluscs were minute creatures, whereas recent gastropods are usually larger. Because larger bodies need more servicing by the nervous system, there were two possible adaptations during evolution (Chase 2002). One was to increase the number of neurons and the other was to increase the size of existing neurons taking over multiple functions (Chase 2002). Comparing animals of equal size of the Caenogastropoda versus Opisthobranchia and Pulmonata, the former have considerably more neurons than the latter one. Consequently, the Caenogastropoda evolved the first scenario, whereas the Opisthobranchia and Pulmonata developed the second alternative (Chase 2002).

Gillette (1991) suggested that the behaviour of the Opisthobranchia and Pulmonata, relative to that of the Caenogastropoda is simpler and also underlaid by a simpler nervous system. Therefore, Opisthobranchia and Pulmonata might have economized on developmental complexity and reduced energy costs by using a small number of very large neurons (Chase 2002). Unfortunately, Gillette says nothing about neuronal giantism in the "Lower Heterobranchia".

As aforementioned, little is known about the nervous system of the "Lower Heterobranchia". Other than Haszprunar's work (1985a), where he introduced the concept of the Heterobranchia, there is only one detailed study existing from Huber (1993). Huber discusses the cerebral nervous system of marine Heterobranchia including the basal groups. Huber (1993) observed giant cerebral nerve cells in *Amathina* (Amathinidae, Pyramidellidae) and Euthyneura. They were absent in the Caenogastropoda, Architectonicoidea, Omalogyridae, Rissoellidae and in small Pyramidellidae (e.g. *Turbonilla*, *Odostomia*). The inclusion of *Amathina* in further molecular phylogenetic analyses is necessary in order to answer different questions regarding the phylogeny of the Heterobranchia particularly Pyramidelloidea. Assuming that *Amathina* would cluster with the Pyramidellidae within the Euthyneura, one has to interpret the findings of Huber (1993) in the following way. First of all, his observation would support the idea of a more derived position of the Pyramidellidae within the Euthyneura due to the presence of giant nerve cells in *Amathina* (see also discussion in the phylogenetic analyses chapter). Secondly, it would support the assumption that giant cells are correlated with body size because they were found in the relatively large *Amathina* but not in the minute *Turbonilla* or *Odostomia*. Furthermore, neuronal giantism could be a reason for a

more successful Euthyneura compared to the "Lower Heterobranchia" because they occur only in the former group. It is most probable that these neuronal cells give the Euthyneura a significant selective advantage as Gillette (1991) already assumed.

What else makes the Euthyneura more successful than their close relatives?

One of the most important innovations related to feeding was the move from grazing of microorganisms to omnivorous and then to carnivorous grazing on sessile animals (Caron et al. 2006). This step causes some of the most important adaptive radiations through dietary specialisation (Aktipis et al. 2008).
A key event was certainly the invasion of freshwater and in particular terrestrial habitats. Moreover, a specialisation on less utilised food resources such as sponges or cnidarians as evolved in several marine clades of the Opisthobranchia possibly leads to this species richness. Opisthobranchia are certainly less diverse in species numbers than other marine gastropods (Wägele & Klussmann-Kolb 2005). However, when comparing species numbers within opisthobranch taxa, it becomes quite obvious that some taxa far outnumber others (Wägele 2004). Wägele (2004) investigated potential key characters in Opisthobranchia and concluded that the examined key characters in her study are morphological characters related to feeding. She assumed them to be triggers for exploring new food sources. Moreover, it was difficult for Wägele (2004) to decide whether the switch to a new food source was the key innovation, followed by a morphological adaption promoting radiation, or vice versa. Nevertheless, the Nudibranchia is the most diverse group within the Opisthobranchia (with more than 2.700 species) while the Cephalaspidea s.str. is the second largest taxon (with at least 840 extant species) (Wägele 2004). Feeding on different kinds of food is surely one reason for the high diversification of these two groups. The same applies for the Sacoglossa which comprises approximately 300 species. The Sacoglossa opened new food resources because of the evolution of a uniseriate radula with just one median tooth per row. This uniseriate radula enables the cutting open of algal cells so that the content can be sucked out (Jensen 1997, Wägele 2004). The Chromodorididae (Nudibranchia) comprising more than 500 species, is considered to be extremely efficient by storing secondary metabolites from their sponge prey in special organs (mantle dermal formations — MDFs) (Wägele 2004).
The only lower heterobranch with a noteworthy number of species are the Architectonicoidea (about 100 species). They are (as far as is known) marine ectoparasites of colonial cnidarians (mainly zoantharian, scleractinian and antipatharian corals) (Robertson 1967; 1970). The

specialised lifestyle of the Architectonicoidea supports the idea of an evolutionary advantage because of food specialisation.

The invasion of freshwater and terrestrial habitats by Pulmonata was doubtlessly a key step in the ongoing evolution of the Euthyneura. This habitat shift has necessitated numerous adaptive changes in their respiratory, nervous, excretory, and reproductive systems, as well as in behaviour and physiology (Mordan & Wade 2008). Moreover, a radiation into these habitats has lead to an enormous increase in species numbers. The most successful Pulmonata are the terrestrial Stylommatophora with about 95% of all pulmonates. They are grouped into about 90 families with more than 10.000 species (Mordan & Wade 2008). The second successful Pulmonata are the principal freshwater taxon Hygrophila with about 1.000 species (Mordan & Wade 2008).
Monophyly of the Pulmonata is supported by various morphological characters like acquisition of a pneumostome and pulmonary vessels, presence of a procerebrum and dorsal bodies (Dayrat & Tillier 2002). All four apomorphies can be related to life outside the sea (Mordan & Wade 2008). Central to the success is the contractile pneumostome, in addition to its role in respiration it also acts as an important water storage area which reduces dehydration (Barker 2001). The procerebrum has direct nervous connections with the cephalic tentacles and is the major central site of olfactory information processing in terrestrial forms (Barker 2001). The medio-dorsal bodies appear to act on the development of both male and female cells in the gonad (Barker 2001).
An additional important key to successful invasion of nonmarine habitats is certainly the regulation of osmotic processes of body fluids. Therefore, many of the adaptations of freshwater and land snails are related to the excretory system (Andrews 1988, Mordan & Wade 2008).
Such terrestrial and freshwater adaptations are uncommon within the “Lower Heterobranchia”. No member has ever invaded a terrestrial habitat. The Valvatidae are the only basal freshwater taxon. Compared to their successful euthyneuran freshwater relatives, the Valvatidae possess no “lung” but rather a secondary gill (Rath 1988), making them probably less successful in freshwater colonisation.

The latest study by Klussmann-Kolb et al. (2008) showed that within the Pulmonata, the freshwater habitat has only been conquered once by the Hygrophila. Their reconstruction of character evolution for the different habitat types at specific nodes indicated that the ancestor

of Hygrophila probably already lived in a freshwater habitat. Moreover, Klussmann-Kolb et al. (2008) conclude, that colonisation of freshwater in Pulmonata occurred via an aquatic pathway directly from the marine habitat and not via a terrestrial step because the ancestor of Eupulmonata and Hygrophila appeared to have lived in a marginal zone (e.g. supralittoral zones, estuaries or mangroves).

The present study provides new insights into freshwater colonisation by heterobranch gastropods due to the inclusion of additional limnic taxa (e.g. *Valvata*, *Glacidorbis*) in the phylogenetic analyses. Within Heterobranchia, the colonisation of freshwater occurred several times independently. Once by the taxon Valvatidae and twice within the Euthyneura. The genus *Valvata* lives in freshwater while the sister taxon Cornirostridae only occurs in the intertidal zone in sheltered, fully marine water. Moreover, the present study clearly shows that *Glacidorbis* is related to the Pulmonata. Consequently, within the Pulmonata, the colonisation of freshwater happened at least twice. Once by the Hygrophila and in parallel by the Glacidorbidae in the Australian region.

In conclusion, it can be stated that the adaptation of different feeding habits has had a noticeable influence on Heterobranchia evolution. Especially the Opisthobranchia have undergone what appears to be explosive adaptive radiations because of food specialization. Moreover, the successful invasion of non-marine habitats has had a profound influence on heterobranch taxa also. Here, the Pulmonata represent by far the most significant invasion of limnic as well as terrestrial environments.

Lower heterobranchs show neither a strong food specialization nor was there a significant shift in environments. Consequently, the “Lower Heterobranchia” show a poorer diversification than their close relatives.

The next chapter deals with the molecular dating of the present phylogenetic hypothesis by using the software Beast. This method reconstructs phylogenies and presents a framework for testing evolutionary hypotheses without conditioning on a single tree topology. The obtained results should help to prove or reject some of the equivocal results, enabling a better understanding of heterobranch evolution.

4. Evolution of the Heterobranchia

4.1 Introduction

"The present is the key to the past" is one of the key concepts in geology and palaeontology. When it comes to the reconstruction of the evolution of gastropods the opposite is also true, and the past represents the key to their present classification (Bandel 1997). It does not matter which evolutionary theory one may prefer, the evidence to support it must come from the fossil record (Bandel 1997). Without fossils the evolutionary history can never be reconstructed in a way that comes close to the truth (Bandel 1997).

Gastropods have remained surprisingly underutilized as models for evolutionary studies. No other animal group offers an equal opportunity to combine the results of morphological and molecular studies of the diverse living fauna with data derived from the extensive fossil record (Bieler 1992). Fossils of the Gastopoda have a long history that can be traced to the Ordovician and have roots in the Cambrian (Fryda & Bandel 1997).

An independent evolutionary history of the Heterobranchia started a long time ago (Bandel & Heidelberger 2002). A Heterobranchia relative is documented from Early Devonian (Emsian) (*Kuskokwimia* Fryda & Blodgett 2001) and similar species lived at the Mid Devonian (*Plaeocarboninia* Bandel & Heidelberger 2002). Both gastropod taxa (*Kuskokwimia* Fryda & Blodgett, 2001 and *Plaeocarboninia* Bandel & Heidelberger, 2002) belong to a group that can be connected with Mesozoic and extant representatives of the Valvatoidea (Bandel 2002).
Bandel (1994, 2002) proposed that the oldest Opisthobranchia appeared in the Triassic (~ 220 Ma). With the exception of the Cephalaspidea and Pteropoda, very little information is available on fossil Opisthobranchia. Like most other molluscan groups, opisthobranchs do have a poor fossil record. Reasons for that are their reduced, thin-walled or complete absent shells. Many families with numerous extant, shelled representatives have never been found in the fossil record (Valdez & Lozouet 2000).
According to Bandel (1994; 2002), the Pulmonata appeared in the Jurassic (~ 190 Ma). The fossil record of the Pulmonata is incomplete, too. Reasons for that are sometimes reduced or lacking shells as well as many shells formed of aragonite which does not preserve well. Additionally, the systematic interpretation is often difficult, because there is much

convergence in shell morphology (Wade et al. 2006). Hence, the fossil data that are available must be interpreted with caution (Mordan & Wade 2008).

Nevertheless, fossil gastropods must be included into a system of the phylogenetic relationships of extant gastropods, if the resulting system is to be considered in an evolutionary framework (Bandel 1997). Without taking the fossil forms into account a reconstruction of the phylogenetic relationships of modern gastropods would be incomplete (Bandel 1997).

Fossil data and molecular clock approaches together are a promising combination for investigating evolutionary events. A molecular clock measures the number of changes, or mutations, which accumulate in the gene sequences of different species over time. The idea of dating evolutionary divergences using calibrated sequence differences was first proposed in 1965 by Zuckerkandl & Pauling. Based on this idea, molecular dating has been used in many studies as a method to investigate mechanisms and processes of evolution (for a review see Rutschmann 2006). Drummond et al. (2006) introduced a new "relaxed" approach for the estimation of phylogenetic divergence times. A "relaxed" molecular clock is a phylogenetic technique that allows the rate of sequence evolution to vary among groups of organisms (Pybus 2006). Furthermore, this new approach estimates phylogeny shape and rate variation among phylogeny branches simultaneously. These are two processes that had to be performed separately in the past (Pybus 2006).

Nevertheless, molecular clocks had a difficult status over the years and one has to keep in mind that there are many sources of error in estimating the actual date of origin of a clade, e.g. an incorrect phylogenetic topology, incomplete fossil record, wrong determination (Donoghue & Benton 2007).

However, for the first time reliably dated trees provide the opportunity to explore a comprehensive field within evolution. Fossil and molecular date estimates are more and more congruent (Benton & Ayala 2003) and this trend has increased (Bromham 2006) as molecular clock analyses have become more sophisticated (Welch & Bromham 2005).

In the present study the phylogenetic relationships of the Heterobranchia with the combined dataset I will be reinvestigated. Based on the resulting phylogeny the evolutionary timescales

of groups belonging to the Heterobranchia will be calculated with the software Beast which is a newly developed relaxed-clock Bayesian dating approach.

4.2 Material and methods

Taxon sampling

(according to chapter 2.2)

DNA extraction, amplification and sequencing

(according to chapter 2.2)

Sequence editing and alignment

(according to chapter 2.2)

Molecular clock

Dataset I (see tab. A5 in the appendix) was used to estimate approximate divergence times using a relaxed clock method (Drummond et al. 2006), as implemented in the software Beast 1.4.8 (Drummond & Rambout 2007). Beast applies Bayesian methods in search for the optimal phylogeny and estimates divergence times simultaneously. Four nodes were chosen as primary calibrations points with a normal distributed prior for the divergence time (tab. 4.1).

Tab. 4.1: Fossil calibration nodes

Calibration	Age (Ma)	Fossil	Source
Heterobranchia	399 ± 11.5	*Palaeocarboninia jankei*	Bandel & Heidelberger 2002
Acteonoidea	210 ± 6,1	*Tornatellaea heberti*	Tracey et al. 1993
Omalogyridae	88 ± 2.4	*Omalogyra sp.*	Tracey et al. 1993
Pteropoda	62 ± 1.8	*Heliconoides mercinensis*	Tracey et al. 1993, Bandel 1997

Divergence times for the remaining nodes in the tree were estimated with Beast using a GTR + I + G model of nucleotide substitution. The Yule process was used to describe speciation. The MCMC chain was run for 20 million generations sampled every 1000 generations. The first 1500 trees were discarded as burnin.

A lineage-through-time plot from the maximum clade probability tree of dataset I was generated with the software Mesquite version 2.5 (Maddison & Maddison 2008).

4.3 Results

The phylogenetic hypothesis obtained with the software Beast (fig. 4.1) appears quite similar to the one reconstructed with MrBayes and RAxML (see fig. 3.2). There are only little differences within the Euthyneura especially Pulmonata. The Siphonarioidea appear as sister group to the Sacoglossa with a reasonable statistical support (0.95) and both are the sister taxon to the remaining Pulmonata. Surprisingly, the Hygrophila are not monophyletic but the phylogenetic position of both included taxa (*Acroloxus* and *Latia*) remains unclear because of a weak statistical support. The same applies for *Phallomedusa* (Amphiboloidea).

Four fossil calibration points to estimate the divergence dates within the Heterobranchia were used (see tab. 4.1). The posterior mean of the divergence time at the root (fig. 4.1/node 1) (origin of the Vetigastropoda and Caenogastropoda + Heterobranchia) was defined at 500 million years ago (Ma), which agrees with palaeontological data suggesting an appearence of the earliest representatives of the group in the Cambrian (Ponder & Lindberg 1997). However, due to the wide uniform priors used for the calibration points in the present study and the high geological age, the 95% confidence interval (CI) remains large at most of the nodes (see fig. 4.2). Therefore, the results of the present study should be considered as a working hypothesis respectively a first insight into the origin and age of the Heterobranchia and their subgroups.

The combined Beast gene analysis (fig. 4.1) dated the most recent ancestor of the Caenogastropoda and the Heterobranchia to Middle Silurian (node 2). The divergence of the major clades of the “Lower Heterobranchia” took place during Middle Carboniferous (fig. 4.1/nodes 3-5) where most of the basal taxa originated (e.g. the two aclidids *Graphis* and *Larochella*, Murchisonellidea, Valvatoidea, Architectonicoidea, Omalogyroidea and Orbitestellidae).

Beast reconstructed a Middle Permian (fig. 4.1/node 6) origin of the Cimidae and the remaining Heterobranchia. The divergence of the latter clade took place later during the Mesozoic. At Middle Triassic (fig. 4.1/node 9) the latest common ancestor of the Acteonoidea + Rissoelloidea as well as the ancestor of the Euthyneura occurs. The origin of the Opisthobranchia was estimated about late Triassic (fig. 4.1/node 11) with the Nudipleura as the earliest offshoot. The initial divergence within the remaining Opisthobranchia occurred during the end of Jurassic to beginning of Cretaceous.

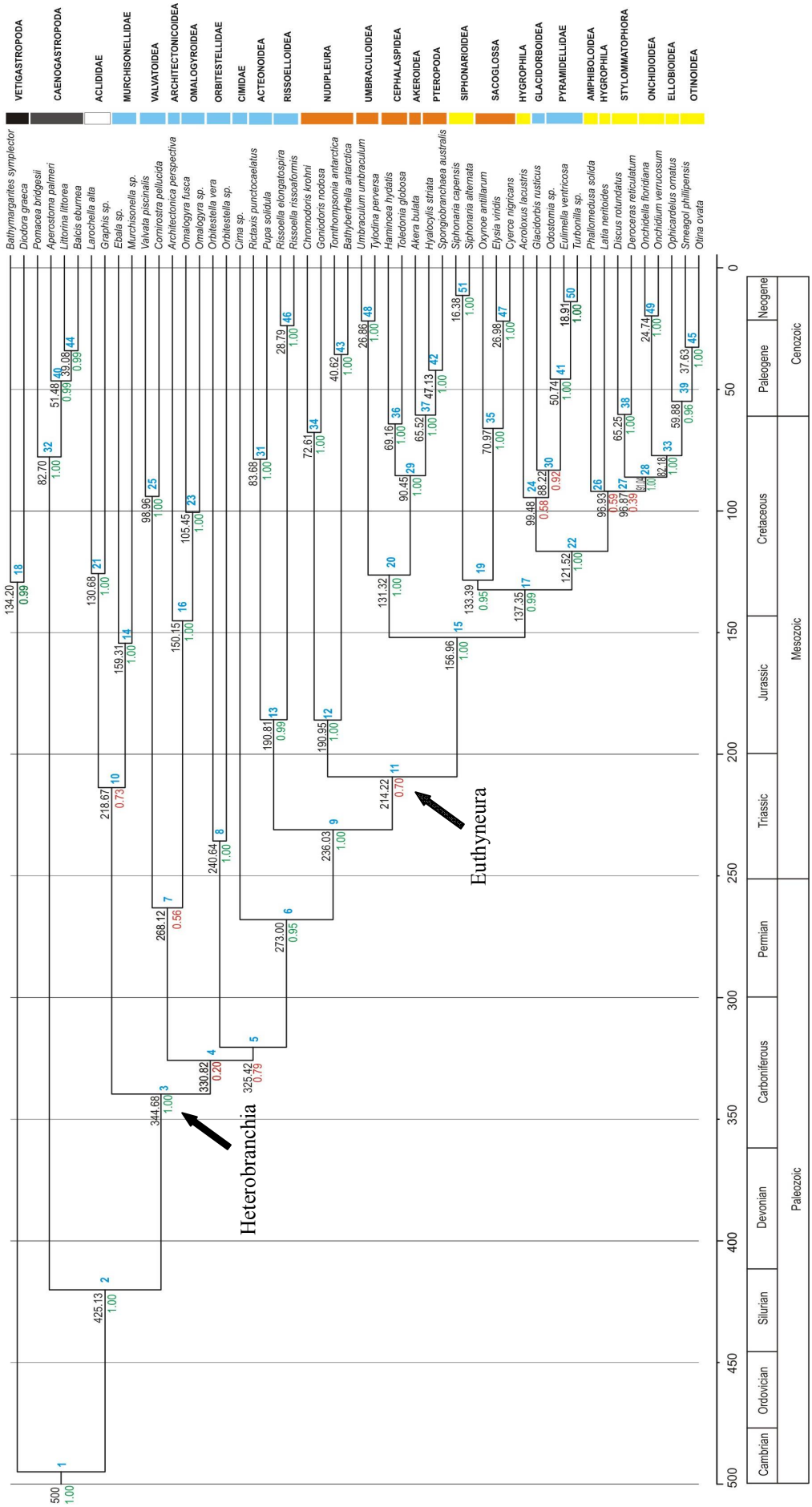


Fig. 4.1: Maximum clade probability tree displayed as a chronogram of dataset I under the relaxed clock analysis using a GTR + G + I substitution model; nodes are numbered in blue according to their appearance date; mean ages are provided above the branches, posterior probabilities below the branches (green: statistically significant, red: statistically insignificant); geological time scale follows the International Commission on Stratigraphy (see fig. A1 in the appendix).

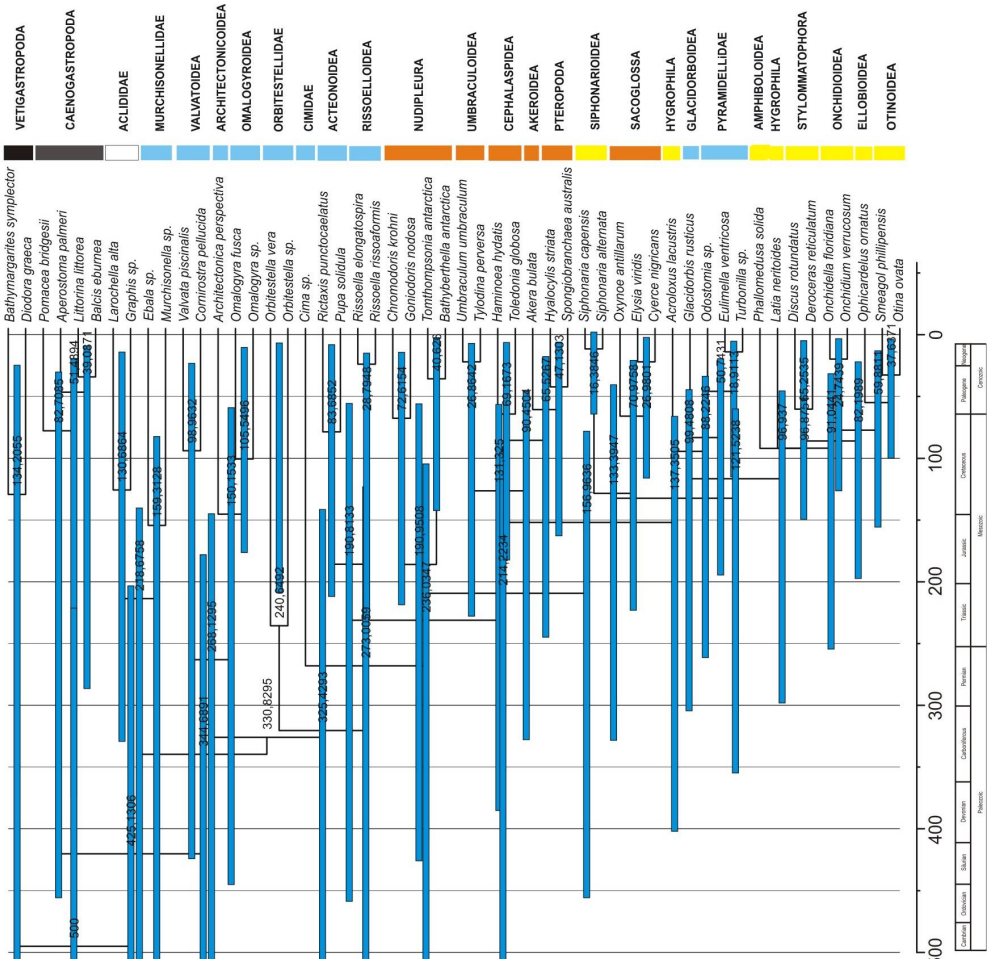


Fig. 4.2: Confidence intervals of the Maximum clade probability tree displayed as a chronogram of dataset I under the relaxed clock analysis using a GTR + G + I substitution model; blue bars represent the 95% highest posterior density intervals (= confidence intervals) for the divergence time estimates; mean ages are provided at the blue bars; geological time scale follows the International Commission on Stratigraphy (see fig. A1 in the appendix).

An initial divergence of the Euthyneura with rise of many of the extant taxa, is estimated to have happened in Early Cretaceous (fig. 4.1/nodes 17, 19, 20, 22). The next divergence events within the Euthyneura were from the Middle up to the Late Cretaceous (e.g. fig. 4.1/nodes 24, 26-30, 33, 34).

Glacidorboidea and Pyramidellidae seem also to have their origin in the Cretaceous (fig. 4.1/nodes 24, 30) but show posterior probabilities with no statistical support at the nodes of their common ancestor. Therefore, it is difficult to infer when they first appeared during the evolutionary history of the Gastropoda.

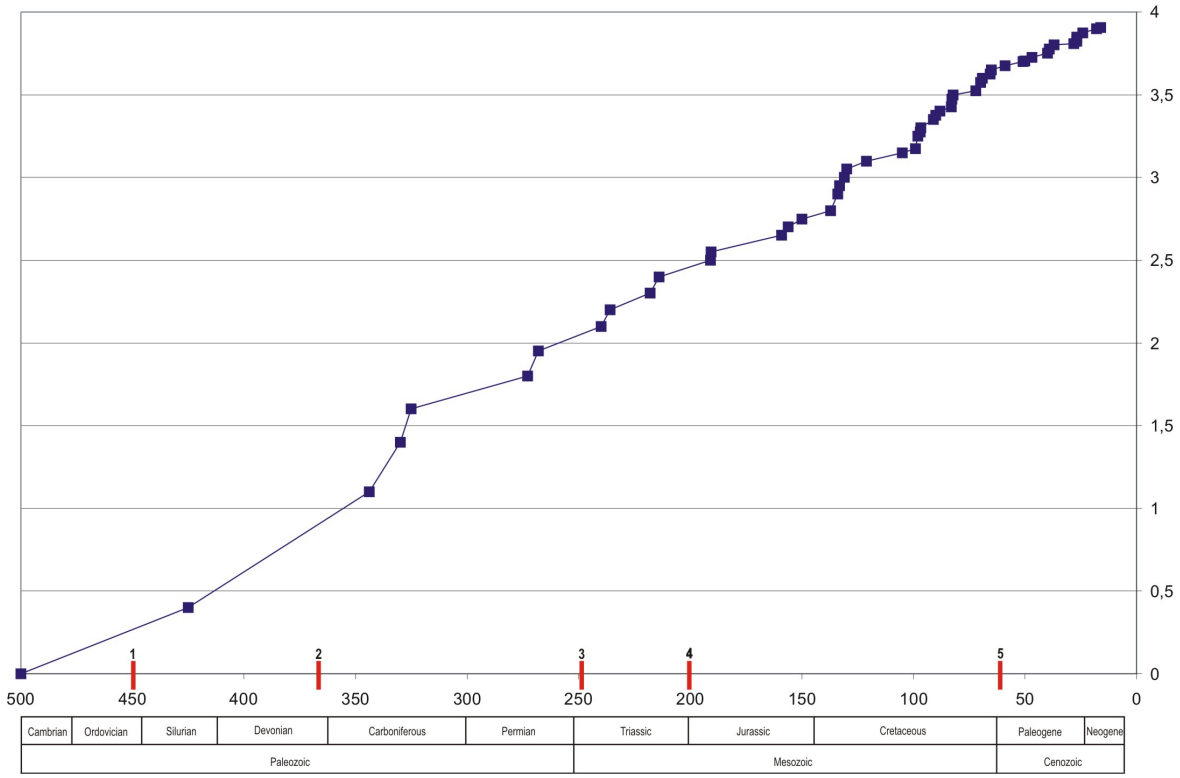


Fig. 4.3: Lineage-through-time plot for the consensus tree of dataset I; x-axis: time in Ma, y-axis: logarithm of number of lineages; the big five mass extinction events are marked with red bars (1: End Ordovician, 2: Late Devonian, 3: End Permian, 4: End Triassic, 5: End Cretaceous); geological time scale follows the International Commission on Stratigraphy (see fig. A1 in the appendix).

The lineage-through-time-plot (fig. 4.3) reveals a more or less continuous lineage increase. Fitting a geological time scale to the plot one can see that during the Carboniferous an increase of lineages is implied. However, the number of new lineages are not significant to make a final statement. From Middle Permian to Late Jurassic a more or less continuous increase of new lineages can be seen. At the beginning of the Cretaceous an increase of lineages occurs but like in the Carboniferous the number of new lineages are not significant.

Adding more taxa would possibly potentiate the observed increase effect. Finally, from Middle Cretaceous to date a continuous lineage increase can be observed.

4.4 Discussion

Origin and age of lineages

Haszprunar (1988) proposed a common ancestry of Heterobranchia and Caenogastropoda. Moreover, he placed heterostroph taxa which are neither Opisthobranchia nor Pulmonata in the paraphyletic taxon "Allogastropoda" ("Lower Heterobranchia"). Because of the right coiled shell of the Caenogastropoda and the left coiled protoconch and right coiled teleoconch of the Heterobranchia it was questionable whether these two taxa are related to each other (Bandel 1997). The known fossil record allows a relationship between members of the Heterobranchia and Caenogastropoda only when their common ancestors have lived before the Devonian (Bandel 2002). They certainly represent different phylogenetic lineages by about Mid Devonian (Bandel 1994; 1995; 1996) but they have no fossil record older to that. The phylogeny of the present study supports the idea of a close relationship between the Caenogastropoda and the Heterobranchia. Moreover, the results of the present study support the assumption that the common ancestor of the Caenogastropoda and Heterobranchia lived before the Devonian (fig. 4.1/node 2).

As far as is known up to date only representatives of the "Lower Heterobranchia" lived during the Paleozoic (Bandel 2002). The results of the present study indicate a first radiation during Middle Carboniferous (fig. 4.1/nodes 3-5) where most of the major lineages of the "Lower Heterobranchia" originated.

The result of the Beast analysis of a non monophyletic Pyramidelloidea supports the result of the phylogenetic reconstruction already discussed in chapter 4.1. The Ebalidae (Murchisonellidae) existed from the Triassic (Schröder 1995, Bandel 1994) and still live in marine habitats (Bandel 2002) whereas the Pyramidellidae with many extant parasitic species appear in the geological record not before Cretaceous (Bandel 1995; 1996; 1997, Kiel & Bandel 2001). Schröder (1995) interpreted a fossil called *Kleinella* from the Lower and Middle Jurassic as member of the Pyramidellidae, but probably it belongs to the Donaldinidae. True Pyramidellidae are found in the Campanian and Maastrichtian of the Cretaceous with fossils called *Creonella* and *Lacrimifromia* (Bandel 1994). The results of the present study also support a very early occurrence of the Ebalidae (Murchisonellidae) and a later occurrence of the Pyramidellidae during the history of Gastropoda. It is therefore hard to believe that both

groups have the same origin. A basal position of the Pyramidellidae seems also implausible because nearly all major lineages of the "Lower Heterobranchia" have their origin during the Paleozoic but not later than Early Mesozoic.
The neritiform Amathinidae are considered to represent close relatives of the Pyramidellidae. They also do not appear prior to the Upper Cretaceous (Bandel 1994). To shed light on the origin of the Pyramidellidae it is absolutely necessary to include members of the Amathinidae (Pyramidelloidea) in further analyses.

The earliest occurrence of Opisthobranchia and Pulmonata in the fossil record has been shifted to the Mesozoic due to current knowledge (Fryda et al. 2008). The results of the present study support the idea of an occurrence of the Euthyneura stemline not before the Mesozoic (fig. 4.1/node 9).

Opisthobranchia can not be traced back prior to the Triassic (Bandel 1991; 1994). What has been considered to represent Opisthobranchia from the Carboniferous (Kollmann & Yochelson 1976) are non-heterostrophic Gastropoda with a convergent shell shape (Bandel 2002). The present study support the assumption of Bandel (1991; 1994) that the ancestor of the Opisthobranchia appeared not before Triassic (fig. 4.1/node 11).
Gründel (1997) proposed a radiation of the Opisthobranchia during the Jurassic where they obtain a first large divergence event. Since then they are an important part of the gastropod fauna.

Opisthobranchia particularly Cephalaspidea are represented by members of the Cylindrobullinoidea from the Late Triassic and Jurassic (Schröder 1995, Bandel 1991; 1994). Acteonoidea can be traced from the Middle Jurassic (Schröder 1995, Gründel 1997) in a continuous lineage to the modern species. Pteropoda as well as Nudipleura appear during the Paleogene (Bandel 1997). In the present study the ancestors of the Nudipleura appear to be the oldest Opisthobranchia. This result is contrary to the assumption that the Nudipleura are highly derived (Wägele & Willan 2000, Wägele et al. 2003, Vonnemann et al. 2005 and Wägele & Klussmann-Kolb 2005) (see also discussion in chapter 4.1). Once again, the author of the present study is unwilling to believe in a basal position of the Nudipleura. Problems during the phylogeny reconstruction respectively estimating divergence times because of deviant base composition and rate heterogeneity should be considered to have caused this unexpected result. Nevertheless, one should keep in mind that the fossil record of the

Nudipleura is problematic. Due to lacking hard body parts there is no reliable fossil record for the Nudibranchia. The sister group to the Nudibranchia are the Pleurobranchoidea (Wägele & Willan 2000). They show a fossil record which can be traced back till the end of the Paleogene beginning of the Neogene (Valdés & Lozouet 2000). Some Pleurobranchoidea have small and fragile internal shells or even lack a shell. It is therefore likely that only a part of their historical diversity has been preserved (Valdés 2004). Nevertheless, the results of the present study support a late appearance of the modern Pleurobranchoidea during the Cenozoic (fig. 4.1/node 43). But the divergence time of the Nudibranchia was estimated to have happened earlier (fig. 4.1/node 34) and the ancestor of both possibly occurred a long time ago during the Mesozoic (fig. 4.1/node 12).
Further analyses with more Nudipleura are needed to shed light on the history of this enigmatic group, which is rich in species diversity but poor in fossil record.

The present study supports the divergence of many Pulmonata clades during the beginning of the Cretaceous (fig. 4.1/nodes 17, 19, 22).
Earliest Pulmonata can not be recognized with any certainty in the Triassic but have been documented from the Jurassic (Kiel & Bandel 2001). Among the Pulmonata the Basommatophora are more ancient appearing in the Jurassic (Bandel 1991) while the Stylommatophora are recognizable during the Late Mesozoic with some doubtful species (Bandel 1991) but with better recognized taxa in the Late Cretaceous (Bandel & Riedel 1994). The present study supports this hypothesis.
This is contrary to the fossil record of Upper Carboniferous terrestrial pulmonates which are regarded to be the earliest stylommatophoran land snails by Solem & Yochelson (1979). Some of them have been re-interpreted as non-stylommatophoran by Bandel (1991; 1997). Moreover, a no Paleozoic origin of the Stylommatophora was supported by sequence studies of 28S rDNA fragments (about 700 bp in total) carried out by Tillier et al. (1996) who inferred the ages of divergence from branch lengths in a tree of molecular distances. He confirmed a Late Mesozoic derivation of this large group of gastropods. This is also supported by the present data. However, because of an incomplete taxon sampling (only two Stylommatophora were included) a statement about radiation events of this group is impossible.
Wade et al. (2006) concluded that a Palaeozoic origin of the Stylommatophora remains a possibility but supporting evidence is lacking. The results of the present study support a late divergence of the Pulmonata in the Jurassic and Cretaceous.

Ellobiidae (Pulmonata) can safely be traced to the Late Jurassic (Bandel 1991) and usually live in intertidal mud flats, coastal forests and swambs. Their close relatives Chilinidae are found in rivers and lakes while the Carychiidae live on land in wet litter and moss (Bandel 1997). The latter one may even be present on land since Late Carboniferous (Bandel 1997) but this contrasts of course the comparatively late appearance of pulmonates like the Basommatophora and the Ellobiidae in the Early Mesozoic. Because Carychiidae and Chilinidae were not included in the present study, a statement about a possible early occurrence of the Carychiidae could not be made. Nevertheless, both taxa should be included in future analyses to get a more detailed picture of pulmonate evolution.

Lineage-through-time plot

The lineage-through-time plot (fig. 4.3) exhibits a more or less continuous diversification through time. Adding more taxa would possibly potentiate two indicated accelerated lineage splitting periods in the Carboniferous and Cretaceous.

When fitting mass extinction events to the lineage-through-time plot then it is evident that the two implied lineage increase periods coincide with recovery phases. Both recovering phases occurred after a mass extinction event. The lineage increase during the Carboniferous could be correlated with the late Devonian mass extinction which took place during the later part of the Devonian at the Frasnian-Famennian boundary. It was one of the “Big Five” major extinction events in the history of the Earth's biota. This crisis primarily affected the marine community. Among marine invertebrates, 70% of the taxa did not survive into the Carboniferous (Elewa 2008). Reasons for the Late Devonian extinction are still speculative. This event was described either to glaciation or meteorite impact leading to an episode of global cooling. Warm water marine species were the most affected organisms in this extinction event (Elewa 2008). Niches which were occupied during the Devonian by various marine invertebrates were now possibly open for the Heterobranchia. Moreover, reef-builders like tabulate corals and stromatoporoids which were the food resources of many marine invertebrates never truly recovered from the extinctions (McGhee 1996).

No major extinction or diversification event separates the Cretaceous from the Jurassic (Stanley 2001, Eleva 2008). Hence, there must be another explanation for an accelerated diversification at the Jurassic-Cretaceous boundary. The Cretaceous was a period with a relatively warm climate and high eustatic sea level. Therefore, a large area of the continents was covered by warm shallow seas (Stanley 2001). Moreover, the oceans were enriched with calcium; best life conditions for most of the heterobranch gastropods. This good life

conditions possibly led to an increase of lineages within this taxon. During that time several occupation events took place like the invasions of land and fresh water, e.g. Provalvata migrated into fresh water during Jurassic and may have given rise to the recent fresh water Valvatoidea (Bandel 1991). Because of an incomplete taxon sampling of the species rich Stylommatophora (only two taxa were included in the present study) it is not possible to see any effect on the lineage-through-time plot of a possible radiation of this group. Nevertheless, Tillier et al. (1996) proposed that fossil evidence indicates an explosive radiation of the Stylommatophora at the Upper Cretaceous-Paleocene period because most of the families known as fossils appear at that time. For further analyses more Stylommatophora should be included to verify the assumption of Tillier et al. (1996) and to get a better idea of Stylommatophora radiation.

Conclusion

The present study is the first comprehensive survey using molecular clock approaches to estimate divergence time within the Heterobranchia. Nevertheless, as aforementioned the confidence intervals (CI) were large in most of the cases meaning that a precise dating of the nodes was impossible. Anyway, many evolutionary hypotheses based on fossils could be confirmed.

Molecular clocks are still in their infancy but it could be shown that the present study or others before (e.g. Krause et al. 2008, Njabo et al. 2008, Zhang et al. 2008 etc.) can make a contribution to a better understanding and reconstruction of evolutionary processes. Therefore, one goal should be the improvement of molecular clock methods (which already happened with the introduction of relaxed molecular clocks). Moreover, the fossil record needs to be enhanced to close fossil gaps which will maximise the accuracy of the calibration points.

.

5. *A posteriori* evaluation of data quality

5.1 Introduction

Given that some of the phylogenetic hypotheses proposed in chapter 3 are contrary to other hypotheses based on morphological (e.g. Haszprunar 1985a; 1988, Ponder & Lindberg 1997, Dayrat & Tillier 2002) and molecular data (e.g. Colgan et al. 2003, Grande et al. 2004a; 2008), it is imperative that the current hypotheses are further tested using independent data sets.

Different approaches exist for testing data *a posteriori*. One option is to apply various statistical tests, e.g. the Approximately Unbiased test (AU), Shimodaira-Hasegawa test (SH) and Kishino-Hasegawa (KH) test from Shimodaira (2002). These tests compare alternative tree topologies and evaluate the differences between trees based on their likelihood scores using bootstrapping.

Investigations of rRNA secondary structures also seem to be promising for evaluating data quality *a posteriori*. The rRNA molecules fold into specific secondary structures, which are important for conservation of their three dimensional structure and their function within the ribosome. The secondary structure is maintained by hydrogen bonds between RNA nucleotides, which form stems (paired regions) or loops (unpaired regions). The stem regions show a high degree of conservation while the loops have a considerable amount of variability (Caetano-Anollés 2002). A feature that makes rRNA markers popular in phylogenetics because different questions with different time scales of diversification can be answered (Higgs 2000).

Nevertheless, information regarding the secondary structure is missing in most phylogenetic studies although the secondary structure has consequences for the use of rRNA molecules in phylogenetic reconstruction. The pairing between the stem nucleotides has important influence on their evolution which differs from that of unpaired loop nucleotides. These differences in evolution should be taken into account when using rDNA sequences for phylogeny estimation (Telford et al. 2005). Specific rDNA evolutionary models have to be applied in order to overcome the problem of co-evolution of paired sites, which violates the basic assumption of the independent evolution of sites made by most phylogenetic methods (Dixon & Hills 1993). Moreover, information about secondary structure also supports the process of aligning rDNA sequences (Kjer 1995, Buckley et al. 2000, Hickson et al. 2000, Misof et al. 2001, Gillespie et al. 2005, Voigt et al. 2008). Both aspects increase the accuracy of phylogenetic reconstructions.

However, secondary structure models are still little used in phylogenetic analyses, probably because establishing a secondary structure for a new sequence is still a time-consuming process and very few software packages allow the simultaneous analysis of paired and unpaired rRNA regions (Voigt et al. 2008). Some databases (e.g. Cannone et al. 2002) provide secondary structure information for a number of organisms, but their records are far from complete.
Lydeard et al. (2002) showed based on a comprehensive analysis of mitochondrial LSU rDNA sequences and subsequently derived secondary structures that the loss or reduction of three helical-loop structures are apomorphies of the Heterobranchia. Hence, it seems possible that comparative analyses of secondary structures of heterobranch taxa will yield phylogenetically informative data for supporting deep evolutionary nodes.

This chapter deals with the *a posteriori* evaluation of data quality. The phylogenetic hypotheses proposed in chapter 3 will be proven or rejected with various statistical tests as well as network analyses and secondary structure reconstruction. For the latter method a comprehensive survey of the complete 18S rRNA and 28S rRNA secondary structure of representatives of most of the major heterobranch groups was performed for the first time. Secondary structures were reconstructed and browsed for possible synapomorphies to support certain nodes in the phylogenetic tree (fig. 3.2). Furthermore, a comparative study was conducted using standard evolutionary models implemented in the software MrBayes (Huelsenbeck & Ronquist 2001) as well as rDNA specific models (which takes paired and unpaired sites into account) implemented in the software package Phase 2.0 beta (http://www.bioinf.manchester.ac.uk/resources/phase/). By accounting the secondary structure in the models of evolution the author hopes to improve the plausibility of the phylogenetic tree.

5.2 Material and methods

Taxon sampling for additional datasets using complete 18S and 28S rDNA sequences

The complete 18S rRNA of 45 gastropod species have been investigated (Dataset III = 2 Vetigastropoda, 4 Caenogastropoda, 12 "Lower Heterobranchia", 14 Opisthobranchia, 11 Pulmonata and 2 taxa not assigned to the Heterobranchia yet) as well as the complete 28S rRNA of 22 gastropod species (Dataset IV = 3 Vetigastropoda, 1 Caenogastropoda, 7 "Lower Heterobranchia", 5 Opisthobranchia, 5 Pulmonata and 1 taxon not assigned to the Heterobranchia yet). For optimal results of the secondary structure reconstruction only sequences without missing data were included in the analyses. For details about the taxonomy and collecting locations of the sampled taxa as well as Genbank accession numbers see tab. A1 in the appendix.

The animals were collected from the field by hand, snorkelling or scuba diving and stored in 70-100% ethanol. Most of the "Lower Heterobranchia" were collected intertidally while collecting algae or substrata where they are living on. The material was washed and sieved and the animals were picked alive under the binocular.

Taxon sampling for dataset I was according to chapter 2.

DNA extraction, amplification and sequencing

(according to chapter 2.2)

Sequence editing and alignment

(according to chapter 2.2)

No ambiguous alignment positions were excluded in dataset III and dataset IV due to the secondary structure reconstruction.

Approximately Unbiased (AU) Test

Alternative tree topologies for dataset I were tested using the Approximately Unbiased (AU) Test developed by Shimodaira (2004).

The likelihood at each nucleotide position was calculated for an unconstrained topology as well as different alternative constrained topologies (monophyletic Opisthobranchia, monophyletic Pulmonata and monophyletic Euthyneura) according to the latest classification of Bouchet & Rocroi (2005) using PAUP 4.0b10 (Swofford 2002). The obtained likelihoods

were used to calculate p-values using default settings of the software CONSEL version 0.1 (Shimodaira & Hasegawa 2001).

Relative rate test and Network analyses

(according to chapter 2.2)

Phylogenetic analyses

Bayesian inference phylogenetic analyses were performed using the software MrBayes 3.1.2 (Huelsenbeck & Ronquist 2001) and Phase 2.0 beta (http://www.bioinf.manchester.ac.uk/resources/phase/) for two different datasets (see tab. A5 in the appendix):

Dataset III = complete 18S rDNA sequences for secondary structure reconstruction; no alignment positions were excluded

Dataset IV = complete 28S rDNA sequences for secondary structure reconstruction; no alignment positions were excluded

Detailed information about MrBayes 3.1.2 (Huelsenbeck & Ronquist 2001) see chapter 2.2.
The analysis for each data set was run for 2.000.000 generations, with a sample frequency of 10. The first 20.000 generations were discarded as burnin. If likelihoods have not reached a plateau the burnin was increased to 40.000 generations.
Support for nodes was expressed as posterior probabilities.
The best-fit models of nucleotide substitution were selected with MrModeltest 2.2 (Nylander, 2004) choosing the Akaike information criterion (AIC) (see tab. A5 in the appendix).

The Phase package is designed specifically for usage with RNA sequences that have a conserved secondary structure, e.g. rRNA. Four simultaneous chains were run for 2.000.000 generations, with a sample frequency of 10. The first 20.000 generations were discarded as burnin. A mixed model was used. Unpaired nucleotides were handled by the model REV (Tavare 1986) and paired nucleotides by the model RNA7D (Tillier & Collins 1998).

Secondary structures

Secondary structures of the rRNA of dataset III and IV were reconstructed using the software RNAsalsa (Stocsits et al. submittet). The result of this application is a complete individual secondary structure for each sequence. For further phylogenetic analysis, a sequence alignment with a consensus structure is produced, which can be used as an input for suitable programs like Phase 2.0 beta (http://www.bioinf.manchester.ac.uk/resources/phase/) where evolutionary models specific to the stem and loop regions of structural RNA molecules are implemented.
RNAsalsa is a method for aligning ribosomal RNA sequences, by adopting thermodynamic folding (using the folding algorithm taken from the Vienna RNA package – RNAfold) and comparative evidence algorithms, combined in a suitable framework. The program simultaneously generates secondary structures for a set of homologous RNA genes and aligns them by taking sequences and structure information into account from so-called constraint sequences.

Constraint sequences (sequences with an already available secondary structure) for 18S rRNA (*Monodonta labio* – Vetigastropoda, Mollusca) and 28S rRNA (*Caenorhabditis elegans* - Rhabditida, Nematoda) were downloaded from the European ribosomal RNA database (http://bioinformatics.psb.ugent.be/webtools/rRNA/).

The program XRNA 1.1.12 beta (http://rna.ucsc.edu/rnacenter/xrna/xrna.html) was used to edit secondary structure diagrams. XRNA is a Java based suite of tools for the visualisation, annotation and modification of RNA secondary structure diagrams.

Differences or similarities between secondary structures of different taxa were treated like morphological characters and included in a character matrix (see tabs. A6 and A7 in the appendix).
Character evolution was reconstructed using MacClade 4.0 (Maddison & Maddison 2000) while characters were treated as unordered and most parsimoniously mapped onto the inferred phylogeny.

5.3 Results

5.3.1 *A posteriori* evaluation to test the congruence between the phylogenetic inference and data quality

5.3.1.1 Relative-Rate-Test

The results of this test were already discussed under *a priori* aspects in chapter 2.3.3.3 and will be re-evaluated under *a posteriori* aspects in this chapter. The aspects focus on unexpected positions of certain heterobranch taxa in the resulting tree topology introduced in chapter 3.
The highest evolutionary rates were observed within the "Lower Heterobranchia" in particular *Omalogyra* sp., *Omalogyra fusca*, *Murchisonella* sp., *Ebala* sp. and *Architectonica perspectiva* (see tab. 5.1a). These high evolution rates are perfectly pictured in the 50% majority rule consensus phylogram (see fig. 3.2) by very long branches.

Tab. 5.1a: Maximum z-scores of dataset I (only noticeable high scores are shown)

Alignment	Species	z-scores (max.)
18S rDNA	***Omalogyra fusca*** vs *Orbitestella* sp.	13.722511
	***Omalogyra* sp.** vs *Orbitestella* sp.	13.472862
	***Murchisonella* sp.** vs *Orbitestella vera*	12.782731
	***Ebala* sp.** vs *Orbitestella* sp.	11.007350
	Architectonica perspectiva vs *Orbitestella vera*	10.742792
	Larochella alta vs *Orbitestella vera*	8.970545
28S rDNA	***Omalogyra* sp**. vs *Orbitestella vera*	8.040198
	Architectonica perspectiva vs *Orbitestella vera*	7.946169
	***Ebala* sp**. vs *Orbitestella vera*	7.815249
	Omalogyra fusca vs *Orbitestella vera*	7.766580
	***Murchisonella* sp.** vs *Orbitestella vera*	7.723543
16S rDNA	***Architectonica perspectiva*** vs *Umbraculum umbraculum*	4.192090
COI position 1	***Omalogyra* sp.** vs *Aperostoma pelermi*	5.273325
	Omalogyra fusca vs *Aperostoma pelermi*	5.118954
COI position 2	***Architectonica perspectiva*** vs *Valvata piscinalis*	4.726909
COI position 3		no data

Tab. 5.1b: Maximum z-scores of dataset I for Pyramidellidae and Glacidorboidea (exemplary for 18S rDNA and 28S rDNA)

Alignment	Species	z-scores (max.)
18S rDNA	***Eulimella ventricosa*** *vs. Orbitestella* sp.	2.901336
	***Odostomia* sp.** *vs. Orbitestella sp.*	2.769155
	***Turbonilla* sp.** *vs. Orbitestella* sp.	3.551441
	Glacidorbis rusticus *vs. Orbitestella sp.*	4.228920
28S rDNA	***Eulimella ventricosa*** *vs. Orbitestella vera*	1.533776
	***Odostomia* sp.** *vs. Orbitestella vera*	0.858330
	***Turbonilla* sp.** *vs. Orbitestella vera*	1.547940
	Glacidorbis rusticus *vs. Orbitestella vera*	0.718196

When comparing the maximum z-scores of dataset I of the Pyramidellidae (*Eulimella*, *Odostomia* and *Turbonilla*) and *Glacidorbis* (exemplary for 18S and 28S rDNA) (tab. 5.1b) with z-scores of the "Lower Heterobranchia" (displayed in tab. 5.1a), then a large discrepancy can be observed. The evolutionary rates of Pyramidellidae and Glacidorbidae are significantly lower than of other "Lower Heterobranchia" like *Murchisonella* or *Ebala*.

5.3.1.2 Approximately Unbiased (AU) Test

The AU test was performed to evaluate whether alternative phylogenetic hypotheses (enforcing monophyly of traditional taxa) can be rejected based on the analysed dataset I.
For AU values of the monophyly of Opisthobranchia, Pulmonata and Euthyneura see tab 5.2.
The unconstrained hypothesis shows the maximum likelihood. The three constrained topologies however cannot be discarded since their likelihoods are not significantly lower in the AU-test (p-values not smaller than the significance level of 0.05).

Tab. 5.2: Statistical test of alternative phylogenetic hypotheses of dataset I; taxa names are coded in a four letter name using the first two letters of genus and species name (for abbreviation see tab. A1 in the appendix).

Constraint		Loglikelihood	AU test (p-values)
Unconstrained	see tree topology (fig. 3.2)	-41798.54	0.829
Opisthobranchia	(Toan, Baan, Chkr, Gono, Umum, Type, Hahy, Togl, Akbu, Hyst, Spau, Oxan, Elvi, Cyni)	-41827.63	0.366
Pulmonata	(Sica, Sial, Lane, Acla, Saso, Diro, Dere, Onfl, Onve, Opor, Smph, Otov)	-41809.44	0.135
Euthyneura	(Toan, Baan, Chkr, Gono, Umum, Type, Hahy, Togl, Akbu, Hyst, Spau, Oxan, Elvi, Cyni, Sica, Sial, Lane, Acla, Saso, Diro, Dere, Onfl, Onve, Opor, Smph, Otov)	-41839.65	0.089

5.3.1.3 SplitsTree

The results of the SplitsTree analysis (fig. 5.1) will be described in context with the results of the tree reconstructions (see fig. 3.2).

The neighbournet graph created with the software SplitsTree was already examined in chapter 2.3.4.1 *a priori*. A high level of conflict was observed indicated 1.) by many parallel edges of

the same lengths and 2.) by several taxa belonging to the Veti- and Caenogastropoda, "Lower Heterobranchia" or Opisthobranchia (in particular Nudipleura) with long terminal branches.

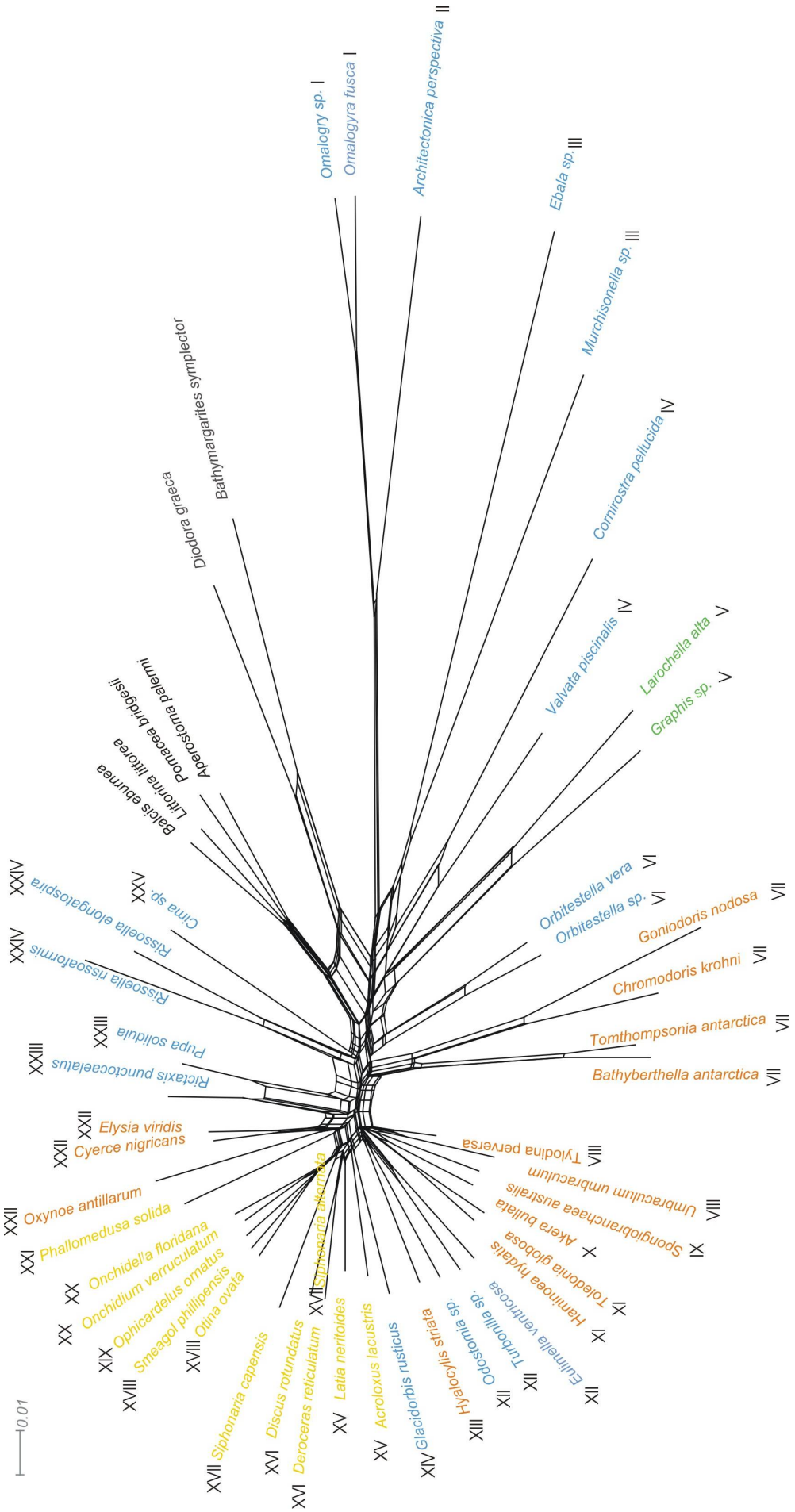
The monophyly of the Heterobranchia was supported by tree reconstructions. This result is also fostered by a good split support in the SplitsTree analysis (see fig. 5.1). The Vetigastropoda + Caenogastropoda are separated by various splits with long parallel edges from the remaining Heterobranchia (see fig. 5.1). The polyphyly of the Pyramidelloidea in the phylogenetic tree (fig. 3.2) can be supported because there is no split support for Murchisonellidae (III) being the sister taxon to Pyramidellidae (XII) (fig. 5.1). They occupy positions in the neighbour net graph far away from each other. There is split support for the Murchisonellidae (III) being the sister taxon to other "Lower Heterobranchia" taxa like 1.) the Omalogyridae (I) + Architectonicoidea (II) or 2.) Valvatoidea (IV) (without Orbitestellidae) + Aclididae (V) but none of them is distinctly stronger than the other. Nevertheless, these two splits support a close affinity of the Murchisonellidae (III) to the "Lower Heterobranchia". In contrast, there is no split support for the Pyramidellidae (XII) to other "Lower Heterobranchia". The Pyramidellidae are situated within a group comprising the Glacidorboidea (XIV), Opisthobranchia (without Nudipleura – VII and Sacoglossa – XXII) and Pulmonata (without Amphiboloidea XXI).

The SplitsTree analysis also supports the result of the tree reconstruction regarding the inclusion of the two aclidids (V) to the Heterobranchia. On the one hand there is split support for a relationship of Aclididae (V) with Valvatoidea (IV) (without Orbitestellidae) + Murchisonellidae (III) and on the other hand for a relationship of Aclididae (V) with Orbitestellidae (VI) but non of these splits has a stronger support (which would be indicated by distinctly longer parallel edges).

There is a good split support indicated by long parallel edges for *Valvata* (IV) being the sister taxon to *Cornirostra* (IV) but no support for a relationship of these two groups with the Orbitestellidae (VI). These results are congruent to the tree topology (fig. 3.2).

The Architectonicoidea (II) and the Omalogyroidea (I) are in a sister group relationship according to tree reconstruction (see chapter 3). This result can be supported by a very strong split indicated by extremely long parallel edges (see fig. 5.1).

The inclusion of the Cimidae (XXV) in the Heterobranchia as well as the single offshoot in the phylogenetic tree can be supported by a rather isolated position within the split network (fig. 5.1).



Vetigastropoda & Caenogastropoda "Lower Heterobranchia" Opisthobranchia Pulmonata Incertae sedis

Fig. 5.1: Neighbournet graph of dataset I; Heterobranchia are numbered in a clockwise direction from I to XXV; I: Omalogyroidea, II: Architectonicoidea; III: Murchisonellidae; IV: Valvatoidea; V: Aclididae; VI: Orbitestellidae; VII: Nudipleura; VIII: Umbraculoidea; IX+XIII: Pteropoda; X: Akeroidea; XI: Cephalaspidea; XII: Pyramidellidae; XIV: Glacidorboidea; XV: Hygrophila; XVI: Stylommatophora; XVII: Siphonarioidea; XVIII: Otinoidea; XIX: Ellobioidea; XX: Onchidioidea; XXI: Amphiboloidea; XXII: Sacoglossa; XXIII: Acteonoidea, XXIV: Rissoelloidea; XXV: Cimidae .

There is a strong split support for the Rissoelloidea (XXIV) and the Acteonoidea (XXIII) respectively but much conflict exists for a sister group relationship between them (see fig. 5.1). Hence, the SplitsTree analyses can neither support the results of the phylogenetic analyses (see fig. 3.2) nor reject them. The paraphyly of Euthyneura as well as the polyphyly of Opisthobranchia and the paraphyly of Pulmonata as recovered in the phylogenetic analyses (see fig. 3.2) are also supported by the neighbournet graph (fig. 5.1) because no conflict free split support for these three groups can be found.

The monophyly of the Nudipleura (VII) is supported by a strong split. The basal position of the Nudipleura (VII) within the system of the Heterobranchia as indicated by the results of the phylogenetic analyses (fig. 3.2) cannot be rejected because in the neighbournet graph they are associated with the "Lower Heterobranchia" rather than with the remaining Opisthobranchia (fig. 5.1).

There is no split support for a clade comprising the Umbraculoidea (VIII), Cephalaspidea (XI), Akeroidea (XI) (Aplysiomorpha) and Pteropoda (IX + XIII) as proposed by the phylogenetic hypothesis although they are grouped close together (with the exception of *Hyalocylis*) in the neighbournet graph (fig. 5.1). In fact there is no specifiable signal uniting these groups in the SplitsTree analysis.

There was no resolution in the phylogenetic tree regarding the exact position of the Siphonarioidea (XVII) and the Sacoglossa (XXII). Within the SplitsTree analyses the Sacoglossa (XXII) as a monophylum are supported by a split but show a rather isolated position with little split support to *Rictaxis punctocaelatus* (Acteonoidea). In contrast, the monophyly of the Siphonarioidea (XVII) is not supported due to the very short terminal branch of *Siphonaria alternata* but both taxa are enclosed within Pulmonata in the neighbournet graph (fig. 5.1).

The position of the Glacidorboidea (XIV) in the neighbournet graph (fig. 5.1) is distant to the remaining "Lower Heterobranchia" (with exception of Pyramidellidae – XII). There is no split support indicating a close relationship with other basal groups which is congruent to the results of the tree reconstruction (fig. 3.2). Because of much conflict in the neighbournet graph (fig. 5.1) within the Pulmonata, little can be said about the position of the Pyramidellidae and Glacidorbidae within the Pulmonata.

The taxon Hygrophila (XV) is supported by a weak split while there is no split support for the Eupulmonata (XVI + XVIII + XIX + XX) which is contrary to the tree reconstruction (see fig. 3.2).

5.3.1.4 SAMS

The results of the SAMS analysis will be described in context with the results of the phylogenetic tree reconstructions (see fig. 3.2).

The split support spectrum created with the Software SAMS was already examined in chapter 2.3.4.2 *a priori*. As already stated there was little binary support (red), some noisy outgroup support (green) and many noisy in- and outgroup support (yellow) for taxa partitions (see fig. 2.11b).

None of the deeper splits were found among the 60 best splits, only signals for groupings of taxa belonging to "Lower Heterobranchia" (e.g. Omalogyroidea, Aclididae), Vetigastropoda and Caenogastropoda and some Opisthobranchia (e.g. Nudipleura) were detected.

Altogether, SAMS detected 3102 partitions in dataset I. 25 groupings can be found in the 1000 most frequent partitions as well as in the corresponding phylogenetic tree (see tab. 5.3, fig. 5.2 and fig. 3.2). All other partitions contain random groupings of species.

Tab. 5.3: Split support values of groupings which can be found in the 1000 most frequent partitions of dataset I as well as in the phylogenetic tree (fig. 3.2); Taxa names are coded in a four letter name using the first two letters of the genus and species name (for abbreviation see tab. A1 in the appendix).

Split No.	Taxa	Outgroup support			Ingroup support			Posterior probability / Bootstrap-support
		binary support	noisy outgroup support	noisy in- and outgroup support	binary support	noisy outgroup support	noisy in- and outgroup support	
1	(Omfu,Omsp)	3	318	0	3	0	269	1.00/100
2	(Arpe,Omfu,Omsp)	8	1	285	8	243	0	1.00/100
4	(Laal,Grsp)	10	214	0	10	0	131	1.00/100
6	(Basy,Digr)	1	0	173	1	142	0	No support
9	(Ebsp,Musp)	1	1	152	1	74	0	1.00/100
13	(Ebsp,Musp,Laal,Grsp)	0	2	139	0	8	31	0.71/>50
22	(Toan,Baan)	0	114	0	0	0	74	1.00/100
34	(Riel,Riri)	0	91	0	0	0	60	1.00/100
49	(Chkr,Gono)	0	62	0	0	0	52	1.00/100
54	(Toan,Baan,Chkr,Gono)	1	0	67	1	14	13	1.00/100
63	(Pobr,Appa,Lili,Baeb,Basy,Digr)	0	25	24	0	0	48	1.00/99
100	(Pobr,Appa,Lili,Baeb)	0	28	23	0	0	35	1.00/-
106	(Vapi,Cope)	0	0	51	0	13	0	1.00/99
126	(Vapi,Cope,Arpe,Omfu,Omsp)	0	0	39	0	7	12	0.97/>50
155	(Diro,Dere)	0	40	0	0	0	8	1.00/100
185	(Elvi,Cyni)	0	32	0	0	0	18	1.00/100
200	(Riel,Riri,Ripu,Puso)	0	0	25	0	4	12	1.00/73
242	(Euve,Tusp)	0	24	0	0	0	6	1.00/98
274	(Ripu,Puso)	0	18	0	0	0	13	1.00/100
283	(Orve,Orsp)	0	16	0	0	0	16	1.00/100
315	(Hyst,Spau)	0	0	20	0	6	0	1.00/97
439	(Hahy,Togl)	0	16	0	0	0	1	1.00/100
485	(Umum,Type)	0	14	0	0	0	2	1.00/100
705	(Smph,Diro,Dere,Onfl,Onve,Opor,Otov)	0	0	8	0	5	0	1.00/94
979	(Euve,Odos,Tusp)	0	4	0	0	0	0	1.00/90

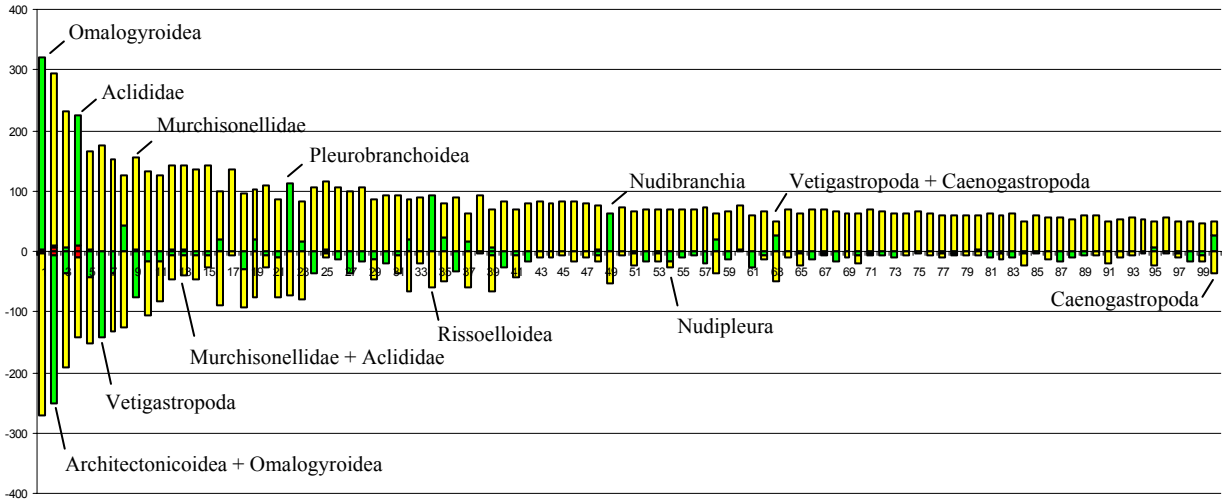


Fig. 5.2: Split support spectrum of the 100 most frequent partitions of dataset I; for the sake of clarity only the first 100 partitions are shown; x-axis = partitions, y-axis = number of sequence positions; above x-axis = outgroup, below x-axis = ingroup; red = binary support, green = noisy outgroup support, yellow = noisy in- and outgroup support.

5.3.2 Utility of the secondary structure of 18S rRNA for phylogenetic inference of the Heterobranchia

This chapter focuses on two major aspects. The first aspect is the reconstruction of the secondary structure of 18S rRNA for representatives of most of the Heterobranchia families with the software RNAsalsa. The single structures will be compared and browsed for positions which have the potential to contain a phylogenetic signal. Identified characters are mapped most parsimoniously on the phylogenetic hypothesis (see fig. 3.2) presented in chapter 3. The second aspect comprises the inclusion of rRNA secondary structure information in alignment and tree reconstruction procedures. The extract consensus structure from RNAsalsa (which provides reliable information on positional interrelation) is used for tree reconstruction and the application of specific rDNA substitution models as implemented in the software package Phase. The application of this new alignment approach on ribosomal sequence data will perhaps allow a more precise identification of positional homologies and thus phylogenetic signal within these data.

For both approaches dataset III for 18S rDNA (see tab. A5 in the appendix) was used.

5.3.2.1 Secondary structure reconstruction of 18S rRNA

The comparative analysis of all secondary structures reveals at least three types of structural domains:

Type I: conserved among all investigated taxa
Type II: variable to a certain degree but conserved among younger phylogenetic groups
Type III: highly polymorphic among all investigated taxa

When browsing the reconstructed secondary structures for type II domains, two promising domains were found (see domain 43 as well as E23 2 & 5 in fig. 5.3) which possibly contain a phylogenetic signal to confirm or reject the phylogenetic hypotheses proposed in chapter 3.

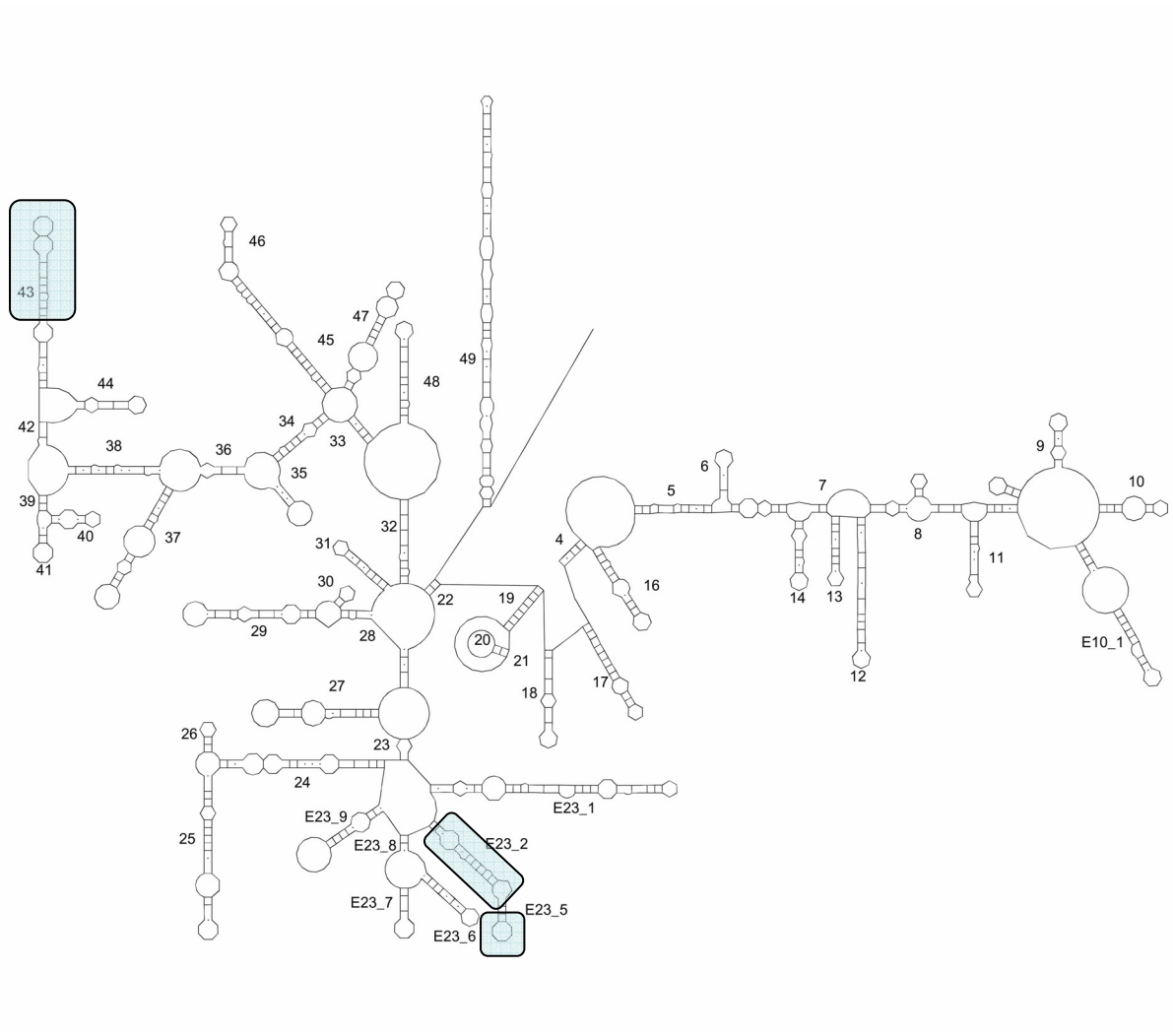
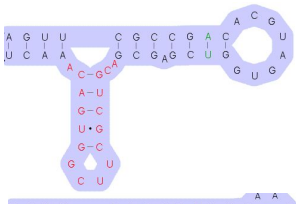
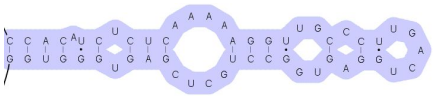
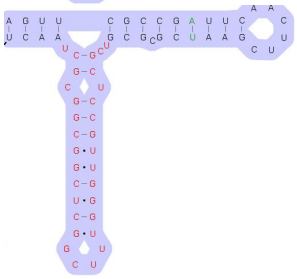
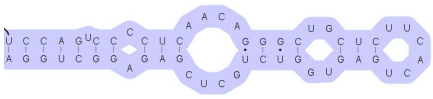
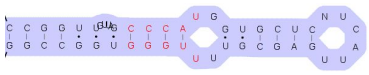
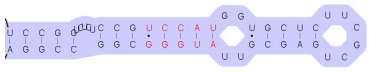
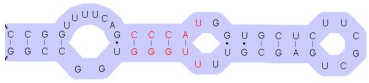
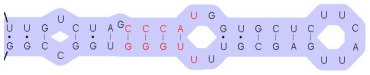
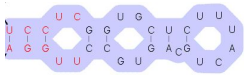
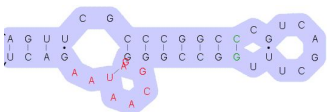
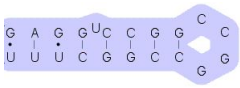
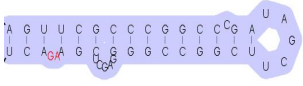
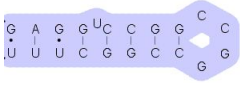
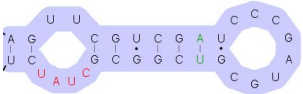
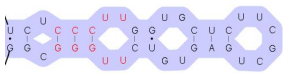
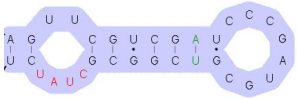
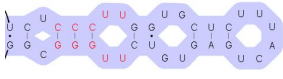
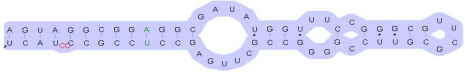


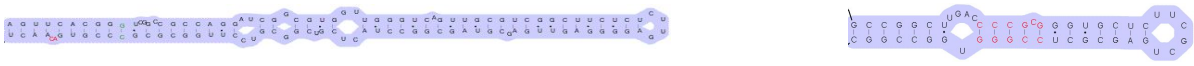
Fig. 5.3: Exemplary secondary structure model of the complete 18S rRNA of *Umbraculum umbraculum* (Opisthobranchia, Umbraculoidea); domain 43 as well as E23 2 & 5 are surrounded by a box; helix numbering according to *E. coli* (comparative RNA WebSite, Cannone et al. 2002).

Tab. 5.4: Secondary structure models of domain 43 and domain E23, 2 & 5 of 18S rRNA

Taxa	Domain 43	Domain E23, 2 & 5
Vetigastropoda		
Bathymargarites symplector		
Diodora graeca		
Caenogastropoda		
Littorina littorea		
Pomacea bridgesii		
Aperostoma palmeri		
Balcis eburnea		
Aclididae		
Larochella alta		
Graphis sp.		
"Lower Heterobranchia"		
Omalogyroidea		
Omalogyra fusca		
Omalogyra sp.		
Orbitestellidae		
Orbitestella vera		
Orbitestella sp.		
Cimidae		
Cima sp.		

Rissoelloidea

Rissoella elongatospira



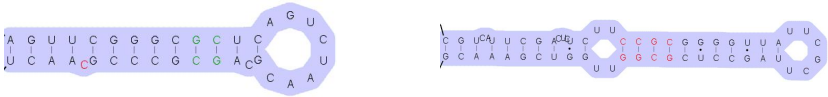
Ebalidae
Murchisonella sp.

Pyramidellidae
Odostomia sp.

Turbonilla sp.

Eulimella ventricosa

Glacidorboidea
Glacidorbis rusticus



Acteonoidea
Pupa solidula

Opisthobranchia
Nudipleura
Tomthompsonia antarctica

Bathyberthella antarctica

Chromodoris krohni

Goniodoris nodosa

Sacoglossa
Elysia viridis

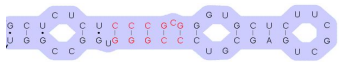
Oxynoe antillarum

Cyerce nigricans

Akeroidea
Akera bullata

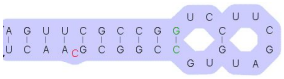
Cephalaspidea
Haminoea hydatis

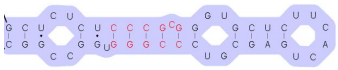
Toledonia globosa



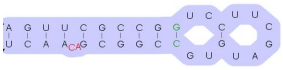
Tylodinoidea

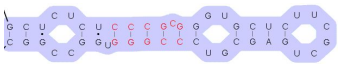
Umbraculum umbraculum



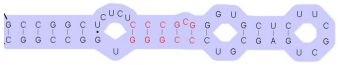


Tylodina perversa

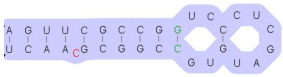




Thecosomata
Hyalocylis striata

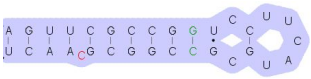


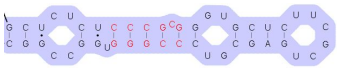
Gymnosomata
Spongiobranchaea australis



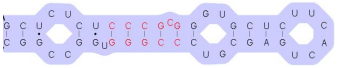
Pulmonata

Otinoidea
Smeagol phillipensis

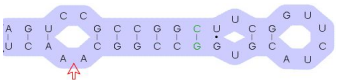


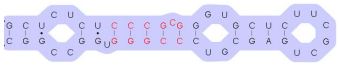


Otina ovata

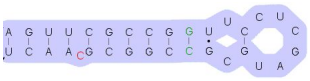


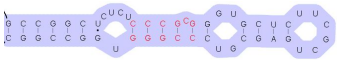
Amphiboloidea
Phallomedusa solida





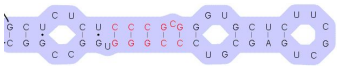
Hygrophila
Latia neritoides



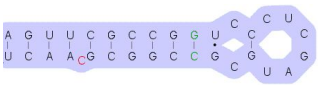


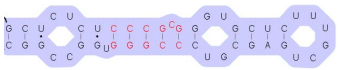
Acroloxus lacustris

Siphonarioidea
Siphonaria alternata

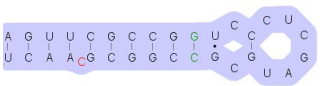


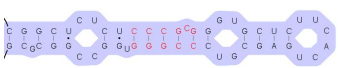
Stylommatophora
Deroceras reticulatum



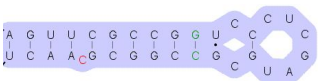


Discus rotundatus

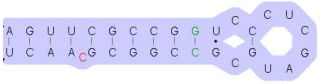


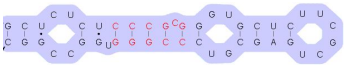


Systellomatophora
Onchidella floridana



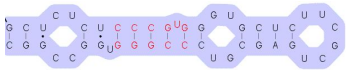
Onchidium verruculatum





Ellobioidea

Ophicardelus ornatus



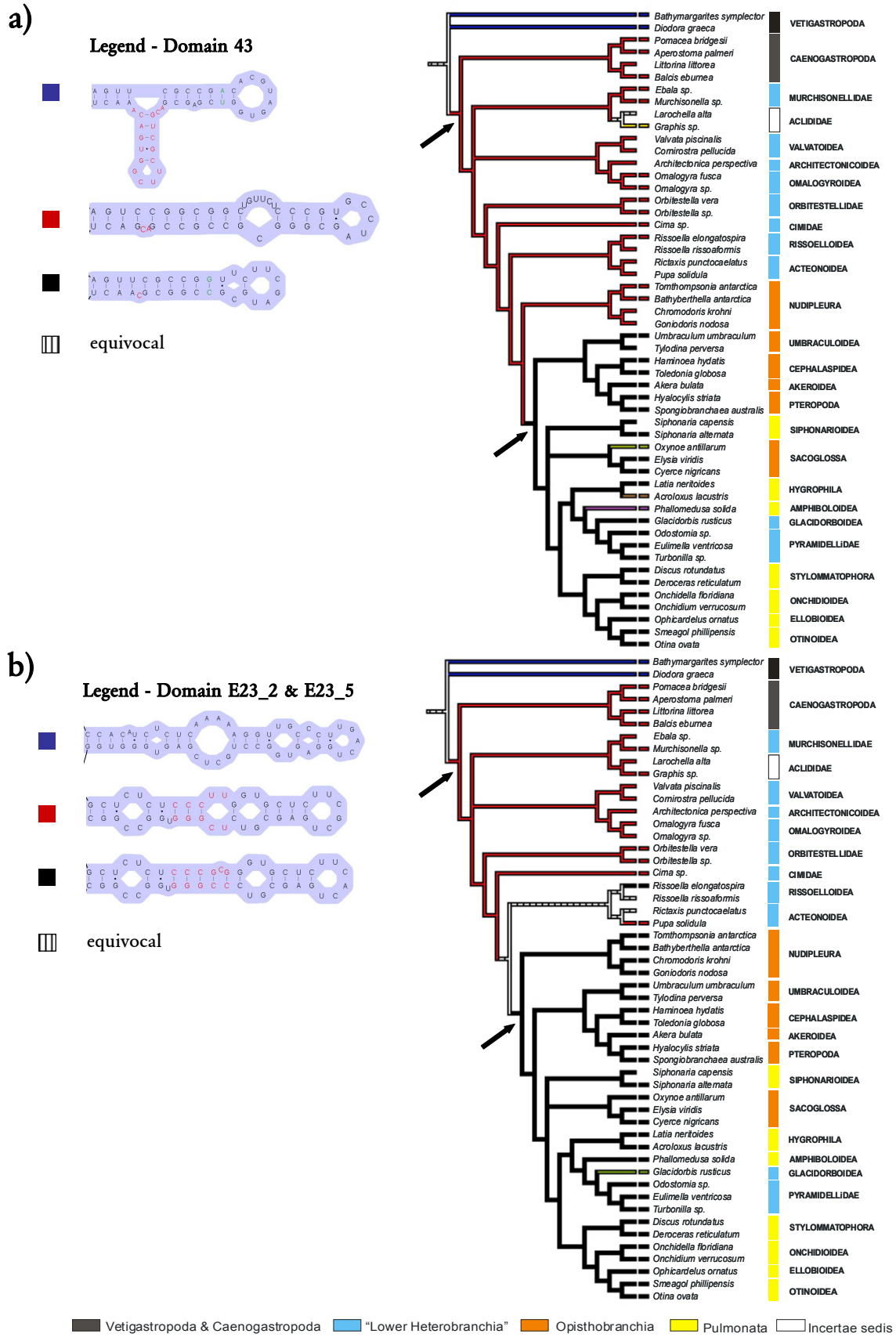


Fig. 5.4: Parsimony optimization mapping of secondary structure models of the 18S rRNA on the inferred phylogeny; a: domain 43 and b: domain E23 2 & 5; The tree is the cladogram of fig. 3.2; blue = character state 0, red = character state 1, black = character state 2, striped = equivocal; arrows mark main character state change.

Both domains show different character states (see tab. 5.4). These states were coded in a character matrix (see tab. A6 in appendix) and mapped most parsimoniously on the phylogenetic tree (see fig. 5.4) with the software MacClade 4.0 (Maddison & Maddison, 2000). To have a clear overview in fig. 5.4, only the three main character states (marked in blue, red and black) are represented in the legend by an exemplary structure.

Within domain 43 one can distinguish two different main clades (see fig. 5.4a). Clade one (red) comprises the Caenogastropoda + "Lower Heterobranchia" (without Pyramidellidae + Glacidorboidea) + Nudipleura. Within clade one, *Graphis* (Aclididae) appears in a different character state (yellow). Clade two (black) comprises the remaining Opisthobranchia + Glacidorboidea + Pyramidellidae + Pulmonata. Within clade two, *Oxynoe* (Sacoglossa) (green), *Acroloxus* (Hygrophila) (brown) and *Phallomedusa* (Amphiboloidea) (purple) appear in different character states.
Both Vetigastropoda taxa appear in the same character state (blue) which is clearly different from other states. Due to tree reconstruction procedures of the software MrBayes the Vetigastropoda are unresolved therefore it is not possible to determine this character state as a clade.
The same problem applies to Domain E23 2 & 5. The Vetigastropoda are unresolved too but both taxa appear in the same character state (blue). Nevertheless, Domain E23 2 & 5 also shows two different main clades. Clade one (red) comprises the Caenogastropoda + "Lower Heterobranchia" (without Rissoelloidea + Pyramidellidae + Glacidorboidea). Because of missing data it was not possible to determine the character state of the ancestor of Acteonoidea and Rissoelloidea. The corresponding branch is therefore shown as equivocal. Clade two (black) comprises the Opisthobranchia + Glacidorboidea + Pyramidellidae + Pulmonata. Rissoelloidea show the same character state. Within clade two, *Glacidorbis* (Glacidorboidea) shows a different character state (green).

5.3.2.2 Comparative tree reconstruction of 18S rDNA (with the software MrBayes and Phase)

Both software programs (MrBayes and Phase) are bases upon the Bayesian inference method. To avoid misunderstandings the author uses the software names instead of the method designation when comparing them with each other.

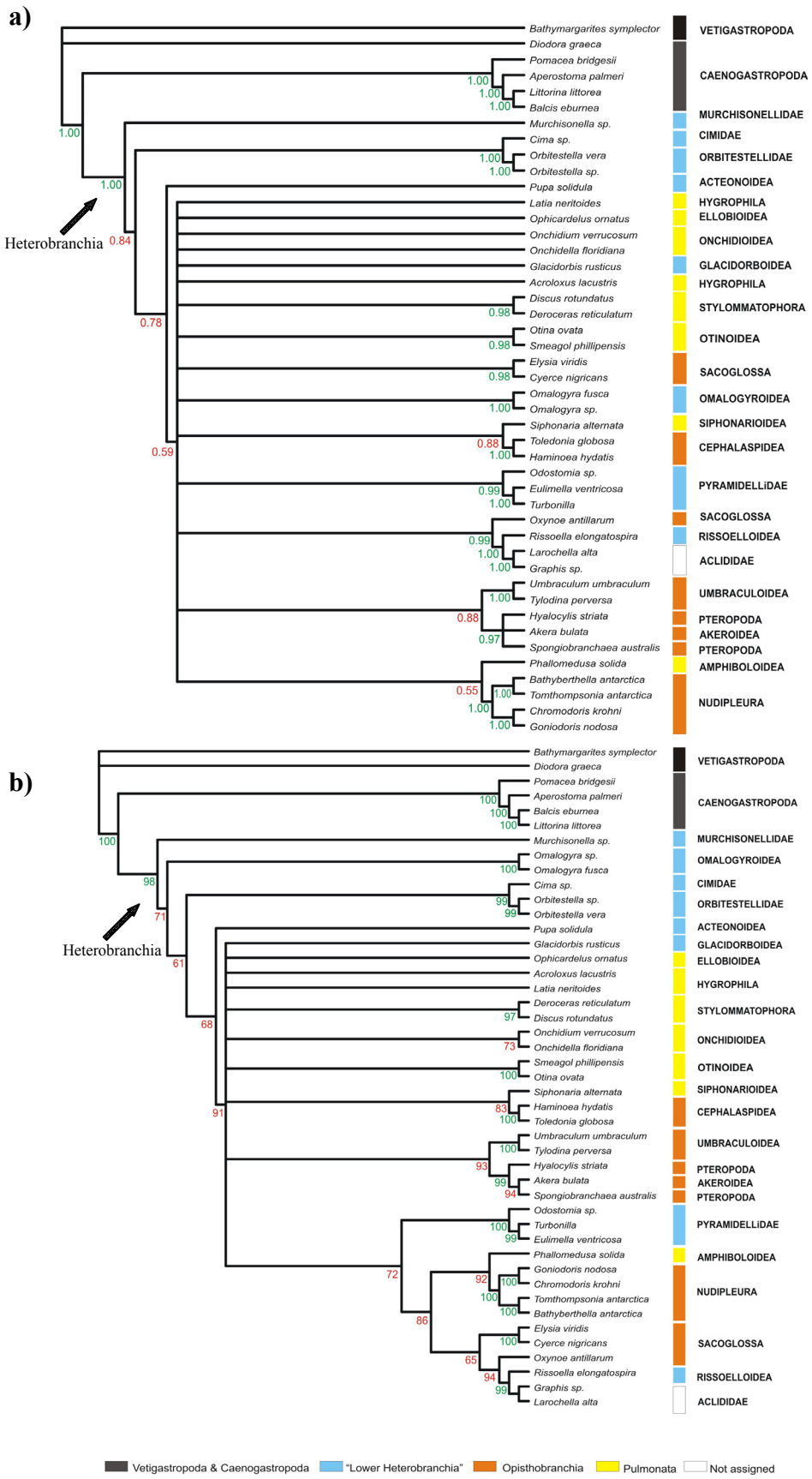


Fig. 5.5: Bayesian inference cladograms of the 50% majority rule consensus tree of dataset III using a: MrBayes and b: Phase; posterior probabilities are provided at the branches respectively (green: statistically significant, red: statistically insignificant); taxonomic classification follows Bouchet & Rocroi (2005).

The trees reconstructed with MrBayes and Phase show comparable results (see fig. 5.5 a & b). As expected, there is a good resolution regarding the deep nodes in both trees. The Caenogastropoda as well as the Heterobranchia are detected in both trees with a high statistical support.
Within the Heterobranchia the resolution is mostly poor. Taking a look at the tree conducted with MrBayes (see fig. 5.5a) one can see a comb-like structure. The following monophyla with a statistically significant support were detected: Orbitestellidae, Stylommatophora, Otinoidea, Omalogyroidea, Cephalaspidea, Pyramidellidae, Aclididae, Umbraculoidea and Nudipleura. Additionally, the following sister group relationships were found: Cimidae as sister taxon to Orbitestellidae and *Oxynoe* as sister taxon to *Rissoella* which is the sister taxon to the Aclididae.

The tree reconstructed with Phase is likewise unresolved (see fig. 5.5b). Within the Heterobranchia the resolution of deep nodes is poor. The same monophyla and most of the sister group relationships as in the topology shown in fig. 5.5a were detected. Only *Oxynoe*, representing the sister taxon to *Rissoella* and Aclididae has no statistically significant support.

5.3.3 Utility of the secondary structure of 28S rRNA for phylogenetic inference of the Heterobranchia

This chapter focuses on the same approaches as chapter 5.3.2. First of all the secondary structure of the 28S rRNA of representatives of the Heterobranchia has been reconstructed with the software RNAsalsa and browsed for domains which possibly contain synapomorphies. The identified characters are mapped most parsimoniously on the tree (see fig. 5.7) shown in chapter 5.3.3.1 which was conducted using dataset IV and MrBayes. Unlike the data of the 18S rDNA, the results could not be mapped on the tree based on dataset I because of a mismatch of many of the taxa.
Like in chapter 5.3.2.2 the next step comprises the inclusion of rRNA secondary structure information in alignment and tree reconstruction procedures. The obtained trees of the analyses with the software Phase and MrBayes are being compared.
For both approaches dataset IV containing sequences of the 28S rDNA (see A5 in the appendix) was used.

5.3.3.1 Secondary structure reconstruction of 28S rRNA

The reconstructed secondary structures of 28S rRNA comprise at least 2 domains which show possible synapomorphies to support groups within the Heterobranchia (see domain E11 and G5_1 in fig. 5.6).

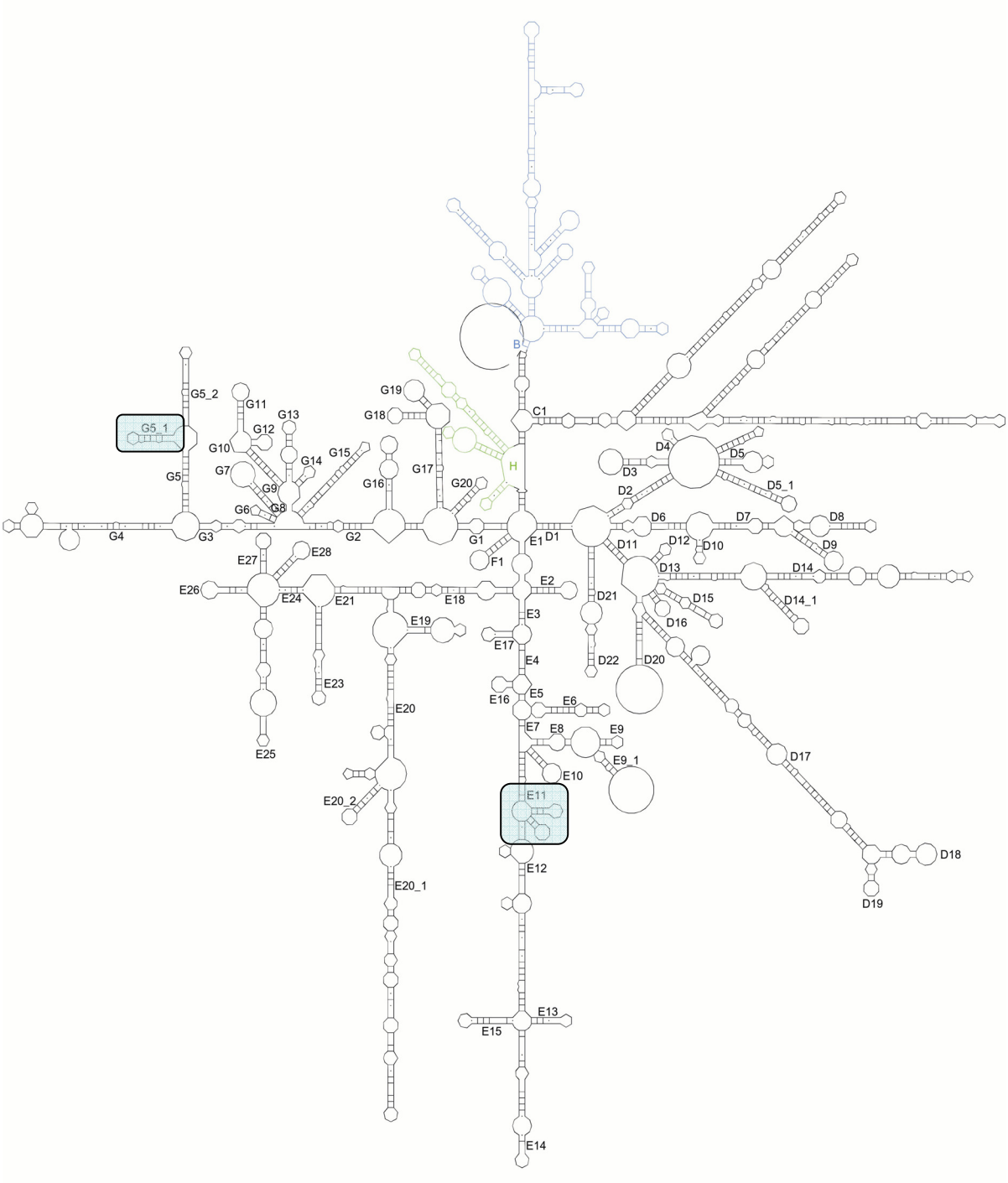
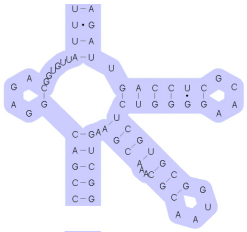
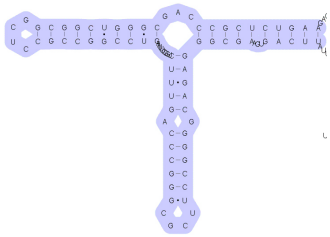
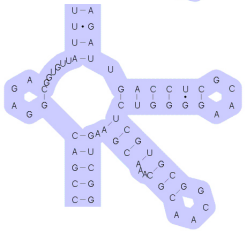
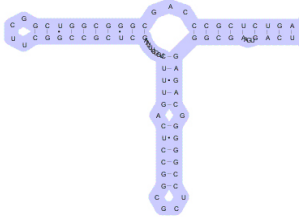
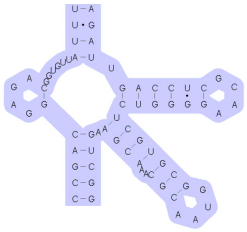
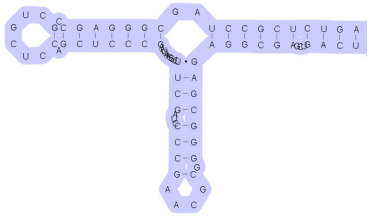
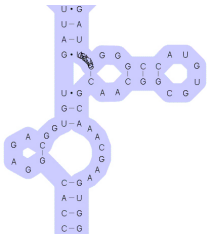
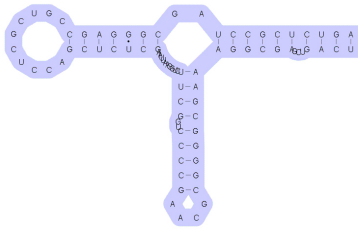
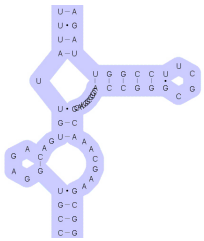
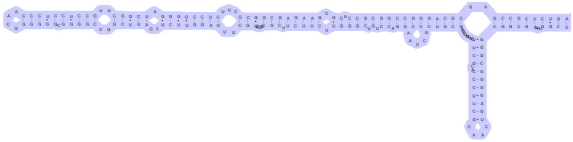
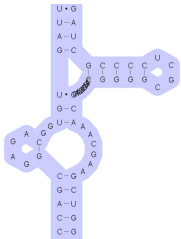
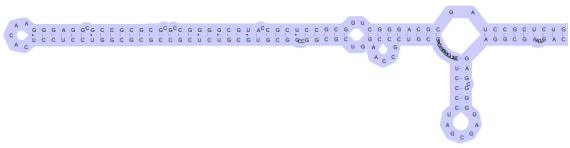
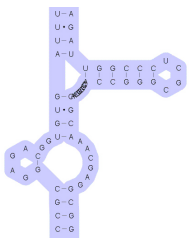
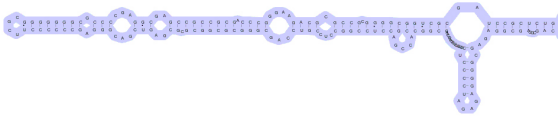


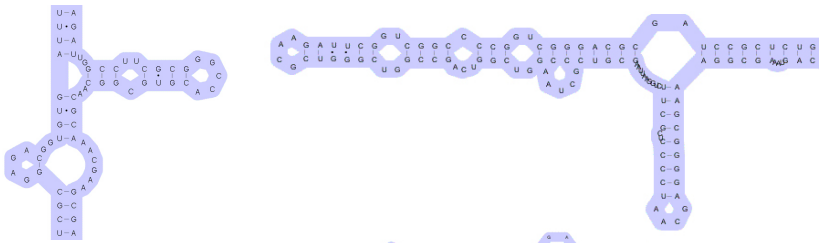
Fig. 5.6: Exemplary secondary structure model of the LSU of *Umbraculum umbraculum* (Opisthobranchia, Umbraculoidea); black = 28S, blue = 5.8S (B), green = H; promising domains are surrounded by a box; helix numbering according to *E. coli* (comparative RNA WebSite, Cannone et al. 2002).

Tab. 5.5: Secondary structure models of domain E11 and domain G5_1 of 28S rRNA

Taxa	Domain E11	Domain G5_1
Vetigastropoda		
Nordotis discus		
Lepetodrilus elevatus		
Gibbula magnus		
Caenogastropoda		
Ilyanassa obsoleta		
Aclididae		
Graphis sp.		
"Lower Heterobranchia"		
Valvatoidea		
Valvata piscinalis		
Cornirostra pellucida		

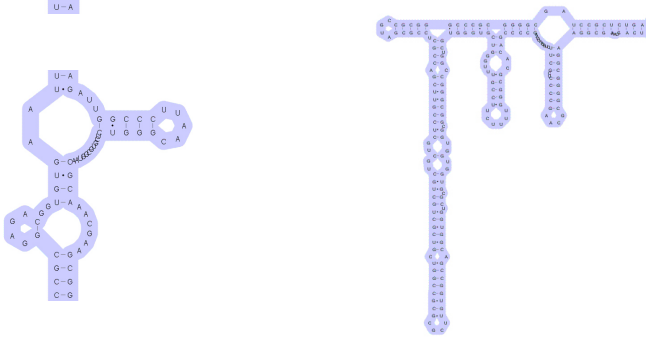
Cimidae

Cima sp.



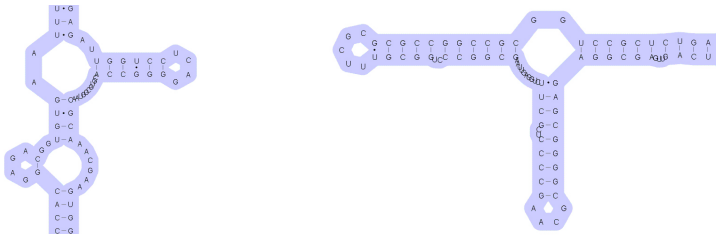
Rissoelloidea

Rissoella rissoaformis



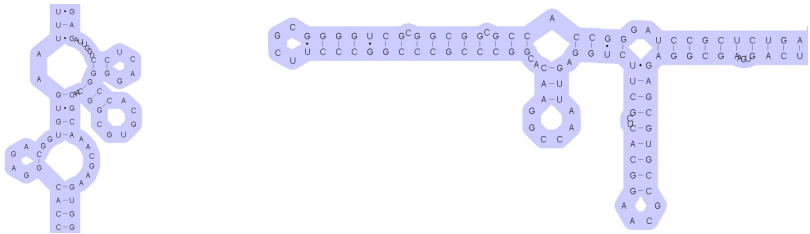
Pyramidellidae

Boonea seminuda



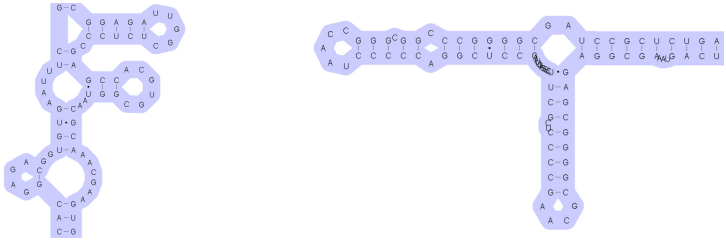
Glacidorboidea

Glacidorbis rusticus



Acteonoidea

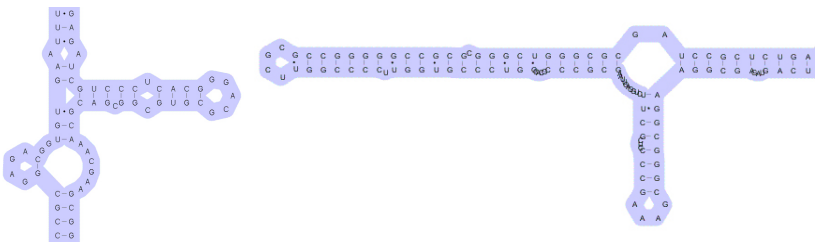
Rictaxis punctocaelatus



Opisthobranchia

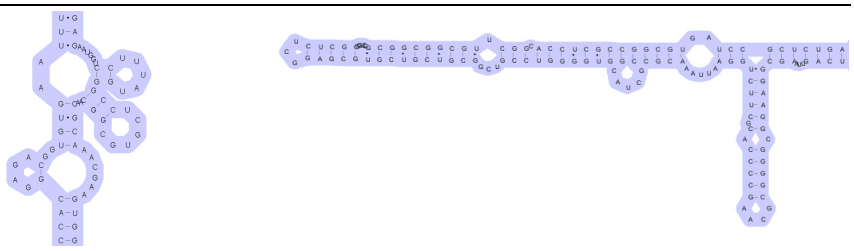
Nudipleura

Diaulula sandiegensis



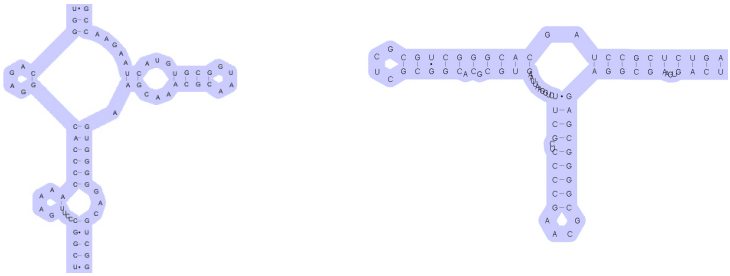
Sacoglossa

Oxynoe antillarum



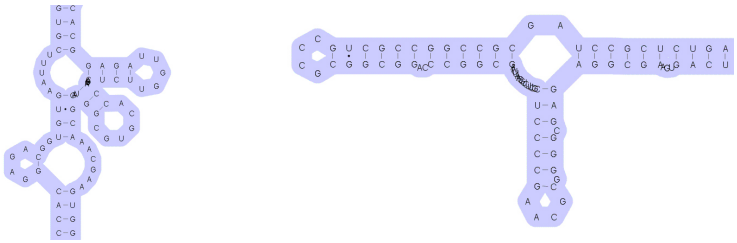
Cephalaspidea

Haminoea solitaria



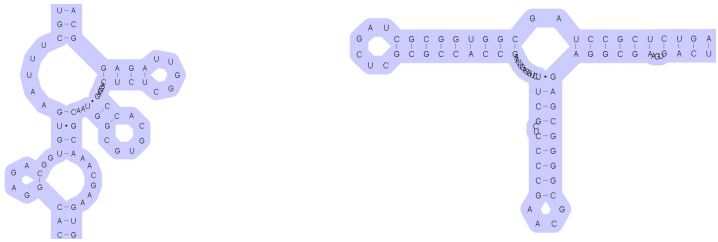
Umbraculoidea

Umbraculum umbraculum



Aplysioidea

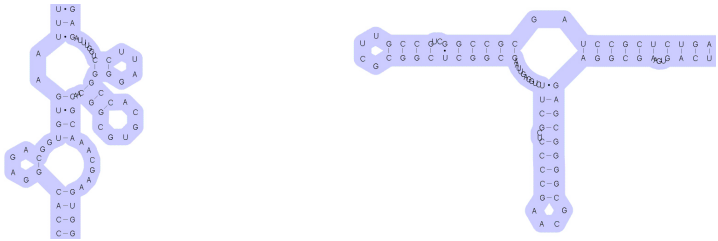
Aplysia californica



Pulmonata

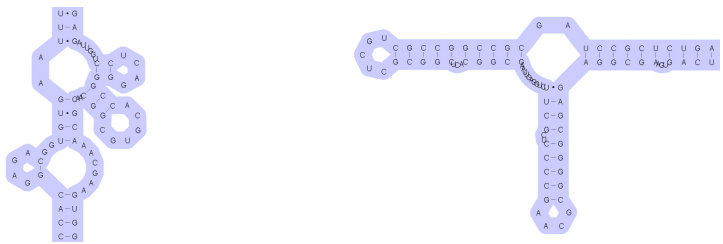
Otinoidea

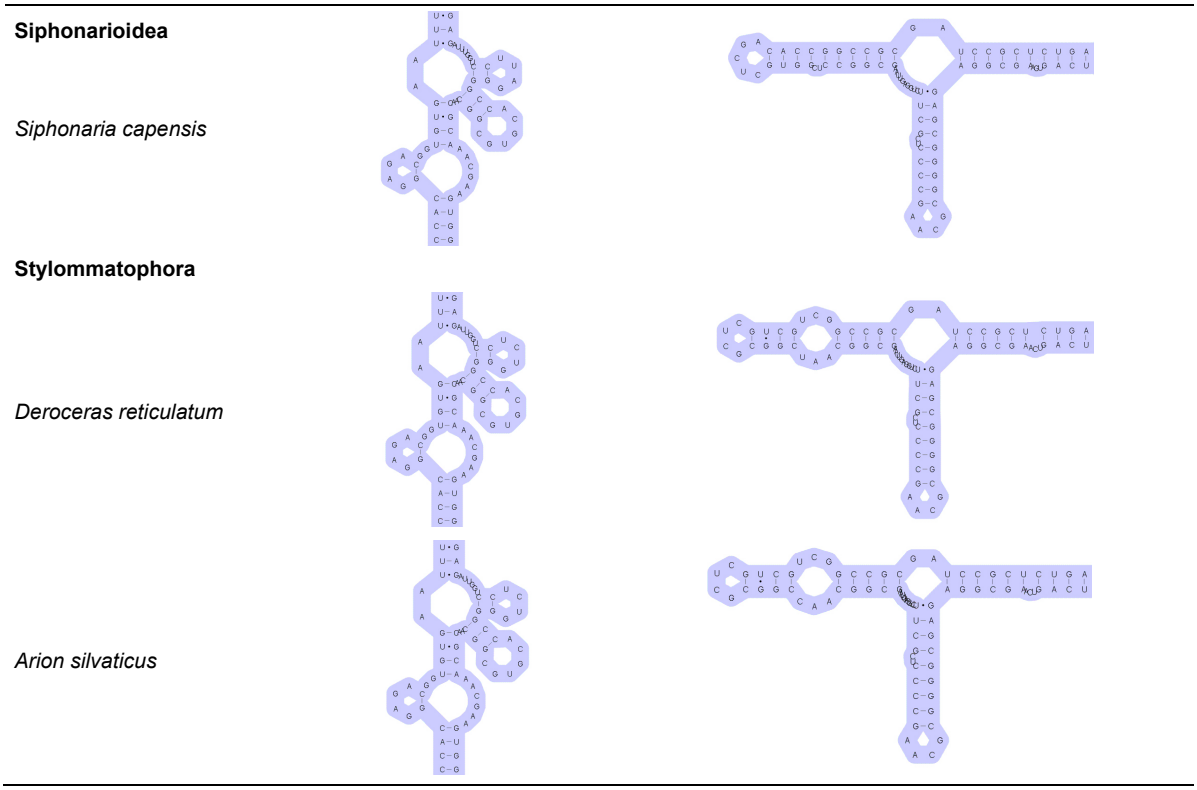
Smeagol phillipensis



Hygrophila

Latia neritoides





Like in chapter 5.3.2.1 the domains appear in different character states (see tab. 5.5) These states were also coded in a character matrix (see tab. A7 in appendix) and mapped most parsimoniously on the phylogenetic tree conducted with dataset IV and the software MrBayes (see fig. 5.7).

Within domain E11 one can distinguish three main character states (see fig. 5.7a). Character state one (blue) comprises the Vetigastropoda, state two (red) the Caenogastropoda + Cimidae + Nudipleura + Valvatoidea + Aclididae + Rissoelloidea + Acteonoidea + Pyramidellidae. State three (black) comprises the remaining Opisthobranchia + Glacidorboidea + Pulmonata while *Haminoea* (Cephalaspidea) appears in a different character state (green).

Within Domain G5_1 one can find two main character states (see fig. 5.7b). Taxa appearing in the first character state (blue) are Vetigastropoda, Caenogastropoda, Acteonoidea, Pyramidellidae, Glacidorboidea, Opisthobranchia (without Sacoglossa) and Pulmonata. Cimidae, Valvatoidea, Aclididae, Rissoelloidea and Sacoglossa appear in the second character state (red). *Glacidorbis* (Glacidorboidea) represents its own character state (black).

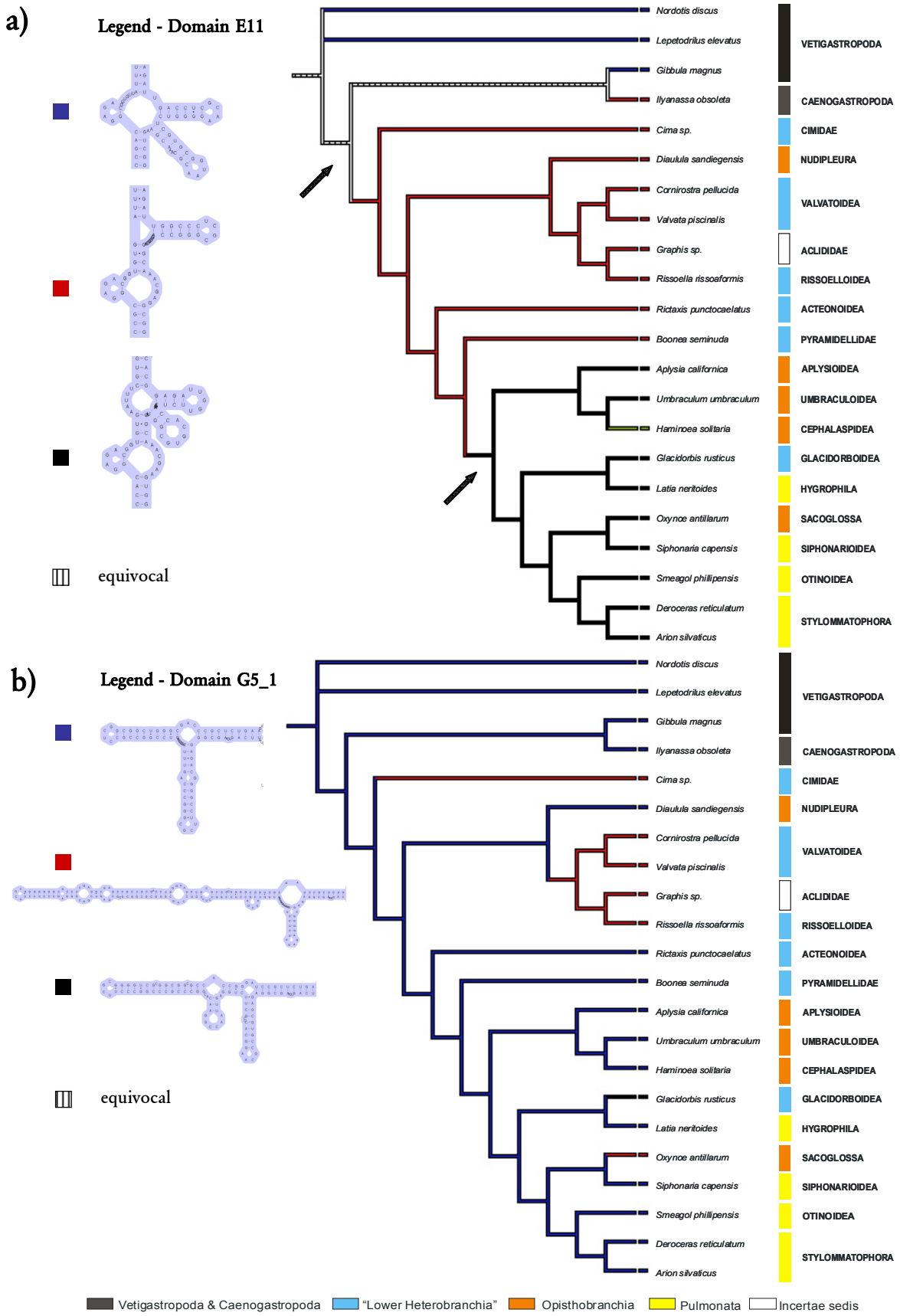


Fig. 5.7: Parsimony optimization mapping of secondary structure models on the inferred Bayesian phylogeny based on dataset IV of a: domain E11 and b: domain G5_1; blue = character state 0, red = character state 1, black = character state 2, striped = equivocal; arrows mark main character state changes.

5.3.3.2 Comparative tree reconstruction of 28S rDNA (with the software MrBayes and Phase)

Compared to the other datasets, dataset IV comprises less taxa. Therefore, most of the major groups are represented only by one taxon. Because this chapter deals with the comparison of two different tree reconstruction methods rather than proposing phylogenetic hypotheses single taxa will be discussed as major groups (e.g. *Diaulula* = Nudipleura).

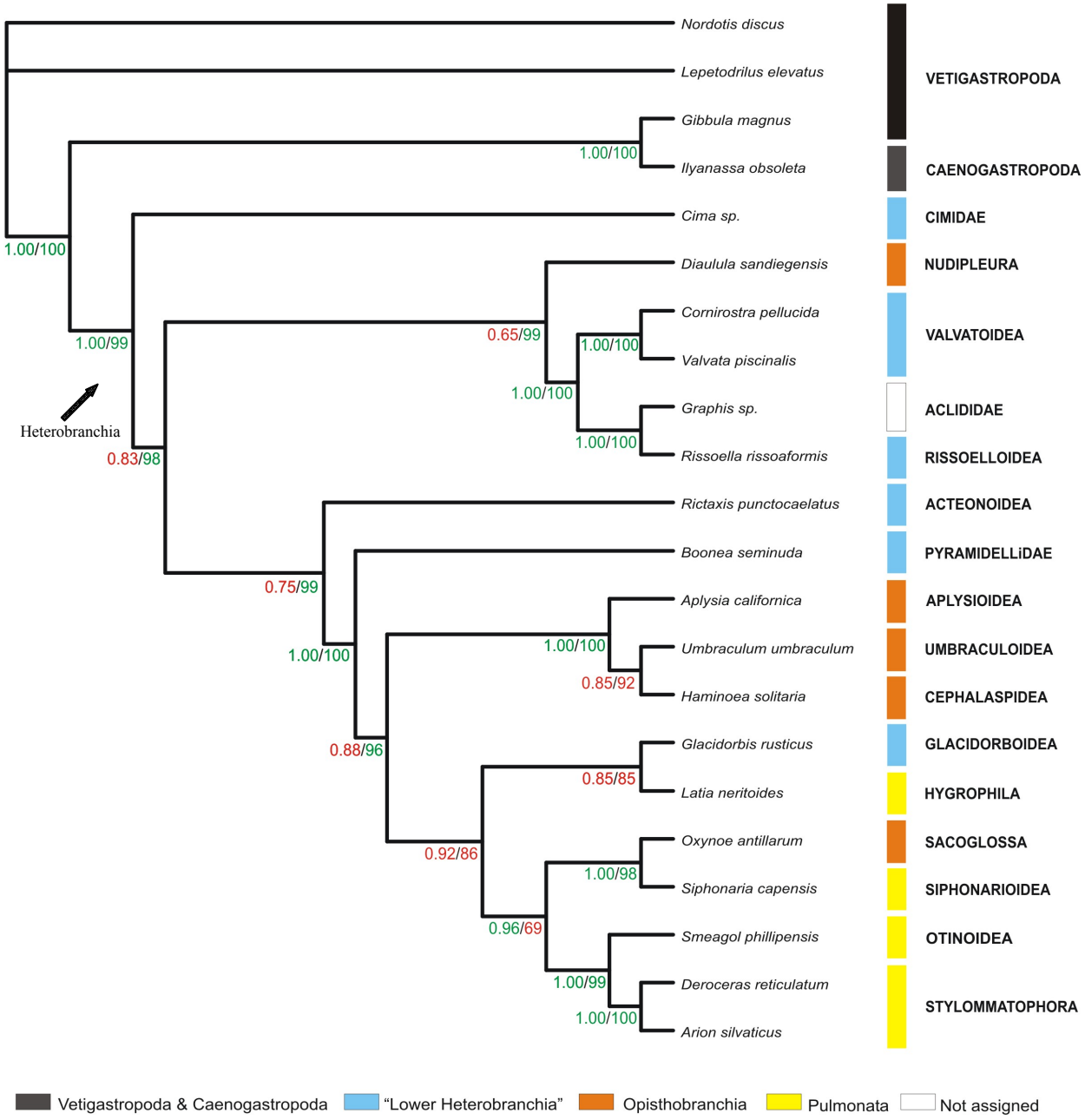


Fig. 5.8: Bayesian inference cladogram of the 50% majority rule consensus tree of dataset IV using Phase and MrBayes; Phase and MrBayes topology are the same; posterior probabilities are provided at the branches (MrBayes/Phase) (green: statistically significant, red: statistically insignificant); taxonomic classification follows Bouchet & Rocroi (2005).

The trees inferred with the two Bayesian approaches (MrBayes and Phase) show the same topology but different support values for the respective clades (see fig. 5.8). Support values are displayed in the following way (MrBayes/Phase). All posterior probabilities are above 0.50/50 but some are below 0.95/95 which renders them statistically insignificant.

The Heterobranchia are monophyletic with a high statistical support in both trees (1.00/99).
The cladogram obtained with MrBayes will be described first concerning the phylogenetic relationships within the Heterobranchia. Some of the nodes have no statistically significant support (values below 0.95/95) and will therefore not be discussed. The Cimidae appear as the first offshoot of the "Lower Heterobranchia". Only a few clades are supported like the Valvatoidea + Aclididae + Rissoelloidea as well as the clade comprising the Aplysioidea + Umbraculoidea + Cephalaspidea. A third clade comprises the Sacoglossa + Siphonarioidea + Otinoidea + Stylommatophora. Within this clade the Sacoglossa are the sister to Siphonarioidea and the Otinoidea are the sister to the Stylommatophora.
The cladogram obtained with Phase shows a better node support within the Heterobranchia. The Cimidae also appear as the first offshoot of the "Lower Heterobranchia". The next clade comprises Nudipleura + Rissoelloidea + Aclididae + Valvatoidea while Rissoelloidea are the sister taxon to Aclididae and both are sister to Valvatoidea.
The Acteonoidea are the next offshoot followed by the Pyramidelloidea. The next clade comprises the Opisthobranchia (without Nudipleura) + Pulmonata + Glacidorboidea. Within this clade Aplysioidea + Cephalaspidea + Umbraculoidea cluster together. The Siphonarioidea are the sister taxon to the Sacoglossa while the Otinoidea are the sister taxon to the Stylommatophora.

5.4 Discussion

The aim of this chapter was to evaluate the data quality *a posteriori* after the tree reconstruction by comparing congruence of statistical tests for phylogenetic signal with tree topology hypotheses proposed in chapter 3.

A posteriori evaluation of data quality by a variety of statistical tests

The Relative-Rate-Test determined high evolutionary rates within the "Lower Heterobranchia" (in particular *Omalogyra* sp., *Omalogyra fusca*, *Murchisonella* sp., *Ebala* sp. and *Architectonica perspectiva* – see Tab. 5.1) which were also visible as long branches in the 50% majority rule consensus tree (see fig. 3.2). These long branches could be a problem for tree reconstruction methods because they show a large number of substitutions which cause signal erosion. The formation of non-monophyletic groups supported mainly by analogies or convergences could be the result (Wägele 2005). Taking a look at the tree topology (especially at taxa with long branches) (see fig. 3.2) with this information in mind it seems that the high substitution rates in various "Lower Heterobranchia" had no (or if only little) influence on the tree reconstruction. Maybe, the sister group relationship of Architectonicoidea and Omalogyridae could be caused by a long branch attraction because both taxa show a long branch. However, this seems to be improbable because there is evidence in the literature for a close relationship of both taxa based on morphological characters (see chapter 3). The clade *Ebala* + *Murchisonella* was recovered as sister group to the clade *Larochella* + *Graphis*. Both clades also show a long branch but as one can see there is no statistical support for this sistergroup relationship. Hence, this result should be ignored as insignificant anyway.

In summary, one can say that evolutionary rates had no visually negative influence (in terms of misarrangements of not related taxa due to long branch attraction) on the tree reconstruction. Actually, the Relative-Rate-Test gives additional hints for an affinity of the Pyramidellidae and Glacidorboidea to the Euthyneura (as already discussed in chapter 3) due to low evolutionary rates compared to other "Lower Heterobranchia" (see tab. 5.1a+b).

However, the basal position of the Nudipleura within the Euthyneura still seems unlikely (see also discussion in chapter 3.4). The Nudipleura do not show long branches (indicating a possible long branch attraction) but show deviating sequences compared to the other more derived Euthyneura which could be the reason for them not clustering with the latter one (see fig. 3.2).

Nevertheless, due to the insignificant results of the performed AU-test alternative phylogenetic hypotheses could not be rejected confidently. Although, the unconstrained hypothesis shows the maximum likelihood (see tab. 5.2) the three constrained trees could not be discarded because the p-values of the AU-test are not smaller than the significance level of 0.05. Nevertheless, a monophyletic Euthyneura (in the traditional sense – excluding Glacidorboidea + Pyramidellidae) seems to be less probable due to a lower p-value (0.089) compared to the unconstrained p-value (0.829). A monophyletic Pulmonata and Opisthobranchia seem also unlikely because both p-values are distinctly lower than the unconstrained p-value. However, the p-values are not significant. Hence, a definitive conclusion is not possible. Reasons for this are probably the unresolved tree topology at the base of the Heterobranchia and within the Euthyneura regarding the position of the Siphonarioidea and Sacoglossa.

Although the dataset shows much conflict (indicated by many parallel edges of the same length), many of the proposed hypotheses of chapter 3 regarding the terminal branches of the "Lower Heterobranchia" are supported by splits in the neighbournet graph (fig. 5.1). Moreover, one obtains additional information particularly regarding the "Lower Heterobranchia" because one big advantage of neighbournet graphs is the possibility to represent more information than a single tree topology could (Huson & Bryant 2006) (e. g the possible relationship of Murchisonellidae or Aclididae with various "Lower Heterobranchia" taxa indicated by different split support). Network methods can extract phylogenetic signals that are missed by tree-based-methods (Huson & Bryant 2006) and give a more complete picture.
The already mentioned conflict is reflected mainly in the deep nodes. Neither the Euthyneura nor the Pulmonata or Opisthobranchia have any split support. With the exception of the Heterobranchia, none of the deeper nodes of the tree topology of chapter 3 receives any split support. Moreover, there is little or no split support for relationships within the Pulmonata (see fig. 5.1) Nevertheless, the tree topology proposed in chapter 3 (fig. 3.2) shows a significant statistical support of many subgroups indicating a good phylogenetic signal.

According to Wägele & Mayer (2007) network analyses as well as split support spectra are not meant to replace tree building methods. The spectra show only distinct conserved patterns. Many clades appearing in phylogenetic trees are not represented among the best splits. This does not mean that such clades do not exist. Instead, spectra and split networks will show whether an alignment contains distinct signals or not, whether a clade is strongly contradicted, and which clades have the best support (Wägele & Mayer 2007). The analysis conducted with

the software SAMS must be classified as a quality analysis. In this case, the stronger the support of the best compatible splits is, the higher is the probability of homology for character states in corresponding supporting positions (Wägele & Mayer 2007). In addition to phylogenetic networks the spectrum shows a ranking order of support quality and it shows splits that are excluded in the network since not all splits can be drawn in planar graph (Wägele & Mayer 2007). Within this study the software SAMS detected 25 groupings within the 1000 most frequent partitions (see tab. 5.3 and fig. 5.2) which also appear in the phylogenetic tree (fig. 3.2). All other partitions are equivalent to random combination of taxa and incompatible with the tree. These incompatible groups show how many chance similarities (noise) occur in this alignment. This implies that many of the spectral signals of the corresponding groupings are not higher than background noise independently from a high posterior probability/bootstrap support (tab. 5.3). Hence, a high support value does not necessarily mean that a clear phylogenetic signal is conserved. This was also observed by Wägele & Rödding (1998) investigating the *a priori* estimation of phylogenetic information conserved in aligned sequences.

There is obviously a high noise level in the alignment utilized in the current study. This can probably be explained by the composition of the dataset. The major part of the used sequences consists of rDNA (18S, 28S and 16S). Besides well-conserved positions, these sequences also contain variable positions. They seem to evolve relatively fast, with the consequence that the phylogenetic signals are destroyed by multiple substitution. An observation that was also made by Wägele & Rödding (1998)

Interestingly, 17 of the 25 detected partitions comprise basal groups belonging to the veti- and caenogastropods as well as lower heterobranch taxa. Seven partitions comprise opisthobranch taxa and only one partition comprises pulmonate taxa. The good support for basal groups is congruent to the observations made with the SplitsTree analyses. A possible reason for this could be the already mentioned high evolutionary rates of taxa belonging to the “Lower Heterobranchia”. Less detected partitions within Pulmonata are probably caused by an incomplete taxon sampling because most of the groups are represent by only one taxon. However, the main focus of the current study lays not on the underrepresented taxa (e.g. Stylommatophora).

Nevertheless, this result also reflects the already stated conflict in the data (particularly concerning Pulmonata and Opisthobranchia) visualized by the neighbournet graph (fig. 5.1).

Utility of the secondary structure of 18S and 28S rRNA for phylogenetic reconstruction

When working with secondary structures one should keep in mind that the inferred secondary structures must be treated with caution and can only be considered as working hypotheses (Misof & Fleck 2003). Therefore, the secondary structures amplified for this study were compared with previously published ones (e.g Erpenbeck et al. 2007, Voigt et al. 2008) to increase the confidence that the reconstructions are generally correct.

However, valuable systematic information can sometimes be achieved by analysing DNA and RNA structure (Erpenbeck et al. 2004). Such "molecular morphology" has been carried out on various invertebrate taxa like cnidarians (Ender & Schierwater 2003, Odorico & Miller 1997), sponges (Chombard & Boury-Esnault 1999) and gastropods (Lydeard et al. 2002).

This study also revealed specific structures for some lineages of Vetigastropoda and Caenogastropoda as well as the Heterobranchia within the 18S rRNA secondary structure (see tab. 5.4 and fig. 5.4). Domain 43 of the 18S molecule appears in three main character states and four further states. In general, one can say that character state 0 (blue) as well as 2 (black) are comparatively constant while within character state 1 (red) some variations appear (see tab. 5.4).

The Vetigastropoda are clearly distinguishable from the remaining taxa by a long stem region within domain 43 which has been reduced during evolution. Character state 1 (red) shows mainly two remaining base pairs and character state 2 (black) only one remaining base pair.

This domain supports the basal position of the Acteonoidea and Nudipleura as already stated in chapter 3 because both taxa show secondary structures of character state 1 (red). Furthermore, a non basal position of the Pyramidellidae and Glacidorbidae is supported because both taxa possess character state 2 (black).

Domain E23 2 & 5 of the 18S molecule also comprises three main character states (0 = blue, 1 = red, 2 = black) while *Glacidorbis* (Glacidorboidea) shows a deviating character state (3 = green). Character state 0 (blue) is defined by the following nucleotide sequence CUCAA, character state 1 (red) is defined by one of the following two nucleotide sequences CCCUU/CCCAU and character state 2 (black) is defined by one of the following three nucleotide sequences CCCGGC/CCCGCG/CCCGUG. These different nucleotide sequences consequently determine different secondary structures.

Within Domain E23 2 & 5 of the 18S molecule, the Vetigastropoda are also clearly distinguishable from the remaining taxa by their nucleotide sequence. The basal position of the Acteonoidea is also supported because it shows character state 1 (red). In contrast the

Nudipleura show a different character state (2 = black) than the more basal taxa (1 = red) which supports a non basal position of this taxon. This is in contrast to domain 43 (see above). Furthermore, like in domain 43 a non basal position of the Pyramidellidae as well as Glacidorboidea is supported because both taxa possess the derived character states 2 (black) or 3 (green).

To sum up, it can be proposed that domain 43 as well as E23 2 & 5 of the 18S molecule are adequate to separate the Vetigastropoda from the remaining taxa and to support a basal position of the Acteonoidea. A non basal position of the Pyramidellidae as well as the Glacidorboidea is also supported by both domains. A final statement about the basal position of the Nudipleura is not possible because of contrary results of domain 43 and E23 2 & 5.

Nevertheless, one should be aware of the conserved condition of the secondary structure of 18S rRNA especially within the Heterobranchia, because sometimes only one base distinguishes one character state from another. Hence, random mutations could lead to wrong results respective to a wrong interpretation of evolution.

Other studies show a less conserved secondary structure for the 18S rRNA. Voigt et al. (2008) for example investigated the complete SSU rRNA secondary structure in Porifera and found structural differences in SSU rRNA among different Porifera groups. He concluded that secondary structure features can provide alternative support for sequence-based topologies and give insights into the evolution of the molecule itself.

Many rDNA molecular phylogenetic studies result in trees that are incongruent to either alternative gene tree reconstructions and/or morphological assumptions. One reason for this outcome might be the application of suboptimal phylogenetic substitution models (Erpenbeck et al. 2007). Partitioned analyses using rDNA specific models have been reported to result in better supported tree topologies (Dohrmann et al. 2006, Erpenbeck et al. 2007, Voigt et al. 2008). Thus, it appears relevant to apply consensus structures in rDNA based phylogenies via rDNA substitution models (Misof et al. 2006). Furthermore, taking secondary and tertiary structures of rRNA genes into account seems to be a promising approach to improve homology estimation in alignments (Kjer 1995; 2004). Therefore, two different analyses with the same dataset (dataset III) of the 18S rDNA were conducted. One analysis was carried out using standard settings (one evolutionary model for the entire molecule) with the software MrBayes. The other one was conducted using the software RNAsalsa and Phase. RNAsalsa uses secondary structure information for adjusting and refining the sequence alignment. This

alignment was used for tree reconstruction with the software Phase taking specific rDNA evolutionary models for paired and unpaired bases into account.
Comparing the two obtained phylogenetic trees (fig. 5.5) it seems that the MrBayes tree shows a better resolution than the Phase tree but taking a look at the statistical support (only posterior probabilities above 0.95/95 are statistically significant) than the first impression needs to be revised. Taking only relationships with a statistically significant node support into account both trees show more or less the same phylogenetic topology. Thus, the here presented results are contrary to other comparable studies where using rDNA specific models have improved phylogenetic results (Dohrmann et al. 2006, Erpenbeck et al. 2007, Voigt et al. 2008). A reason for the missing improvement of the phylogeny could be the limited number of gene markers the tree reconstruction was based on. The 18S rDNA sequences alone does not contain enough phylogenetic signal to solve deep as well as terminal nodes independent of the used tree reconstruction method or evolutionary model.

The large nuclear ribosomal subunit (LSU) is a popular phylogenetic marker in Metazoa research with its most variable regions located in the expansion or D segments (Erpenbeck et al. 2004). Especially the domains D1-D3 of the 28S rRNA have been used for phylogenetic analyses of gastropod taxa (Dayrat et al. 2001, Klussmann-Kolb & Dinapoli 2006, Vonnemann et al. 2005, Klussmann-Kolb et al. 2008).
Within the 28S rRNA secondary structure of Vetigastropoda, Caenogastropoda and Heterobranchia particular motifs were found that are specific for some lineages (see tab. 5.5 and fig. 5.7) (see also discussion about the 18S rRNA secondary structure above).
Domain E11 (fig. 5.7a) appears in three main character states (state 0 = blue, state 1 = red, state 2 = black) while *Haminoea* (Cephalaspidea) shows a deviating character state (green). All four states are clearly distinguishable from each other by their different structure.
The Vetigastropoda are clearly distinguishable from the Caenogastropoda and Heterobranchia showing character state 0 (blue).
Domain E11 supports the already stated basal position of the Nudipleura and Acteonoidea (see discussion above) because both taxa possess character state 1 (red). In contrast to the secondary structure of the 18S rRNA the Pyramidellidae also show character state 1 which implicates a basal position for this taxon. Nevertheless, a non basal position of the Glacidorbidae is supported because *Glacidorbis* presents the derived character state 2 (black).
Domain G5_1 (fig. 5.7b) appears in two main character states (0 = blue, 1 = red) while *Glacidorbis* (Glacidorboidea) shows an aberrant character state (2 = black).

The character state change from 0 (blue) to 1 (red) within domain G5_1 is characterised by an insertion event leading to an extreme long stem region. This event happened within groups belonging to the "Lower Heterobranchia" (*Cima*, *Cornirostra*, *Valvata*, *Graphis* and *Rissoella*) and Sacoglossa. Nudipleura, Acteonoidea and Pyramidellidae show the original character state (0 = blue) which means that the results of Domain G5_1 favour the hypothesis that these three groups are not closely related to the "Lower Heterobranchia" and Sacoglossa.
At this point the question arises whether this insertion event happened only once or several times independently during evolution. If it only happened once than one must assume mistakes during tree reconstruction and the Nudipleura for example are less basal while the Sacoglossa are less derived than the phylogenetic topology implies. However, when looking at the results of chapter 5.3.3.2 (fig. 5.8) some of the posterior probabilities are below the statistical significant level of 0.95 and thus a misarrangement of the tree topology is within the realms of possibility.
Summing up, it seems that domain E11 contains a phylogenetic signal for separating the Vetigastropoda from the Caenogastropoda + Heterobranchia and groups within the Heterobranchia from each other. Domain G5_1 possibly contains a phylogenetic signal to characterise taxa belonging to the "Lower Heterobranchia". Both domains favour the idea of non basal Glacidorbidae while domain E11 favours basal Nudipleura, Acteonoidea and Pyramidellidae whereas domain G5_1 favours the latter three taxa in a non basal position.

As expected, the secondary structure of 28S rRNA is less conserved than the secondary structure of 18S rRNA.
The fact that variable structures are found in domain 43 and E23 of the 18S rRNA as well as in domain E11 and G5_1 of the 28S rRNA in various taxa indicates that these regions are under less functional constraints than are the core regions of the small and large ribosomal subunit. Wuyts et al. (2001), by investigating the tertiary structure of rRNA showed that the substitution rates are generally low near the centre of the ribosome (where the nucleotides essential for its function are situated) and that nucleotide variability increases towards the surface.

As already discussed, it seems to be a promising approach to use secondary structure information to improve the alignment and to use rDNA specific models.
The same two analyses conducted for the 18S rDNA dataset (dataset III) were carried out for the 28S rDNA dataset (dataset IV) (see fig. 5.8).

Comparing the two Bayesian phylogenetic trees (one with MrBayes and one with Phase) (fig. 5.8) both trees show the same topology but with a different statistical support. The posterior probabilities from the MrBayes analysis are often lower than support values from the Phase analysis. However, the author feels unable to decide whether the lower posterior probabilities are due to a lower or noisy phylogenetic signal or due to the different evolutionary models applied with the different softwares. Once again these results are contrary to already existing studies (Dohrmann et al. 2006, Erpenbeck et al. 2007, Voigt et al. 2008) where specific rDNA models have noticeably improved the phylogenetic results. So, one should be aware that the use of secondary structures in aligning rDNA sequences does not guarantee obtaining the correct alignment for every base. This especially counts for sequence variable loops (Buckley et al. 2000)

However, the phylogenetic tree based on dataset IV (complete 28S rDNA sequences only) has a noticeably better resolution than the phylogenetic tree based on dataset III (complete 18S rDNA sequences only) which leads to the conclusion that complete 28S rDNA sequences contain a useful phylogenetic signal to reconstruct Heterobranchia phylogeny. Because amplifying the complete 28S rDNA sequences is a time consuming and expensive procedure only few taxa belonging to different groups within the Heterobranchia were selected for amplification for the current study. In addition, all available Genbank sequences were added to the alignment. Because of this positive result more taxa belonging to the Heterobranchia should be sequenced in the future to improve the taxon sampling and to get a better impression of gastropod evolution.

A combination of the complete 18S and 28S rDNA datasets could also possibly improve the phylogenetic signal. Unfortunately, for this study, there was little taxa overlap between the two datasets. This should be taken into account in further studies.

Conclusion

The results of this chapter regarding the data quality are congruent to the conclusions made in chapter 2 (*a priori* evaluation of data quality). Much conflict was observed in the dataset. Nevertheless, the *a posteriori* investigation also supports many of the obtained results of the phylogenetic tree reconstruction of chapter 3 regarding the phylogeny of the "Lower Heterobranchia" like the affinity of the Pyramidellidae and Glacidorbidae to the Euthyneura or the sister group relationship between the Architectonicoidea and Omalogyroidea. Hence, an *a*

posteriori evaluation is helpful in estimating whether the complete alignment contains distinct or noisy signals and which clades have a strong or weak support.

Within this study the first comprehensive survey of the complete 18S rRNA and 28S rRNA secondary structures of representatives of the main lineages of the Heterobranchia was performed and evaluated how secondary structure information and features can contribute to improve phylogenetic reconstructions.
The 18S and 28S rRNA secondary structure provide valuable phylogenetic information in addition to the primary sequence. This study demonstrated that secondary structure analyses can increase the potential phylogenetic information of already available rDNA sequences because the secondary structures of both rRNA molecules show taxon specific structural variation within Vetigastropoda and Ceanogastropoda as well as Heterobranchia. Thus, the importance of 18S and 28S rRNA secondary structure information for phylogenetic reconstruction is still generally underestimated, at least among Gastropoda.

It is well established that phylogenetic methods perform better when the model of evolution is appropriate (Sullivan & Swofford 1997, Posada & Crandall 2001). This also concerns the specific rDNA models which have noticeable improved phylogenetic results (Dohrmann et al. 2006, Erpenbeck et al. 2007, Voigt et al. 2008). So far, the automatisation of alignment procedures using secondary structures failed due to inadequate formalizations of the alignment process and moreover because of difficulties in generating secondary structure models for rDNA sequences (Higgs 2000). This study shows that the new phylogenetic software RNAsalsa is a fast and capable tool to reconstruct secondary structures even if a significant improvement of the phylogenetic results of this study could not be recognized although specific rDNA evolutionary models were used.
Nevertheless, improvement and innovation of phylogenetic reconstruction methods is essential to advance the reconstruction and thus the understanding of phylogenetic and evolutionary processes. Better partitioned analyses and refined evolutionary models will certainly bring us closer to this goal.

6. General conclusion

The present study focused on two major goals. The first goal was the reconstruction of the Heterobranchia phylogeny and evolution based on molecular data whereas the main focus was on the "Lower Heterobranchia". The second goal comprised the adoption of newly developed approaches in phylogenetic inference. These approaches include the detection of ambiguously aligned positions in sequence alignments with the software Aliscore, consideration of rRNA secondary structures in tree reconstruction and alignment procedures with the software RNAsalsa, and the application of specific rDNA substitution models with the software Phase. Moreover, a case study using "relaxed" molecular clock approaches to estimate divergence times within the Heterobranchia was accomplished using the software Beast.

6.1 New insights into heterobranch phylogeny and evolution

Due to an outstanding taxon sampling the proposed phylogenetic hypothesis enables many new insights into heterobranch phylogeny and evolution. Various important "lower" heterobranch groups which have not received much attention in former morphological and molecular investigations (e.g. *Omalogyra*, *Rissoella*, *Orbitestella*, *Glacidorbis, Ebala, Murchisonella*) as well as additional members of several groups with uncertain systematic affinities (e.g. *Larochella*, *Graphis*), were included for the first time in a phylogenetic approach based on molecular data. Mikkelsen (2002) already stated that many phylogenetic analyses are biased towards Pulmonata and Opisthobranchia due to an unbalanced taxon sampling. Until the present study, insufficient "Lower Heterobranchia" taxa have been included in former studies to receive a reliable phylogenetic hypothesis for the main lineages within the Heterobranchia.

The monophyly of the Heterobranchia which was already proposed based on morphological data could be confirmed in the present study. Within the Heterobranchia, many new findings concerning the "Lower Heterobranchia" as well as the Opisthobranchia and Pulmonata were recovered.

The inclusion of the two murchisonellid taxa *Ebala* and *Murchisonella* as well as three pyramidellid taxa (*Odostomia*, *Eulimella* and *Turbonilla*) in the current study shows new insights into Pyramidelloidea phylogeny. The monophyly of the Pyramidelloidea was rejected while a derived position of the Pyramidellidae within the Euthyneura seems ever more likely. The same applies for the inclusion of *Glacidorbis* in the phylogenetic analyses rejecting a basal position of the Glacidorboidea as proposed by Haszprunar (1988) and Dayrat & Tillier (2002). In fact, the current study supports a pulmonate relationship of *Glacidorbis* with an affinity to the Amphiboloidea as proposed by Ponder (1986) and Ponder & Avern (2000).

The inclusion of further "Lower Heterobranchia" taxa and the resultant outcomes of the present study, like the sister group relationship between the Acteonoidea and the Rissoelloidea, gives also new research impulses. Further research must be undertaken to clarify whether this newly resolved sister group relationship is based on tree reconstruction artefacts or represents true relationships supported by further data and potential synapomorphies.

The inclusion of taxa other than "Lower Heterobranchia", like the pulmonate taxon *Smeagol*, provides additional insights into gastropod phylogeny. Within the current study, *Smeagol* appears as sister group to *Otina* (Otinidae) as already proposed by Tillier (1984), Tillier & Ponder (1992) and Barker (2001) based on morphological characters. *Smeagol* is not related to the Onchidioidea as proposed by Haszprunar & Huber (1990) and Nordsieck (1992) based on the nervous system.

Hence, the present study yields an important contribution for the better understanding of the phylogeny and evolutionary history of Gastropoda. Furthermore, the recovered phylogeny provides the basis for further comparative studies within the Gastropoda.

The data of the present study were evaluated *a priori* to tree reconstruction in order to detect the data with the most appropriate phylogenetic signal for investigating the Heterobranchia phylogeny. Furthermore, an estimation of data quality was made to get an idea of how reliable the expected results of the tree reconstruction will be.
This proved advantageous when it came to deciding which inference method would be the best for reconstructing heterobranch phylogeny. The *a priori* evaluation revealed rate heterogeneity, some deviating base composition and much conflict in the dataset possibly due

to noisy nucleotide positions. For this reason, I decided to use only model based inference methods to compensate possible reconstruction difficulties.

More and more tools are available to investigate data independent of tree reconstruction methods (see also discussion in chapter 6.2). A more significant picture of phylogenetic relationships and evolutionary events can be shown. Hence, *a priori* evaluation of data should be accomplished in all future studies.

Some of the results introduced in the current molecular study are contrary to already existing hypotheses based primarily on morphological but also molecular data. The derived position of the Pyramidellidae and Glacidorbidae as well as the resulting paraphyly of the Euthyneura demonstrates this. The *a posteriori* evaluation of the data helped the author to prove the plausibility of the new phylogenetic hypothesis obtained via tree reconstruction methods and accomplished by various statistical tests as well as secondary structure reconstruction methods (see also discussion in chapter 6.2). The received additional outcomes complement the results obtained with tree reconstruction methods. This practical approach should be considered in future studies.

6.2 Novel methodological approaches using newly invented software

Many of the software packages used in the present study were applied for the first time to answer questions regarding gastropod phylogeny and evolution. Therefore, a brief review about the utility of these (partially) newly invented methods will be given in this chapter.

The software Aliscore (Misof & Misof in press) was introduced by Bernhard Misof and his former working group at the Forschungsmuseum König in Bonn. This method identifies random similarity within multiple sequence alignments based on a Monte Carlo resampling. Random similarity of sequences can impact phylogenetic reconstruction as well as interfere negatively with the estimation of substitution model parameters. The identification and removal of possible random similarity in advance of model estimation as well as tree reconstruction is therefore recommended. After testing this novel invented software I believe that Aliscore is a promising tool to examine sequence alignments *a priori*. Although I decided to use the by eye reviewed dataset, the dataset modified by Aliscore achieved in most of the *a priori* investigations (e.g. Chi-Square-Test) better results than the original alignment (see also

discussion in chapter 2.4). Though, in the case of the very heterogeneous Heterobranchia sequences a more conservative evaluation of the sequence alignment was necessary to make sure that most of the random similarities were excluded. Nevertheless, Aliscore provides an automised and more objective evaluation and should therefore be utilised in further analyses.

Two different split network analyses (split decomposition and split support spectra) were used to visualize variations in signal distinctness. The split decomposition was applied with SplitsTree 4.10 (Huson & Bryant 2006) and the split support spectra with SAMS 1.4 beta (Mayer & Wägele 2005).

Both tools are appropriate for the examination of data quality *a priori* as well as *a posteriori* because the networks of supporting positions can be generated without reference to any tree topology. But the most important advantage of network analyses approaches is the possibility to visualize various possible evolutionary scenarios rather than only one evolutionary pathway like a tree topology does (Huson & Bryant 2006).

The network approaches are not meant to replace tree building methods but will show whether an alignment contains distinct signal or not and which clades have the best support (Wägele & Mayer 2007).

Regarding the results of the current study, a strong split support of the sister group relationship between the Architectonicoidea and Omalogyroidea can be seen. This split support is indicated by long parallel edges in the neighbournet graph (fig. 5.1) supporting the results of the tree reconstruction (fig. 3.2). Moreover, this cluster receives also very good support in the split support spectrum (the second strongest split in fig. 5.2). All three results together are strong arguments to propose a phylogenetic relationship between the Architectonicoidea and the Omalogyroidea based on molecular data. These findings are contrary to recently published results of Bäumler et al. (2008) who studied the anatomy of the taxon *Omalogyra atomus* via 3-D reconstruction and concluded that a closer relationship with the Architectonicoidea is unlikely (see also discussion in chapter 3.4).

Another example for the utility of network analyses concerning data evaluation is the support of the recovered polyphyly of the Pyramidelloidea in the current phylogenetic tree (fig. 3.2). The monophyly of this taxon also gains no support in the neighbournet graph (fig. 5.1) where Murchisonellidae and Pyramidellidae occupy positions far away from each other and share no split support at all.

The software package RNAsalsa (Stocsits et al. submittet) was introduced by Roman Stocsits in cooperation with the working group of the Forschungsmuseum König in Bonn. RNAsalsa can be used to reconstruct secondary structures as well as to adjust and refine sequence alignments automatically and is therefore a less time consuming approach compared to previous reconstruction methods. Hence, in the present study little computing time was needed (less than three days for the folding procedure – depending on computer power) to reconstruct the complete 18S rRNA secondary structures of 45 taxa and the complete 28S rRNA secondary structure of 22 taxa. Taxon specific secondary structure motifs were found in the 18S rRNA as well as 28S rRNA which contribute additional important information to the phylogenetic hypothesis achieved with traditional tree reconstruction methods. These findings improve at least the understanding of the phylogeny and evolution of the Heterobranchia. The secondary structures of both domains (domain 43 and E23 2 & 5, see fig. 5.4) of 18S rRNA for example, support a derived position of the Pyramidellidae and Glacidorboidea within the Euthyneura according to the tree reconstruction (fig. 3.2) of the current study.

Moreover, especially the secondary structure characters of the 18S rRNA (domain 43 and E23 2 & 5, see fig. 5.4) and 28S rRNA (domain E11 and domain G5_1, see fig. 5.7), separating the Vetigastropoda from the Caenogastropoda and Heterobranchia, were clearly distinguishing. Hence, it can be assumed that RNAsalsa will also be an appropriate algorithmic framework to reconstruct the secondary structure of higher taxonomic levels such as Gastropoda or even Mollusca.

Additionally to the facility of secondary structure reconstruction, RNAsalsa provides informations about paired and unpaired base pairs. These informations help to refine tree reconstruction methods by optimising rDNA specific evolutionary models. Using appropriate rDNA models have improved phylogenetic reconstruction approaches in earlier studies (for a discussion see also chapter 5.4) but not in the present study. A significant improvement of the phylogenetic reconstruction was missing even though rDNA specific evolutionary models were used. Reason for that is possibly the inability of one molecular marker rather than a combination of different markers to reconstruct Heterobranchia phylogeny due to missing phylogenetic signal. However, the further improvement of phylogenetic reconstruction tools as well as the refinement of evolutionary models has to be the aim. To date an incredible amount of data has been generated, possibly already containing the answers we are still

searching for. Maybe we just do not know how to extract the appropriate information because of improper phylogenetic methods.

The software package Beast (Drummond & Rambaut 2007) is a program for Bayesian MCMC analysis of molecular sequences using strict or relaxed molecular clock models.
The results of the present investigations on the evolution of the Heterobranchia using the software Beast have to be seen as preliminary mainly due to large 95% confidence intervals (CI) (see also discussion in chapter 4.4). Further analyses are necessary readjusting the prior settings by examining the fossil record in order to provide even more reliable estimates for clock calibration. In addition, the taxon sampling has to be extended because some taxa are underrepresented to answer specific questions such as the occurrence of the first Pulmonata or the radiation of the Stylommatophora.
Nevertheless, molecular clock methods are now more sophisticated than they were a few years ago and it seems that “relaxed” clock methods with their sensible date estimates complement the fossil record as our guide to evolutionary history.

7. Outlook

7.1 Inclusion of additional taxa

Although the present study provides the most extensive taxon sampling of the "Lower Heterobranchia" there are still important taxa missing to answer different phylogenetic questions regarding the Heterobranchia.

The inclusion of the Amathinidae (Pyramidelloidea) could shed light on the phylogenetic position of the Pyramidellidae in the system of the Heterobranchia as well as the monophyly of the Pyramidelloidea.

The same applies for the minute deep water gastropods of Xylodisculidae Waren, 1992 and the minute hydrothermal vent gastropods of Hyalogyrinidae Warén & Bouchet, 1993. According to Bouchet & Rocroi (2005) the Xylodisculidae (like the Orbitestellidae) are unassigned to any superfamily yet and the Hyalogyrinidae are assigned to the Valvatoidea. A molecular confirmation of the systematic position of both taxa is missing to date. Including both taxa in further analyses could possibly give new insights into the evolutionary relationships of the "Lower Heterobranchia" particularly Valvatoidea. Moreover, new findings regarding the phylogenetic position of the Orbitestellidae and Xylodisculidae within the Heterobranchia could possibly result.

The inclusion of further Architectonicoidea (like the Mathildidae) could provide new insights into the controversially discussed relationship between the Architectonicoidea and Omalogyroidea.

Last but not least, the inclusion of more "Lower Heterobranchia" in general could help to resolve the unsolved base of the tree topology (see fig. 3.2) presented in chapter 3.

To answer further questions regarding the occurrence of the first pulmonate more pulmonate taxa have to be included in future analyses. Carychiidae may have been present during the Late Carboniferous as proposed by Bandel (1997) but this hypothesis needs further testing. The inclusion of Carychiidae, Chilinidae and more Ellobiidae as well as Systellommatophora could give new insights into pulmonate evolution. In addition, more Stylommatophora should be included to get a better inside into pulmonate radiation during the Cretaceous.

7.2 Utility of novel analysing tools

The results of the current molecular study request the reassessment of morphological characters because of incongruence of morphological and molecular hypothesis. Due to their small size, many lower heterobranch species are not suitable for anatomical investigations by dissecting. The tinyness and complexity of e.g. central nervous and reproductive systems makes the data interpretation derived from histology difficult. Computer based 3D-reconstruction techniques have already been utilized successfully to resolve complex anatomical regions in minute gastropods such as Acochlidia (Neusser et al. 2006, Neusser & Schrödel 2007, Jörger et al. 2008), Omalogyridae (Bäumler et al. 2008) or skeneimorph gastropods (Kunze et al. 2008). Especially the software Amira (TGS Template Graphics Software, San Diego, CA) seems to be a capable tool for efficient analysis and presentation of the microanatomy of small specimens. Therefore, taxa whose systematic positions are still ambiguous like Pyramidellidae and Glacidorbidae should be re-evaluated morphologically using 3D reconstruction methods.

More and more analysing tools especially for investigating molecular data are available such as alignment programs which have also become more sophisticated over the years.
A new multiple sequence alignment program for unix-like operating systems called Mafft Version 6 (Katoh & Toh 2008) seems to be very promising in obtaining more accurate alignments in extremely difficult cases (Patrick Kück, pers. comm.) and should therefore be tested in further analyses.

In general, the amount of data will rise within the next years due to a more sophisticated automatisation of sequence amplification and the associated decreasing costs. To handle this huge amount of data new automatised computerised pipelines are needed, for example to evaluate data quality *a priori* as shown in the current study.

7.3 Novel phylogenetic markers

Over the years, comparing gene sequences have enriched our knowledge about gastropod phylogeny a lot. Nevertheless, to avoid errors in tree reconstruction one has to be aware that a large number of genes from many species have to be taken into account (Philippe & Telford 2006). In recent years, sequence capacity is increasing while sequencing costs are decreasing.

Hence, expressed sequence tags (EST's) provide a reliable alternative to common PCR-based sequencing approaches of single genes. The guideline of an EST approach is that a cDNA library is made from each taxon of the dataset, from which a few thousands clones are then sequenced (Philippe & Telford 2006). Dunn et al. (2008) recently showed, how a broad phylogenomic sampling improves resolution of the animal tree of life, using 77 taxa and 150 genes. Within gastropod phylogeny many relationships remain disputed and support for deep nodes is often low. Conducting phylogenetic approaches in future studies using EST data will possibly help to overcome these problems.

Compared to other vertebrate (e.g. fishes, amphibians, mammals) or invertebrate (e.g. arthropods) phyla, the mitochondrial genomic structure is unusually variable in the molluscan phylum (Kurabayashi & Ueshima 2000a). Mitochondrial gene contents as well as gene arrangements can vary between different molluscan groups (Kurabayashi & Ueshima 2000a). Wilding et al. (1999) demonstrated that mitochondrial gene arrangements are not only highly variable among different Mollusca, but also within the Gastropoda. Mitochondrial gene arrangements in Gastropoda exhibit high levels of variability and may provide valuable information for phylogenetic reconstruction. Over the years, mitochondrial gene arrangements have attracted the attention of evolutionary biologists as new phylogenetic markers (Boore & Brown 1998, Dowton 1999, Kurabayashi & Ueshima 2000a; b, Grande et al. 2008). Kurabayashi & Ueshima (2000b) investigated the mitochondrial genome organization of *Omalogyra atomus* ("Lower Heterobranchia") in context of gastropod phylogeny. They found unique gene order which can be regarded as synapomorphies of the Heterobranchia. This kind of data will possibly provide valuable information for phylogenetic reconstruction of the Heterobranchia. Grande et al. (2008) investigated the evolution of gastropod mitochondrial genome arrangements and already included a remarkable number of Opisthobranchia and Pulmonata. However, "Lower Heterobranchia" taxa are missing and have to be included in further analyses to complete the puzzle of Heterobranchia evolution, of which the present study has provided first important pieces.

Moreover, in the future scientists will possibly work increased with whole genomes due to the already mentioned further development of sequencing techniques as well as the possibility to handle this huge amount of data with high-performance computers and the respective software.

References

Aartsen, J. J. v. (1995) *Anisocycla monterosato*, 1880 or *Ebala leach* in Gray, 1847: That is the Question. *Bollettino Malacologico* **31**(1-4), 65-68.

Aktipis, S. H., Giribet, G., Lindberg, D. R. & Ponder, W. F. (2008) Gastropod Phylogeny: an Overview and Analysis. In: *Phylogeny and Evolution of the Mollusca* (edited by Ponder, F. W. & Lindberg, D. R.). University of California Press, Berkeley, 199- 236.

Albrecht, C. (2005) Patterns and Processes of Diversification in the Basommatophora (Gastropoda: Pulmonata). *Dissertation*.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**(3), 403-410.

Andrews, E. B. (1988) Excretory systems of molluscs. In: *The Mollusca* (edited by Wilbur, K. M.) **11**. Academic Press, San Diego, 381-448.

Bandel, K. (1988) Repräsentieren die Euomphaloidea eine natürliche Einheit der Gastropoden? *Mitteilungen des Geologisch-Paläontologischen Instituts der Universität Hamburg* **67**, 1-33.

Bandel, K. (1991) Gastropods from brackish and fresh water of the Jurassic - Cretaceous transition (a systematic reevaluation). *Berliner geowissenschaftliche Abhandlungen A* **134**(9-55).

Bandel, K. (1994) Triassic Euthyneura (Gastropoda) from St. Cassian Formation (Italian Alps) with a discussion on the evolution of the Heterostropha. *Freiberger Forschungshefte* **C**(452), 79-100.

Bandel, K. (1995) Mathildoidea (Gastropoda, Heterostropha) from the Late Triassic St. Cassian Formation. *Scripta Geologica* **111**, 1-83.

Bandel, K. (1996) Some heterostrophic gastropods from Triassic St. Cassian Formation with a discussion on the classification of the Allogastropoda. *Paläontologische Zeitschrift* **70**(3/4), 325-365.

Bandel, K. (1997) Higher classification and pattern of evolution of the Gastropoda. *Courier Forschungsinstitut Senckenberg* **201**, 57-81.

Bandel, K. (2002) About the Heterostropha (Gastropoda) from the Carboniferous and Permian. *Mitteilungen des Geologisch-Paläontologischen Instituts der Universität Hamburg* **86**, 45-80.

Bandel, K. & Riedel, F. (1994) The late Cretaceous gastropod fauna from Ajka (Bakony Mountains, Hungary). *A Revision - Annalen des Naturhistorischen Museums Wien* **96A**(1-65).

Bandel, K. & Heidelberger, D. (2002) A Devonian member of the subclass Heterostropha (Gastropoda) with valvatoid shell shape. *Neues Jahrbuch für Geologie und Paläontologie* **9**, 533-550.

Barker, G. M. (2001) Gatropods on Land: Phylogeny, Diversity and Adaptive Morpholgy. In: *The Biology of Terrestrial Molluscs* (edited by Barker, G. M.). CABI Publishing, Wallingford, 1-139.

Bäumler, N., Haszprunar, G. & Ruthensteiner, B. (2008) 3D interactive microanatomy of *Omalogyra atomus* (Philippi, 1841) (Gastropoda, Heterobranchia, Omalogyridae). *Zoosymposia* **1**, 101–118.

Benton, M. J. & Ayala, F. J. (2003) Dating the tree of life. *Science* **300**(5626), 1698-1700.

Bieler, R. (1992) Gastropod phylogeny and systematics. *Annual Review of Ecology and Systematics* **23**, 311-338.

Bieler, R., Ball, A. D. & Mikkelsen, P. M. (1998) Marine Valvatoidea - Comments on Anatomy and Systematic with Description of a New Species from Florida (Heterobranchia: Cornirostridae). *Malacologia* **40**(1-2), 305-320.

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), 668-674.

Bouchet, P. & Rocroi, J.-P. (2005) Classification and Nomenclator of Gastropod Families. *Malacologia* **47**(1-2), 1-397.

Brandley, M. C., Schmitz, A. & Reeder, T. W. (2005) Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. *Systematic Biology* **54**(3), 373-390.

Bromham, L. (2006) Molecular dates for the Cambrian Explosion: is the light at the end of the tunnel an oncoming train? *Palaeontologia Electronica* **9**(1).

Bucka-Lassen, K., Caprani, O. & Hein, J. (1999) Combining many multiple alignments in one improved alignment. *Bioinformatics* **15**(2), 122-130.

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), 565-580.

Caetano-Annolés, G. (2002) Tracing the evolution of RNA structure in ribosomes. *Nucleic Acids Research* **30**(11), 2575-2587.

Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Muller, K. M., Pande, N., Shang, Z., Yu, N. & Gutell, R. R. (2002) The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics* **3**, 2.

Caron, J. B., Scheltema, A. H., Schander, C. & Rudkin, D. (2006) A soft-bodied mollusc with radula from the Middle Cambrian Burgess Shale. *Nature* **442**(7099), 159-163.

Castresana, J. (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17**(4), 540-552.

Chase, R. (2002) *Behavior and Its Neural Control in Gastropod Molluscs*. Oxford University Press, Oxford, 336p.

Chombard, C. & Boury-Esnault, N. (1999) Good congruence between morphology and molecular phylogeny of Hadromerida, or how to bother sponge taxonomists. *Memoirs of the Queensland Museum* **44**, 100.

Climo, F. M. (1980) Smeagolida, a new order of gymnomorph mollusc from New Zealand based on a new genus and species. *New Zealand Journal of Zoology* **7**, 513-522.

Colgan, D. J., Ponder, W. F. & Eggler, P. E. (2000) Gastropod evolutionary rates and phylogenetic relationships. *Zoologica Scripta* **29**(1), 29-63.

Colgan, D. J., Ponder, W. F., Beacham, E. & Macaranas, J. M. (2003) Gastropod phylogeny based on six segments from four genes representing coding or non-coding and mitochondrial or nuclear DNA. *Molluscan Research* **23**, 123-148.

Colgan, D. J., Ponder, W. F., Beacham, E. & Macaranas, J. (2006) Molecular phylogenetics of Caenogastropoda (Gastropoda : Mollusca). *Molecular Phylogenetics and Evolution* **42**(3), 717-737.

Dayrat, B. & Tillier, S. (2000) Taxon sampling, character sampling and systematics: how gradist presuppositions created additional ganglia in gastropod euthyneuran taxa. *Zoological Journal of the Linnean Society* **129**, 403-418.

Dayrat, B. & Tillier, S. (2002) Evolutionary relationships of euthyneuran gastropods (Mollusca): a cladistic re-evaltuation of morphological characters. *Zoological Journal of the Linnean Society* **135**, 403-470.

Dayrat, B. & Tillier, A. (2003) Goals and limits of Phylogenetics: The euthyneuran Gastropods. In: *Molecular Systematics of Mollusks* (edited by Lydeard, C. & Lindberg, D.). Smithsonian Institution Press, Washington DC, 161-184.

Dayrat, B., Tillier, A., Lecointre, G. & Tillier, S. (2001) New Clades of Euthyneuran Gastropods (Mollusca) from 28S rRNA Sequences. *Molecular Phylogenetics and Evolution* **19**(2), 225-235.

Dixon, M. T. & Hills, D. M. (1993) Ribosomal RNA secondary structure: compensatory mutations and implications for phylogenetic analysis. *Molecular Biology and Evolution* **10**, 256-267.

Dohrmann, M., Voigt, O., Erpenbeck, D. & Worheide, G. (2006) Non-monophyly of most supraspecific taxa of calcareous sponges (Porifera, Calcarea) revealed by increased taxon sampling and partitioned Bayesian analysis of ribosomal DNA. *Molecular Phylogenetics and Evolution* **40**(3), 830-843.

Donoghue, P. C. J. & Benton, M. J. (2007) Rocks and clocks: calibrating the Tree of Life using fossils and molecules. *Trends in Ecology & Evolution* **22**(8), 424-431.

Dowton, M. (1999) Relationships among the cyclostome braconid (Hymenoptera: Braconidae) subfamilies inferred from a mitochondrial tRNA gene rearrangement. *Molecular Phylogenetics and Evolution* **11**(2), 283-287.

Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. (2006) Relaxed phylogenetics and dating with confidence. *Plos Biology* **4**(5), 699-710.

Drummond, A. J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**(214), 1-8.

Dunn, C. W., Hejnol, A., Matus, D. Q., Pang, K., Browne, W. E., Smith, S. A., Seaver, E., Rouse, G. W., Obst, M., Edgecombe, G. D., Sorensen, M. V., Haddock, S. H., Schmidt-Rhaesa, A., Okusu, A., Kristensen, R. M., Wheeler, W. C., Martindale, M. Q. & Giribet, G. (2008) Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* **452**(7188), 745-749.

Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**(5), 1792-1797.

Elewa, A. M. T. (2008) *Mass Extinction*. Springer Verlag, Heidelberg, 252p.

Ender, A. & Schierwater, B. (2003) Placozoa are not derived cnidarians: evidence from molecular morphology. *Molecular Biology and Evolution* **20**(1), 130-134.

Erpenbeck, D., McCormack, G. P., Breeuwer, J. A. & van Soest, R. W. (2004) Order level differences in the structure of partial LSU across demosponges (Porifera): new insights into an old taxon. *Molecular Phylogenetics and Evolution* **32**(1), 388-395.

Erpenbeck, D., Nichols, S. A., Voigt, O., Dohrmann, M., Degnan, B. M., Hooper, J. N. A. & Worheide, G. (2007) Phylogenetic analyses under secondary structure-specific substitution models outperform traditional approaches: Case studies with diploblast LSU. *Journal of Molecular Evolution* **64**(5), 543-557.

Farris, J. S., Källersjö, M., Kluge, A. G. & Bult, C. (1994) Testing significance of congruence. *Cladistics* **10**, 315–319.

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), 294-299.

Fretter, V. & Graham, A. (1982) The Prosobranch molluscs of Britain and Denmark, Part 7 - Heterogastropoda (Cerithiopsacea, Triforacea, Epitoniacea, Eulimacea). *Journal of Molluscan Studies* **11** Suppl., 363-434.

Fryda, J. & Bandel, K. (1997) New Early Devonian gastropods from the Plectonotus (Boucotonotus)-Palaeozygopleura community in the Prague Basin (Bohemia). *Mitteilungen des Geologisch-Paläontologischen Instituts der Universität Hamburg* **80**, 1-57.

Fryda, J. & Blodgett, R. B. (2001) The oldest known heterobranch gastropod, *Kuskokwimia* gen. nov., from the Early Devonian of west-central Alaska, with notes on the early phylogeny of higher gastropods. *Bulletin of the Czech Geological Survey* **76**, 39-53.

Fryda, J., Nützel, A. & Wagner, P. (2008) Paleozoic Gastropoda. In: *Phylogeny and Evolution of the Mollusca* (edited by Ponder, F. W. & Lindberg, D.). University of California Press, Berkeley, 239-270.

Gatesy, J., DeSalle, R. & Wheeler, W. (1993) Alignment-ambiguous nucleotide sites and the exclusion of systematic data. *Molecular Phylogenetic and Evololution* **2**(2), 152-157.

Gillespie, J. J., Munro, J. B., Heraty, J. M., Yoder, M. J., Owen, A. K. & Carmichael, A. E. (2005) A secondary structural model of the 28S rRNA expansion segments D2 and D3 for Chalcidoid wasps (Hymenoptera: Chalcidoidea). *Molecular Biology and Evolution* **22**(7), 1593-1608.

Gillette, R. (1991) On the significance of neuronal gigantism in gastropods. *The Biological Bulletin* **180**, 234-240.

Golikov, A. & Starobogatov, Y. I. (1975) Systematics of prosobranch gastropods. *Malacologia* **11**, 185-232.

Gosliner, T. M. (1985) Parallelism, parsimony and the testing of phylogenetic hypothesis: The case of opisthobranch gastropods. In: *Species and Speciation* (edited by Museum, T.) **4**. Vrba, E S, Pretioria, 105-107.

Gosliner, T. M. (1991) Morphological parallelism in opisthobranch gastropods. *Malacologia* **32**, 313-327.

Gosliner, T. M. & Ghiselin, M. T. (1984) Parallel evolution in opisthobranch gastropods and its implications for phylogenetic methodology. *Systematic Zoology* **33**(3), 255-274.

Grande, C., Templado, J., Cervera, J. L. & Zardoya, R. (2004a) Molecular Phylogeny of Euthyneura (Mollusca: Gastropoda). *Molecular Biology and Evolution* **21**(2), 303-313.

Grande, C., Templado, J., Cervera, J. L. & Zardoya, R. (2004b) Phylogenetic relationships among Opisthobranchia (Mollusca: Gastropoda) based on mitochondrial *cox 1*, *trn V*, and *rrnL* genes. *Molecular Phylogenetics and Evolution* **33**, 378-388.

Grande, C., Templado, J. & Zardoya, R. (2008) Evolution of gastropod mitochondrial genome arrangements. *BMC Evolutionary Biology* **8**(61), 1-15.

Gründel, J. (1997) Heterostropha (Gastropoda) aus dem Dogger Norddeutschland und Nordpolens. III. Opisthobranchia. *Berliner geowissenschaftliche Abhandlungen E* **25**, 177-223.

Hall, T. A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**, 95-98.

Hasegawa, M. & Hashimoto, T. (1993) Ribosomal RNA trees misleading? *Nature* **361**(6407), 23.

Haszprunar, G. (1985a) The Heterobranchia - a new concept of the phylogeny of the higher Gastropoda. *Zeitschrift für Zoologie, Systematik und Evolutionsforschung* **23**, 15-37.

Haszprunar, G. (1985b) The fine morphology of the osphradial sense organs of the Mollusca. II. Allogastropoda (Architectonicidae, Pyramidellidae). *Philosophical Transactions of the Royal Society London B* **307**, 497-505.

Haszprunar, G. (1988) On the origin and evolution of major gastropod groups, with special reference to the Streptoneura. *Journal of molluscan Studies* **54**, 367-441.

Haszprunar, G. (1990) Towards a phylogenetic system of Gastropoda Part I: traditional methodology - a reply (to Bieler´s (1990) critique). *Malacologia* **32**, 195-202.

Haszprunar, G. & Huber, G. (1990) On the Central-Nervous-System of Smeagolidae and Rhodopidae, 2 Families Questionably Allied with the Gymnomorpha (Gastropoda, Euthyneura). *Journal of Zoology* **220**, 185-199.

Haszprunar, G., Schander, C. & Halanych, K. M. (2008) Relationships of Higher Molluscan Taxa. In: *Phylogeny and Evolution of the Mollusca* (edited by Ponder, F. W. & Lindberg, D.). University of California Press, Berkeley, 19-32.

Healy, J. M. (1988) Sperm morphology and its systematic importance in the Gastropoda. *Malacological Review Suppl* **4**, 251-266.

Healy, J. M. (1993) Comparative sperm ultrastructure and spermiogenesis in basal heterobranch gastropods (Valvatoidea, Architectonicoidea, Rissoelloidea, Omalogyroidea, Pyramidelloidea) (Mollusca). *Zoologica Scripta* **22**(3), 263-276.

Healy, J. M. (1996) Molluscan sperm ultrastructure: correlation with taxonomic units within the Gastropoda, Cephalopoda and Bivalvia. In: *Origin and evolutionary radiation of the mollusca* (edited by Taylor, J. D.). Oxford Press, Oxford, 99-113.

Hendy, M. D. & Penny, D. (1993) Spectral Analysis of Phylogenetic Data. *Journal of Classification* **10**, 5-23.

Hickson, R. E., Simon, C. & Perrey, S. W. (2000) The performance of several multiple-sequence alignment programs in relation to secondary-structure features for an rRNA sequence. *Molecular Biology and Evolution* **17**(4), 530-539.

Higgs, P. G. (2000) RNA secondary structure: physical and computational aspects. *Quarterly Reviews of Biophysics* **33**(3), 199-253.

Hillis, D. M. & Dixon, M. T. (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. *Quarterly Review of Biology* **66**(4), 411-453.

Holland, B. R., Huber, K. T., Dress, A. & Moulton, V. (2002) Delta plots: a tool for analyzing phylogenetic distance data. *Molecular Biology Evolution* **19**(12), 2051-2059.

Huber, G. (1993) On the cerebral nervous system of marine Heterobranchia (Gastropoda). *Journal of Molluscan Studies* **59**, 381-420.

Huelsenbeck, J. P. & Ronquist, F. (2001) MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**(8), 754-755.

Huson, D. H. & Bryant, D. (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**(2), 254-267.

Jensen, K. R. (1997) Evolution of the Sacoglossa (Mollusca, Opisthobranchia) and the ecological associations with their food plants. *Evolutionary Ecology* **11**, 301-335.

Jeremiin, L. S. (1997) K2WuLi: A program to conduct relative-rate tests from DNA sequences. *http://jcsmr.anu.edu.au/dmm/humgen.*

Jörger, K. M., Neusser, T. P., Haszprunar, G. & Schrödl, M. (2008) Undersized and underestimated: 3D visualization of the Mediterranean interstitial acochlidian gastropod *Pontohedyle milaschewitchii* (Kowalevsky, 1901). *Organisms Diversity & Evolution* **8**(3), 194-214.

Katoh, K. & Toh, H. (2008) Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics* **9**(4), 286-298.

Kiel, S. & Bandel, K. (2001) About Heterostropha (Gastropoda) of the Campanian of Torallola, Spain. *Journal of the Czech Geological Society* **46**(3-4), 319-334.

Kjer, K. M. (1995) Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: an example of alignment and data presentation from the frogs. *Molecular Phylogenetics and Evolution* **4**(3), 314-330.

Kjer, K. M. (2004) Aligned 18S and insect phylogeny. *Systematic Biology* **53**(3), 506-514.

Klussmann-Kolb, A. & Dinapoli, A. (2006) Systematic position of the pelagic Thecosomata and Gymnosomata within Opisthobranchia (Mollusca, Gastropoda) - revival of the Pteropoda. *Journal of Zoological Systematics and Evolutionary Research* **44**(2), 118-129.

Klussmann-Kolb, A., Dinapoli, A., Kuhn, K., Streit, B. & Albrecht, C. (2008) From sea to land and beyond - new insights into the evolution of euthyneuran Gastropoda (Molllusca). *BMC Evolutionary Biology* **8**(57), 1-16.

Knudsen, B., Kohn, A. B., Nahir, B., McFadden, C. S. & Moroz, L. L. (2006) Complete DNA sequence of the mitochondrial genome of the sea-slug, *Aplysia californica*: Conservation of the gene order in Euthyneura. *Molecular Phylogenetics and Evolution* **38**, 459-469.

Kollmann, H. A. & Yochelson, E. L. (1976) Survey of Paleozoic gastropods possibly belonging to the subclass Opisthobrachia. *Annalen des Naturhistorischen Hofmuseums Wien* **80**, 207-220.

Krause, J., Unger, T., Nocon, A., Malaspinas, A. S., Kolokotronis, S. O., Stiller, M., Soibelzon, L., Spriggs, H., Dear, P. H., Briggs, A. W., Bray, S. C., O'Brien, S. J., Rabeder, G., Matheus, P., Cooper, A., Slatkin, M., Paabo, S. & Hofreiter, M. (2008) Mitochondrial genomes reveal an explosive radiation of extinct and extant bears near the Miocene-Pliocene boundary. *BMC Evolutionary Biology* **8**(220), 1-12.

Kunze, T., Brückner, N., Beck, F., Hess, M. & Haszprunar, G. (2008) Skeneimorph gastropods in Neomphalida and Vetigastropoda - a preliminary report. In: *Micromolluscs: Methodological Challenges – Exiting Results* (edited by Geiger, D. L. & Ruthensteiner, B.) Zoosymposia, 119-131.

Kurabayashi, A. & Ueshima, R. (2000a) Complete sequence of the mitochondrial DNA of the primitive opisthobranch gastropod *Pupa strigosa*: Systematic implication of the genome organization. *Molecular Biology and Evolution* **17**(2), 266-277.

Kurabayashi, A. & Ueshima, R. (2000b) Partial Mitochondrial Genome Organization of the Heterostrophan Gastropod *Omalogyra atomus* and its Systematic Significance. *Venus* **59**(1), 7-18.

Lassmann, T. & Sonnhammer, E. L. (2005) Automatic assessment of alignment quality. *Nucleic Acids Research* **33**(22), 7120-7128.

Lopez, P., Forterre, P. & Philippe, H. (1999) The root of the tree of life in the light of the covarion model. *Journal of Molecular Evolution* **49**(4), 496-508.

Loytynoja, A. & Milinkovitch, M. C. (2001) SOAP, cleaning multiple alignments from unstable blocks. *Bioinformatics* **17**(6), 573-574.

Lydeard, C., Holznagel, W. E., Ueshima, R. & Kurabayashi, A. (2002) Systematic Implications of Extreme Loss or Reduction of Mitochondrial LSU rRNA Helical-Loop Structures in Gastropoda. *Malacologia* **44**(2), 349-352.

Lyons-Weiler, J., Hoelzer, G. A. & Tausch, R. J. (1996) Relative apparent synapomorphy analysis (RASA). I: The statistical measurement of phylogenetic signal. *Molecular Biology and Evolution* **13**(6), 749-757.

Maddison, D. R. & Maddison, W. P. (2000) MacClade 4. *Sinauer Associates, Sunderland, Massachusetts*.

Maddison, W. P. & Maddison, D. R. (2008) Mesquite: a modular system for evolutionary analysis. *Version 2.5 http://mesquiteproject.org*.

Mayer, C. & Wägele, J. W. (2005) SAMS (Split analysis methods) Version 1.4 beta. *Distributed by the authors*.

McGhee, G. R. (1996) *The Late Devonian mass extinction*. Columbia University Press, New York, 303.

Mikkelsen, P. M. (1996) The evolutionary relationships of Cephalaspidea s. l. (Gastropoda: Opisthobranchia): a phylogenetic analysis. *Malacologia* **37**(2), 375-442.

Mikkelsen, P. M. (2002) Shelled Opisthobranchs. In: *Advances in Marine Biology* (edited by Southward, A. J.) **42**. Academic Press, London, 67-136.

Misof, B. & Fleck, G. (2003) Comparative analysis of mt LSU rRNA secondary structures of Odonates: structural variability and phylogenetic signal. *Insect Molecular Biology* **12**(6), 535-547.

Misof, B. & Misof, K. (in press) A monte carlo approach successfully identifies randomness in multiple sequence alignments: A more objective means of data exclusion. *Systematic Biology*.

Misof, B., Rickert, A. M., Buckley, T. R., Fleck, G. & Sauer, K. P. (2001) Phylogenetic signal and its decay in mitochondrial SSU and LSU rRNA gene fragments of Anisoptera. *Molecular Biology and Evolution* **18**(1), 27-37.

Misof, B., Niehuis, O., Bischoff, I., Rickert, A., Erpenbeck, D. & Staniczek, A. (2006) A hexapod nuclear SSU rRNA secondary-structure model and catalog of taxon-specific structural variation. *Journal of Experimental Zoology B Molecular and Developmental Evolution* **306**(1), 70-88.

Mooers, A. O. & Holmes, E. C. (2000) The evolution of base composition and phylogenetic inference. *Trends in Ecology & Evolution* **15**(9), 365-369.

Mordan, P. B. & Wade, C. M. (2008) Heterobranchia II: Pulmonata. In: *Phylogeny and Evolution of the Mollusca* (edited by Ponder, F. W. & Lindberg, D.). University of California, Berkeley, 409-426.

Morgenstern, B., Goel, S., Sczyrba, A. & Dress, A. (2003) AltAVisT: comparing alternative multiple sequence alignments. *Bioinformatics* **19**(3), 425-426.

Neusser, T. P., Hess, M., Haszprunar, G. & Schrödl, M. (2006) Computer-based three-dimensional reconstruction of the anatomy of *Microhedyle remanei* (Marcus, 1953), an interstitial acochlidian gastropod from Bermuda. *Journal of Morphology* **267**(2), 231-247.

Neusser, T. P. & Schrödl, M. (2007) Tantulum elegans reloaded: a computer-based 3D visualization of the anatomy of a Caribbean freshwater acochlidian gastropod. *Invertebrate Biology* **126**, 18-39.

Newcomb, J. M., Fickbohm, D. J. & Katz, P. S. (2006) Comparative mapping of serotonin-immunoreactive neurons in the central nervous systems of nudibranch molluscs. *Journal of Comparative Neurology* **499**(3), 485-505.

Njabo, K. Y., Bowie, R. C. K. & Sorenson, M. D. (2008) Phylogeny, biogeography and taxonomy of the African wattle-eyes (Aves: Passeriformes: Platysteiridae). *Molecular Phylogenetics and Evolution* **48**(1), 136-149.

Nordsieck, H. (1992) Phylogeny and system of the Pulmonata. *Archiv für Molluskenkunde* **121**(1/6), 31-52.

Notredame, C., Higgins, D. G. & Heringa, J. (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Mololecular Biology* **302**(1), 205-217.

Nylander, J. A. A. (2004) MrModeltest v2. *Program distributed by the author*.

Odorico, D. M. & Miller, D. J. (1997) Internal and external relationships of the Cnidaria: implications of primary and predicted secondary structure of the 5'-end of the 23S-like rDNA. *Proceedings of the Royal Society of London Series B-Biological Sciences* **264**(1378), 77-82.

Passamaneck, Y. J., Schander, C. & Halanych, K. M. (2004) Investigation of molluscan phylogeny using large-subunit and small-subunit rRNA sequences. *Molecular Phylogenetics and Evolution* **32**, 25-38.

Philippe, H. & Laurent, J. (1998) How good are deep phylogenetic trees? *Current Opinion in Genetics & Development* **8**(6), 616-623.

Philippe, H. & Forterre, P. (1999) The rooting of the universal tree of life is not reliable. *Journal of Molecular Evolution* **49**(4), 509-523.

Philippe, H. & Telford, M. J. (2006) Large-scale sequencing and the new animal phylogeny. *Trends in Ecology and Evolution* **21**(11), 614-620.

Ponder, W. F. (1984) A review of the Genera of the Rissoidae (Mollusca: Mesogastropoda: Rissoacea). *Records of the Australian Museum* Suppl **4**, 1-221.

Ponder, W. F. (1986) Glacidorbidae (Glacidorbacea: Basommatophora), a new family and superfamily of operculate freshwater gastropods. *Zoological Journal of the Linnean Society* **87**, 53-83.

Ponder, W. F. (1990a) The anatomy and relationships of a marine valvatoidean (Gastropoda: Heterobranchia). *Journal of Molluscan Studies* **56**, 533-555.

Ponder, W. F. (1990b) The anatomy and relationships of the Orbitestellidae (Gastropoda: Heterobranchia). *The Malacological Society of London* **57**, 515-532.

Ponder, W. F. (1991) Marine valvatoidean gastropods - implications for early heterobranch phylogeny. *Journal of Molluscan Studies* **57**, 21-32.

Ponder, W. F. & Waren, A. (1988) Appendix. Classification of the Caenogastropoda and Heterostropha - a list of the family-group names and higher taxa. *Malacological Review* **4** Suppl., 288-328.

Ponder, W. F. & Lindberg, D. R. (1997) Towards a phylogeny of gastropod molluscs: an analysis using morphological characters. *Zoological Journal of the Linnean Society* **119**, 83-265.

Ponder, W. F. & Avern, G. J. (2000) The Glacidorbidae (Mollusca: Gastropoda: Heterobranchia) of Australia. *Records of the Australian Museum* **52**, 307-353.

Posada, D. & Crandall, K. A. (2001) Selecting the best-fit model of nucleotide substitution. *Systematic Biology* **50**(4), 580-601.

Pybus, O. G. (2006) Model selection and the molecular clock. *Plos Biology* **4**(5), e151.

Rath, E. (1988) Organization and systematic position of the Valvatidae. *Malacological Review* **4** Suppl., 194-204.

Remigio, E. A. & Hebert, P. D. N. (2003) Testing the utility of partial COI sequences for phylogenetic estimates of gastropod relationships. *Molecular Phylogenetics and Evolution* **29**, 641-647.

Robertson, R. (1967) *Heliacus* (Gastropoda: Architectonicidae) symbiotic with Zooanthinaria (Coelenterata). *Science* **156**(3772), 246 - 248.

Robertson, R. (1970) The feeding, larval dispersal, and metamorphosis of *Philippia* (Gastropoda: Architectonicidae). *Pacific Science* **24**, 55-65.

Rutschmann, F. (2006) Molecular dating of phylogenetic trees: A brief review of current methods that estimate divergence times. *Diversity and Distributions* **12**(1), 35-48.

Salvini-Plawen, L. v. (1980) A reconsideration of systematics in the Mollusca (phylogeny and higher classification). *Malacologia* **19**, 249-278.

Salvini-Plawen, L. v. (1990) Origin, Phylogeny and Classification of the Phylum Mollusca. *Iberus* **9**, 1-33.

Salvini-Plawen, L. & Steiner, G. (1996) Synapomorphies and plesiomorphies in higher classification of Mollusca. In: *Origin and evolutionary radiation of the Mollusca*. J. Taylor (Hrsg.), The Malacological Society London:, London, 29-51.

Schröder, M. (1995) Frühontogenetische Schalen jurassischer und unterkretazischer Gastropoden aus Norddeutschland und Polen. *Paläontographica, A* **238**, 1-95.

Shimodaira, H. (2002) An Approximately Unbiased Test of Phylogenetic Tree Selection. *Systematic Biology* **51**, 492-508.

Shimodaira, H. (2004) Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling. *The Annals of Statistics* **32**(6), 2616–2641.

Shimodaira, H. & Hasegawa, M. (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* **17**, 1246-1247.

Simon, S., Frati, F., Beckenbach, A., Crespi, B., Li, H. & Flook, P. (1994) Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* **87**(651-701).

Solem, A. & Yochelson, E. L. (1979) North American Paleozoic land snails, with a summary of other Paleozoic non-marine snails. *U.S. Geological Survey Professional Paper* **1072**, 1-42.

Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**(21), 2688-2690.

Stanley, S. M. (2001) *Historische Geologie: Eine Einführung in die Geschichte der Erde und des Lebens*. Spektrum Akademischer Verlag, 710p.

Staubach, S. (2008) The Evolution of the Cephalic Sensory Organs within the Opisthobranchia. *Dissertation*.

Steel, M. A., Lockhart, P. J. & Penny, D. (1993) Confidence in evolutionary trees from biological sequence data. *Nature* **364**(6436), 440-442.

Stocsits, R., Letsch, H., Hertel, H., Misof, B. & Stadler, P. F. (submitted) Accurate and efficient reconstruction of deep phylogenies from structured RNAs. *Nucleic Acids Research*.

Sullivan, J. & Swofford, D. L. (1997) Are Guinea Pigs Rodents? The Importance of Adequate Models in Molecular Phylogenetics. *Journal of Mammalian Evolution,* **4**(2), 77-86.

Swofford, D. L. (2002) PAUP*: Phylogenetic analysis using parsimony (* and other methods). 4.0Beta for Linux/UNIX. *Sinauer Associates, Sunderland, MA*.

Tavare, S. (1986) Some probabilistic and statistical problems on the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences* **17**, 262–272.

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), 1129-1136.

Thollesson, M. (1999) Phylogenetic analysis of Euthyneura (Gastropda) by means of the 16S rRNA gene: use of a fast gene for higher-level phylogenies. *Proceedings of the Royal Society of London Series B* **266**, 75-83.

Tillier, A. (1984) Relationships of gymnomorph gastropods (Mollusca: Gastropoda). *Zoological Journal of the Linnean Society* **82**, 345-362.

Tillier, A. (1989) Comparative morphology, phylogeny and classification of land snails and slugs (Gastropoda: Pulmonata: Stylommatophora). *Malacologia* **30**, 1-303.

Tillier, A. & Ponder, W. F. (1992) New species of Smeagol from Australia and New Zealand , with a discussion of the affinities of the genus (Gastropoda: Pulmonata). *Journal of Molluscan Studies* **58**, 135-155.

Tillier, E. & Collins, R. (1998) High apparent rate of simultaneous compensatory basepair substitutions in ribosomal RNA. *Genetics* **12**, 1993–2002.

Tillier, S., Masselot, M., Guerdoux, J. & Tillier, A. (1994) Monophyly of major Gastropod taxa tested form partial 28S rRNA sequences, with emphasis on Euthyneura and hot-vent limpets Peltospiroidea. *The Nautilus* **2**, 122-140.

Tillier, S., Masselmot, M. & Tillirt, A. (1996) *Phylogenetic relationships of the Pulmonate Gastropods from rRNA sequences, and tempo and age of the Stylommatophoran radiation*. Oxford Press, Oxford, 267-284.

Tracey, S., Todd, J. A. & Erwin, D. H. (1993) Mollusca: Gastropoda. In: *Fossil Record* (edited by Benton, M. J.) **2**. Chapman & Hall, London, 131-167.

Valdés, Á. (2004) Phylogeography and phyloecology of dorid nudibranchs (Mollusca, Gastropoda). *Biological Journal of the Linnean Society* **83**, 551-559.

Valdés, Á. & Lozouet, P. (2000) Opisthobranch molluscs from the tertiary of the Aquitaine basin (South-Western France), with descriptions of seven new species and a new genus. *Palaeontology* **43**(3), 457-479.

Visser, M. H. C. (1988) The significance of terminal duct structures and the role of neoteny in the evolution of the reproductive system of Pulmonata. *Zoologica Scripta* **17**(3), 239–252.

Voigt, O., Erpenbeck, D. & Wörheide, G. (2008) Molecular evolution of rDNA in early diverging Metazoa: first comparative analysis and phylogenetic application of complete SSU rRNA secondary structures in Porifera. *BMC Evolutionary Biology* **8**(69), 1-16.

Vonnemann, V., Schrödl, M., Klussmann-Kolb, A. & Wägele, H. (2005) Reconstruction of the Phylogeny of the Opisthobranchia (Mollusca: Gastropoda) by means of 18S and 28S rRNA Gene Sequences. *Journal of Molluscan Studies* **71**, 113-125.

Wade, C. M. & Mordan, P. B. (2000) Evolution within the gastropod molluscs; using the ribosomal RNA gene-cluster as an indicator of phylogenetic relationships. *The Malacological Society of London* **66**, 565-570.

Wade, C. M., Mordan, P. B. & Naggs, F. (2006) Evolutionary relationships among the Pulmonate land snails and slugs (Pulmonata, Stylommatophora). *Biological Journal of the Linnean Society* **87**(4), 593-610.

Wägele, H. (2004) Potential key characters in Opisthobranchia (Gastropoda, Mollusca) enchancing adaptive radiation. *Organsims Diversity & Evolution* **4**, 175-188.

Wägele, H. & Willan, R. C. (2000) On the phylogeny of the Nudibranchia. *Zoological Journal of the Linnean Society* **130**, 83-181.

Wägele, H. & Klussmann-Kolb, A. (2005) Opisthobranchia (Mollusca, Gastropoda) - more than just slimy slugs. Shell reduction and its implications on defence and foraging. *Frontiers in Zoology* **2**(3), 1-18.

Wägele, H., Vonnemann, V. & Wägele, J. W. (2003) Towards a phylogeny of the Opisthobranchia. In: *Molecular Systematics of Mollusks* (edited by Lydeard, C. & Lindberg, D.). Smithsonian Institution Press, Washington DC, 185-228.

Wägele, J. W. (1996) First principles of phylogenetic systematics, a basis for numerical methods used for morphological and molecular characters. *Vie Milieu*, 125-138.

Wägele, J. W. (2005) *Foundations of Phylogenetic Systematics*. Pfeil Verlag, München, 365.

Wägele, J. W. & Rödding, F. (1998) A priori estimation of phylogenetic information conserved in aligned sequences. *Molecular Phylogenetics and Evolution* **9**(3), 358-365.

Wägele, J. W. & Mayer, C. (2007) Visualizing differences in phylogenetic information content of alignments and distinction of three classes of long-branch effects. *BMC Evolutionary Biology* **7**(147), 1-24.

Warén, A. (1992) New and little known "skeneimorph" gastropods from the Mediterranean sea and the adjacent Atlantic Ocean. *Bollettino Malacologico* **27**(10-12), 149-248.

Warén, A. (1993) New and little known Mollusca from Iceland and Scandinavia. Part 2. *Sarsia* **78**(3-4), 159-201.

Warén, A. (1994) Systematic Position and Validity of *Ebala* Gray, 1847 (Ebalidae fam. n., Pyramidelloidea, Heterobranchia). *Bollettino Malacologico* **30**(5-9), 203-210.

Warén, A. & Bouchet, P. (1993) New records, species, genera, and a new family of gastropods from hydrothermal vents and hydrocarbon seeps. *Zoologica Scripta* **22**(1), 1-90.

Welch, J. J. & Bromham, L. (2005) Molecular dating when rates vary. *Trends in Ecology and Evololution* **20**(6), 320-327.

Wilding, C. S., Mill, P. J. & Grahame, J. (1999) Partial sequence of the mitochondrial genome of *Littorina saxatilis*: relevance to gastropod phylogenetics. *Journal of Molecular Evolution* **48**(3), 348-359.

Wilkinson, M. (1998) Split support and split conflict randomization tests in phylogenetic inference. *Systematic Biology* **47**(4), 673-695.

Winnepenninckx, B. M. H., Reid, D. G. & Backeljau, T. (1998) Performance of 18S rRNA in Littorinid Phylogeny. *Journal of Molecular Evolution* **47**, 586-596.

Wise, J. B. (1996) Morphology and phylogenetic relationships of certain pyramidellid taxa (Heterobranchia). *Malacologia* **37**(2), 443-511.

Wollscheid, E. & Wagele, H. (1999) Initial results on the molecular phylogeny of the Nudibranchia (Gastropoda, Opisthobranchia) based on 18S rDNA data. *Molecular Phylogenetics and Evolution* **13**(2), 215-226.

Wu, C. I. & Li, W. H. (1985) Evidence for higher rates of nucleotide substitution in rodents than in man. *Proceedings of the National Academiy of Science, USA* **82**(6), 1741-1745.

Wuyts, J., Van de Peer, Y. & De Wachter, R. (2001) Distribution of substitution rates and location of insertion sites in the tertiary structure of ribosomal RNA. *Nucleic Acids Research* **29**(24), 5017-5028.

Xia, X. (2000) Data analysis in molecular biology and evolution. Kluwer Academic publishers, Boston, 276.

Xia, X. & Xie, Z. (2001) DAMBE: data analysis in molecular biology and evolution. *Journal of Heredity* **92**, 371-373.

Xia, X., Xie, Z., Salemi, M., Chen, L. & Wang, Y. (2003) An index of substitution saturation and its application. *Molecular Phylogenetics and Evolution* **26**, 1-7.

Yang, Z. (1998) On the best evolutionary rate for phylogenetic analysis. *Systematic Biology* **47**(1), 125-133.

Yoon, S. H. & Kim, W. (2000) Phylogeny of some gastropod mollusks derived from 18S rDNA sequences with emphasis on the Euthyneura. *The Nautilus* **114**(3), 84-92.

Zhang, P., Papenfuss, T. J., Wake, M. H., Qu, L. H. & Wake, D. B. (2008) Phylogeny and biogeography of the family Salamandridae (Amphibia: Caudata) inferred from complete mitochondrial genomes. *Molecular Phylogenetics and Evolution* **49**(2), 586-597.

Zinßmeister, C. (2008) Phylogenie der Pyramidellidae. *Diploma thesis*.

Zuckerkandl, E. & Pauling, L. (1965) Evolutionary divergence and convergence in proteins. In: *Evolving genes and proteins* (edited by Bryson, V. & Vogel, H.). Academic Press, New York, 97–166.

Acknowledgments

First of all I would like to thank Annette Klussmann-Kolb for providing and supervising this work. I am very grateful for the good time in her working group and her always open door. She supported me in all ways related to my work and beyond it. I was always impressed of her leadership abilities and networking skills. Hopefully, I was able to adopt some by her features during my years as PhD student.

I like to thank Bruno Streit for being the co-supervisor of this work and the good cooperation between his working group Ecology and Evolution and the working group Phylogeny and Systematic, where I was imbedded.

Special thanks to Claudia Nesselhauf (who is the good spirit of the working group Phylogeny and Systematic) for her help with all the little and large problems occurring during my PhD and of course for taking care of my snails while I was absent.

Many thanks to the former senior PhD student Sid Staubach who was always ready to help, no matter what or when.

I also like to thank Katrin Göbbeler, who was a great research fellow and also became a friend during my PhD. Australia without you: unimaginable!

For a linguistic review of the manuscript and a very good time on the Azores, I like to thank Adrianne Jochum.

For entertaining me during the years in various ways I like to thank all former master students like Frauke Bagusche, Tim Wollesen, Simone Faller and Alen Christof.

I like to thank Alexander Weigand for helping me with the secondary structure reconstruction of 18S rRNA and Corinna Schulze for helping me with the secondary structure reconstruction of the 28S rRNA.

For investigating the Pyramidellidae in her diploma thesis I like to thank Carmen Zinßmeister.

For helping me to build up the molecular laboratory and establishing various pcr protocols I would like to thank Kerstin Kuhn, who has an enormous know how in molecular techniques.

Many thanks to Bruce Marshall (Wellington) and Richard Taylor (Leigh) for their support during my collecting trip in New Zealand as well as Marc Hasenbank for his hospitality in Wellington. Moreover, I like to thank Georg Mayer and John Healy for their support during my field trip in Australia. Georg for providing me with chemicals and John for preparing and sending me the samples (in the end it was more work than expected).

For providing me with heterobranch material I am grateful to Michael Schrödl and Enrico Schwabe from the Zoologische Staatssammlung in München (ZSM) as well as Christian Albrecht (University Gießen).

For making my life as software user easier I like to thank various people. First of all the bunch of people from Bonn. I am grateful to Harald Letsch who helped me with the software RNAsalsa. Furthermore, I like to thank Patrick Kück for his support with the software Aliscore. For assisting me with the software SplitsTree I like to thank Katharine Händeler. Moreover, I appreciate the permanent interests of Heike Wägele in my work and her hospitality during my stay in Bonn.
Special thanks to Heiko Schmidt from Vienna and his numerous emails regarding the AU-Test. For providing me with the software SAMS I like to thank Christoph Mayer from Bochum. Last but not least I like to thank Markus Pfenninger from Frankfurt for giving me a patient introduction to the software Beast.

I am deeply grateful to the German Science Foundation (DFG) for supporting my PhD project (KL 1303/4-1, 4-2). Moreover, I like to thank the German Academic Exchange Service (DAAD) for making the field trip to New Zealand possible as well as the ‘Vereinigung von Freunden und Förderern der Goethe-Universität’ for financing my flight to Australia. The ‘Herman-Wilkomm-Stiftung’ I would like to thank for funding a centrifuge for our lab.

Last but not least I like to thank my parents for their support during my studies and of course my husband Tobias Schöpfner for his encouraging love and immense support over these years.

Appendix

Tab. A1 Taxon sampling: taxonomic classification (following Bouchet & Rocroi, 2005), collecting sites, accession numbers; gene sequences amplified for this study are marked with an *.

Taxon	Abbreviation	Family	Locality	18S	28S		16S	COI
				complete	partial	complete	partial	complete
LOWER HETEROBRANCHIA								
VALVATOIDEA								
Valvata piscinalis (Müller 1774)	Vapi	Valvatidae	Lake Prespa, Macedonia	FJ917222*/ FJ91723*	FJ917224*	FJ917224*	FJ917248*	FJ917267*
Cornirostra pellucida (Laseron 1954)	Cope	Cornirostridae	Port Stephens, Australia	FJ917215*	FJ917225*	FJ917225*	FJ917249*	FJ917282*
ARCHITECTONICOIDEA								
Architectonica perspectiva (Linné, 1758)	Arpe	Architectonicidae	Dingo Beach, Australia	FJ917220*/ FJ917221*	FJ917231*	-	FJ917251*	FJ917269*
RISSOELLOIDEA								
Rissoella elongatospira Ponder, 1966	Riel	Rissoellidae	Wellington, New Zealand	FJ917203*	FJ917232*	-	-	FJ917270*
Rissoella rissoaformis (Powell, 1939)	Riri	Rissoellidae	Wellington, New Zealand	FJ917214*	FJ917226*	FJ917226*	FJ917252*	FJ917271*
OMALOGYROIDEA								
Omalogyra fusca Suter,1908	Omfu	Omalogyridae	Leigh, New Zealand	FJ917217*	FJ917233*	-	FJ917253*	FJ917272*
Omalogyra sp.	Omsp	Omalogyridae	Ahipara, New Zealand	FJ917204*	FJ917234*	-	FJ917254*	FJ917273*
PYRAMIDELLOIDEA								
Eulimella ventricosa (Forbes, 1844)	Euve	Pyramidellidae	off Gnejna Bay, Malta	FJ917213*	FJ917235*	-	FJ917255*	FJ917274*
Odostomia sp.	Odsp	Pyramidellidae	Banyuls sur mer, France	AY427526	AY427491	-	FJ917256*	FJ917275*
Turbonilla sp.	Tusp	Pyramidellidae	Pahia, New Zealand	FJ917216*	FJ917236*	-	FJ917257*	FJ917276*
Boonea seminuda (Adams, 1839)	Bose	Pyramidellidae	Genbank	-	-	AY145395	-	-
Ebala sp.	Ebsp	Murchisonellidae	Moreton Bay, Australia	FJ917218*/ FJ917219*	FJ917237*	-	FJ917258*	FJ917277*
Murchisonella sp.	Musp	Murchisonellidae	Moreton Bay, Australia	FJ917205*	FJ917238*	-	FJ917259*	FJ917278*
GLACIDORBOIDEA								
Glacidorbis rusticus Ponder, 2000	Glru	Glacidorbidae	Wilsons Promontory, Australia	FJ917211*	FJ917227*	FJ917227*	FJ917264*	FJ917284*
ACTEONOIDEA								
Rictaxis punctocaelatus (Carpenter, 1864)	Ripu	Acteonidae	California, USA	EF489346	EF489318	FJ917243*	EF489393	EF489370
Pupa solidula (Linné, 1758)	Puso	Acteonidae	Genbank	AY427516	AY427481	-	EF489319	DQ238006
UNASSIGNED TO SUPERFAMILY								
Cima sp.	Cisp	Cimidae	Port Stephens, Australia	FJ917206*	FJ917228*	FJ917228*	FJ917260*	FJ917279*
Orbitestella vera Powell, 1940	Orve	Orbitestellidae	Wellington, New Zealand	FJ917207*	FJ917239*	-	FJ917250*	FJ917268*
Orbitestella sp.	Orsp	Orbitestellidae	Genbank	EF489352	EF489377	-	EF489333	EF489397
PULMONATA								
BASOMMATOPHORA								

SIPHONARIOIDEA								
Siphonaria capensis Quoy & Gaimard, 1833	Sica	Siphonariidae	South Africa	EF489335	EF489354	FJ917244*	EF489301	EF489379
Siphonaria alternata Say, 1826	Sial	Siphonariidae	Genbank	AY427523	AY427488	-	-	-
AMPHIBOLOIDEA								
Phallomedusa solida (Martens, 1878)	Phso	Amphibolidae	Genbank	DQ093440	DQ279991	-	DQ093484	DQ093528
HYGROPHILA								
ACROLOXOIDEA								
Acroloxus lacustris (Linné, 1758)	Acla	Acroloxidae	Genbank	AY282592	EF489364	-	EF489311	AY282581
CHILINOIDEA								
Latia neritoides Gray,1850	Lane	Latiidae	Waikato, New Zealand	EF489339	EF489359	FJ917245*	EF489307	EF489384
EUPULMONATA								
ELLOBIOIDEA								
Ophicardelus ornatus (Ferussac, 1821)	Opor	Ellobiidae	Genbank	DQ093442	DQ279994	-	DQ093486	DQ093486
OTINOIDEA								
Otina ovata (Brown, 1827)	Otov	Otinidae	Genbank	EF489344	EF489363	-	EF489310	EF489389
Smeagol phillipensis Tillier & Ponder, 1992	Smph	Smeagolidae	Phillip Island, Australia	FJ917210*	FJ917229*	FJ917229*	FJ917263*	FJ917283*
SYSTELLOMMATOPHORA								
ONCHIDIOIDEA								
Onchidella floridana (Dall, 1885)	Onfl	Onchididae	Genbank	AY427522	AY427487	-	EF489316	EF489391
Onchidium verruculatum Cuvier, 1830	Onve	Onchididae	Genbank	AY427521	AY427486	-	EF489317	EF489392
STYLOMMATOPHORA								
PUNCTOIDEA								
Discus rotundatus (Müller, 1776)	Diro	Endodontidae	Frankfurt, Germany	FJ917212*	FJ917240*	-	FJ917265*	FJ917285*
LIMACOIDEA								
Deroceras reticulatum (Müller, 1776)	Dere	Agriolimacidae	Ober-Olm, Germany	AY145373.1	FJ917241*	AY145404	FJ917266*	FJ917286*
ARIONIDAE								
Arion silvaticus Lohmander, 1937	Arsi	Arioninae	Genbank	-	-	AY145392	-	-
OPISTHOBRANCHIA								
UMBRACULOIDEA								
Umbraculum umbraculum (Lightfoot, 1786)	Umum	Tylodinidae	New South Wales, Australia	AY427499	AY427457	FJ917246*	EF489322	DQ256200
Tylodina perversa (Gmelin, 1791)	Type	Tylodinidae	Genbank	AY427496	AY427458	-	-	AF249809
APLYSIOMORPHA								
AKEROIDEA								
Akera bullata Müller, 1776	Akbu	Akeridae	Genbank	AY427502	AY427466	-	AF156127	AF156143
APLYSIOIDEA								
Aplysia californica Cooper, 1863	Apca	Aplysiidae	Genbank	-	-	AY026366	-	-
THECOSOMATA								
CAVOLINIOIDEA								
Hyalocylis striata (Rang, 1828)	Hyst	Cavoliniidae	Genbank	DQ237966	DQ237985	-	-	DQ237999
GYMNOSOMATA								
CLIONOIDEA								

*Spongiobranchaea australis d'*Orbigny, 1836	Spau	Pneumodermatidae	Genbank	DQ237969	DQ237988	-	-	DQ238002
SACOGLOSSA								
PLACOBRANCHIDOIDEA								
Elysia viridis (Montagu, 1804)	Elvi	Placobranchidae	Genbank	AY427499	AY427462	-	AJ223398	DQ237994
POLYBRANCHIOIDEA								
Cyerce nigricans (Pease, 1866)	Cyni	Polybranchiidae	Genbank	AY427500	AY427463	-	EU140843	DQ237995
OXYNOOIDEA								
Oxynoe antillarum Mörch, 1863	Oxan	Oxynoeidae	Isla de Cubagua, Venezuela	FJ917441	FJ917466	FJ917466/ FJ917247*	FJ917425	FJ917483
CEPHALASPIDEA								
HAMINOEOIDEA								
Haminoea hydatis (Linné, 1758)	Hahy	Haminoeidae	Genbank	AY427504	AY427468	-	EF489323	DQ238004
Haminoea solitaria Say, 1822	Haso	Haminoeidae	Genbank	-	-	AY145408	-	-
DIAPHANOIDEA								
Toledonia globosa Hedley, 1916	Togl	Diaphanidae	Genbank	EF489350	EF489375	-	EF489327	EF489395
NUDIPLEURA								
PLEUROBRANCHOIDEA								
Tomthompsonia antarctica (Thiele, 1912)	Toan	Pleurobranchidae	Genbank	AY427492	AY427452	-	EF489330	DQ237992
Bathyberthella antarctica Willan & Bertsch, 1987	Baan	Pleurobranchidae	Genbank	AF249219	AY427453	-	Katrin	AY345027
EUDORIDOIDEA								
Chromodoris krohni (Vérany, 1846)	Chkr	Chromodorididae	Genbank	AJ224774	AY427445	-	AF249239	AY345036
*Diaulula sandiegensis (*Cooper, 1863)	Disa	Discodorididae	Genbank	-	-	AY144352	-	-
ANADORIDOIDEA								
Goniodoris nodosa (Montagu, 1808)	Gono	Goniodorididae	Genbank	AJ224783	AY014157	-	AF249226	AF249788
UNASSIGNED TO HETEROBRANCHIA								
Larochella alta Powell,1927	Laal	Aclididae	Leigh, New Zealand	FJ917208*	FJ917242*	-	FJ917261*	FJ917280*
Graphis sp.	Grsp	Aclididae	Leigh, New Zealand	FJ917209*	FJ917230*	FJ917230*	FJ917262*	FJ917281*
CAENOGASTROPODA								
Pomacea bridgesii (Reeve, 1856)	Pobr	Ampullariidae	Genbank	DQ093436	DQ279984	-	DQ093480	DQ916496
Aperostoma palmeri (Bartsch & Morrison, 1942)	Appa	Cyclophoridae	Genbank	DQ093435	DQ279983	-	DQ093479	DQ093523
Littorina littorea (Linné, 1758)	Lili	Littorinidae	Genbank	X91970	AJ488672	-	DQ093481	AY345020
Balcis eburnea (Muehlfeld, 1824)	Baeb	Eulimidae	Genbank	AF120519	AF120576	-	DQ280051	AF120636
Ilyanassa obsoleta Say, 1822	Ilob	Nassariidae	Genbank	-	-	AY145411	-	-
VETIGASTROPODA								
Diodora graeca (Linné, 1758)	Digr	Fissurellidae	Genbank	AF120513	DQ279980	-	DQ093476	AY923915
Bathymargarites symplector Warén & Bouchet, 1989	Basy	Trochidae	Genbank	DQ093433	DQ279982	-	DQ093477	DQ093521
Gibbula magnus (Linné, 1758)	Gima	Trochidae	Genbank	-	-	AY145406	-	-
Lepetodrilus elevatus McLean, 1988	Leel	Lepetodrilidae	Genbank	-	-	AY145413	-	-
Nordotis discus (Reeve, 1846)	Nodi	Haliotidae	Genbank	-	-	AY145418	-	-

Tab. A2: List of used chemicals and kits (in alphabetical order)

Chemical/Kit	Company
Agarose	Carl-Roth, Karlsruhe, Germany
100 bp-DNA-Leiter extended	Carl-Roth, Karlsruhe, Germany
BSA (Albumin)	Invitrogen, Karlsruhe, Germany
Buffer	Invitrogen, Karlsruhe, Germany
dH2O	destilated Water
DMSO (Dimethyl Sulfoxide)	Carl-Roth, Karlsruhe, Germany
DNeasy Tissue Kit	Qiagen, Hilden, Germany
dNTP	Invitrogen, Karlsruhe, Germany
Ethidium bromide	Carl-Roth, Karlsruhe, Germany
Lamda Hind III-Ladder	Carl-Roth, Karlsruhe, Germany
MgCl2	Invitrogen, Karlsruhe, Germany
Primer	Invitrogen, Karlsruhe, Germany
Primer	Invitrogen, Karlsruhe, Germany
QIAquick Gel Extraction Kit	Qiagen, Hilden, Germany
Taq polymerase recombinant	Invitrogen, Karlsruhe, Germany
TBE Buffer	
-Tris	NeoLab Migge Laborbedarf-Vertriebs GmbH, Heidelberg, Germany
-Boracid	NeoLab Migge Laborbedarf-Vertriebs GmbH, Heidelberg, Germany
-EDTA	Mallinckrodt Baker, Griesheim, Germany
TMAC (Tetramethyl Ammonium Chloride)	Carl-Roth, Karlsruhe, Germany

Tab. A3: Primer design following the IUPAC 1-letter code abbreviations (S = G/C, W= T/C, Y = C/T, R = AG, M = A/C)

Reference	Sequence 5' → 3'	Direction	Fragment length	Primer
			ca. 1800 bp	18S COMPLETE
				PCR amplification
Wollscheid & Wägele 1999	CCT ACT TCT GGT TGA TCC TGC CAG T	Forward		18A1
Wollscheid & Wägele 1999	TAA TGA TCC TTC CGC AGG TT	Reverse		1800
				Sequencing
Vonnemann et al. 2005	CTG GTT GAT CCT GCC AGT CAT ATG C	Forward		18A1seq
Wollscheid & Wägele 1999	ACG GGT AAC GGG GAA TCA GGG	Forward		400F
Vonnemann et al. 2005	CAG CAG GCA CGC AAA TTA CCC	Forward		470F
Vonnemann et al. 2005	GTC TGG TGC CAG CAG CCG CG	Forward		700F
Wollscheid & Wägele 1999	CTG AAA CTT AAA GGA ATT GAC GG	Forward		1155F
Vonnemann et al. 2005	CGT CCC TGC CCT TTG TAC ACA CC	Forward		1600F
Vonnemann et al. 2005	GAT CCT TCC GCA GGT TCA CCT ACG	Reverse		1800seq
Vonnemann et al. 2005	CAT CTA GGG CAT CAC AGA CC	Reverse		1500R
Wollscheid & Wägele 1999	CCG TCA ATT CCT TTA AGT TTC AG	Reverse		1155R
Vonnemann et al. 2005	CGC GGC TGC TGG CAC CAG AC	Reverse		700R
Wollscheid & Wägele 1999	CCC TGA TTC CCC GTT ACC CGT	Reverse		400R
			ca. 1000 bp	28S PARTIAL (D1-D3)
				PCR amplification
Dayrat et al. 2001	ACC CGC TGA ATT TAA GCA T	Forward		28SC1
Colgan et al. 2003	ACC CSC TGA AYT TAA GCA T	Forward		28SD1
Vonnemann et al. 2005	GAC GAT CGA TTT GCA CGT CA	Reverse		28SD3
				Sequencing
According to PCR primer	ACC CGC TGA ATT TAA GCA T	Forward		28SC1
Dayrat et al. 2001	GAA AAG AAC TTT GAA GAG AGA GT	Forward		28SC2F (C2')*
Vonnemann et al. 2005	CCC GTC TTG AAA CAC GGA CCA AGG	Forward		28SD2F
According to PCR primer	GAC GAT CGA TTT GCA CGT CA	Reverse		28SD3
Vonnemann et al. 2005	CCT TGG TCC GTG TTT CAA GAC GGG	Reverse		28SD2R
Dayrat et al. 2001	ACT CTC TCT TCA AAG TTC TTT TC	Reverse		28SC2R (C2)*
			ca. 3500 bp	28S COMPLETE
				PCR amplification
Passamaneck et al. 2004	ACC CGC TGA AYT TAA GCA TAT	Forward		F63.2
Passamaneck et al. 2004	TWC YRM CTT AGA GGC GTT CAG	Reverse		R3264.2
				Sequencing

Passamaneck et al. 2004	CAA GTA CCG TGA GGG AAA GTT G'	Forward		28F5
Passamaneck et al. 2004	TCC TTG GTC CGT GTT TC AAG ACG	Reverse		28 MT4.1
Passamaneck et al. 2004	GGA ACC AGC TAC TAG ATG GTT CG	Reverse		28nn
Passamaneck et al. 2004	GYW GGG ACC CGA AAG ATG GTG AAC	Forward		28F1-2
Passamaneck et al. 2004	GCA GAA CTG GCG CTG AGG GAT GAA C	Forward		28F2-2
Hillis & Dixon 1991	GGT GAG TTG TTA CAC ACT CCT TAG CGG	Reverse		28ff
Hillis & Dixon 1991	ATC CGC TAA GGA GTG TGT AAC AAC TCA CC	Forward		28ee
Passamaneck et al. 2004	GAG GCT GTK CAC CTT GGA GAC CTG CTG CG	Reverse		28R2
Passamaneck et al. 2004	CGC AGC AGG TCT CCA AGG TGM ACA GCC TC	Forward		28F4
Passamaneck et al. 2004	GAG CCA ATC CTT ATC CCA AAG TTA CGG ATC	Reverse		28R4
Passamaneck et al. 2004	GAT GAC GAG GCA TTT GGC TAC C	Reverse		28R3
Hillis & Dixon 1991	GAC GAG GCA TTT GGC TAC CTT AAG	Reverse		28gg
Hillis & Dixon 1991	AAG GTA GCC AAA TGY CTC GTC ATC	Forward		28V
Hillis & Dixon 1991	GTG AAT TCT GCT TCA CAA TGA TAG GAA GAG CC	Reverse		28X
			ca. 500 bp	16S PARTIAL
				PCR amplification and sequencing
Simon et al. 1994	CGC CTG TTT ATC AAA AAC AT	Forward		16S-H
Simon et al. 1994	CCG GTC TGA ACT CAG ATC ACG T	Reverse		16S-R
			ca. 700 bp	COI PARTIAL
				PCR amplification and sequencing
Folmer et al. 1994	GGT CAA CAA ATC ATA AAG ATA TTG G	Forward		LCOI
Folmer et al. 1994	TAA ACT TCA GGG TGA CCA AAA AAT CA	Reverse		HCOI
Colgan et al. 2003	CWA ATC AYA AAG ATA TTG GAA C	Forward		Cox AF
Colgan et al. 2003	AAT ATA WAC TTC WGG GTG ACC	Reverse		Cox AR
Colgan et al. 2003	GGT AAR TYT ATT GTA ATA GCW CC	Reverse		Cox 623R

Tab. A4: Alignment information

Alignment	Length of alignment (prior to removal of ambiguous positions)	Length of alignment (after removal of ambiguous positions)	Excluded nucleotide positions
Ambiguous alignment positions were excluded by visual judgement			
18S rDNA complete	2716 bp	1775 bp	175-240, 271-381, 434-454, 854-1200, 1212-1299, 1401-1412, 1455-1466, 1701-1728, 1778-1799, 2055-2243, 2564-2594, 2657-2676
28S rDNA partial (D1-D3)	1980 bp	830 bp	413-481, 506-668, 695-736, 765-843, 879-1018, 1049-1093, 1128-1290, 1301-1336, 1345-1387, 1472-1496, 1520-1841
28S rDNA complete	4084 bp	4084 bp	-
16S rDNA partial	722 bp	278 bp	18-32, 143-158, 238-409, 431-564, 594-634, 662-700
COI	579 bp	386 bp	3rd codon positions
Ambiguous alignment positions were determined with the software Aliscore			
18S rDNA complete	2716 bp	2636 bp	47-52, 165-168, 231-234, 283-292, 312-315, 436-438, 1241-1244, 1305-1313, 1406-1408, 2192-2209, 2362-2364, 2618- 2629
28S rDNA partial (D1-D3)	1980 bp	1809 bp	58-63, 113- 119, 526-535, 542-549, 577-578, 586-590, 612-621, 625-630, 672-675, 854-861, 873-877, 894-915, 936-941, 950-954, 962-967, 1006-1014, 1033, 1209-1213, 1223-1229, 1247-1248, 1250-1251, 1253-1264, 1269, 1780-1787, 1795-1808
28S rDNA complete	4084 bp	4084 bp	not analysed
16S rDNA partial	722 bp	569 bp	1-7, 20-27, 122, 141-165, 182-184, 212-220, 371-397, 452-478, 533-543, 562-573, 617-618, 639-642, 676-687, 692-706
COI position 1	193 bp	176 bp	17-21, 84-90, 107-108, 167-169
COI position 2	193 bp	193 bp	-
COI position 3	193 bp	18 bp	1-35, 42-70, 74-134, 143-193

Tab. A5: Models of sequence evolution

Alignment	Number of taxa	Model	Proportion of invariable sites	Gamma distribution shape parameter	Base frequencies	Substitution rate matrix
Dataset 0 (all positions)						
18S rDNA complete	52	GTR+I+G	0.1386	0.3587	freqA = 0.2052 freqC = 0.2631 freqG = 0.3059 freqT = 0.2257	R(a) [A-C] = 0.8825 R(b) [A-G] = 1.8087 R(c) [A-T] = 0.8778 R(d) [C-G] = 0.6711 R(e) [C-T] = 3.5935 R(f) [G-T] = 1.0000
28S rDNA partial (D1-D3)	52	GTR+I+G	0.1290	0.6352	freqA = 0.1482 freqC = 0.3121 freqG = 0.3415 freqT = 0.1982	R(a) [A-C] = 0.6206 R(b) [A-G] = 1.9812 R(c) [A-T] = 1.3390 R(d) [C-G] = 0.5563 R(e) [C-T] = 4.2856 R(f) [G-T] = 1.0000
16S rDNA partial	47	GTR+I+G	0.1451	0.5874	freqA = 0.3771 freqC = 0.1004 freqG = 0.1432 freqT = 0.3793	R(a) [A-C] = 1.5400 R(b) [A-G] = 4.2779 R(c) [A-T] = 1.5752 R(d) [C-G] = 0.6856 R(e) [C-T] = 5.4464 R(f) [G-T] = 1.0000
COI position 1	51	GTR+I+G	0.2135	0.8081	freqA = 0.2830 freqC = 0.1419 freqG = 0.2386 freqT = 0.3365	R(a) [A-C] = 0.9756 R(b) [A-G] = 2.3303 R(c) [A-T] = 1.6616 R(d) [C-G] = 0.4750 R(e) [C-T] = 18.8917 R(f) [G-T] = 1.0000
COI position 2	51	GTR+G	0	0.4753	freqA = 0.1055 freqC = 0.2601 freqG = 0.1708 freqT = 0.4636	R(a) [A-C] = 3.7541 R(b) [A-G] = 11.6163 R(c) [A-T] = 2.6554 R(d) [C-G] = 16.6939 R(e) [C-T] = 4.2370 R(f) [G-T] = 1.0000
COI position 3	51	HKY+G	0	0.4712	freqA = 0.3598 freqC = 0.0880 freqG = 0.1306 freqT = 0.4216	Ti/tv ratio = 28.9173
Dataset I (inserts and ambiguous alignment positions were excluded by visual judgement) Outgroup = *Bathymargarites symplector* (Vetigastropoda)						
18S rDNA complete	52	GTR+I+G	0.2536	0.4831	freqA = 0.2676 freqC = 0.2163 freqG = 0.2775 freqT = 0.2386	R(a) [A-C] = 1.1589 R(b) [A-G] = 2.0870 R(c) [A-T] = 0.7485 R(e) [C-T] =0.8912 R(e) [C-T] = 4.3018 R(f) [G-T] = 1.0000
28S rDNA partial (D1-D3)	52	GTR+I+G	0.1839	0.5076	freqA = 0.1666 freqC = 0.3011 freqG = 0.3482 freqT = 0.1841	R(a) [A-C] = 0.7743 R(b) [A-G] = 2.1984 R(c) [A-T] = 1.4786 R(d) [C-G] = 0.5559 R(e) [C-T] = 5.0289 R(f) [G-T] = 1.0000

16S rDNA partial	47	GTR+I+G	0.3017	0.6511	freqA = 0.3445 freqC = 0.1208 freqG = 0.1685 freqT = 0.3662	R(a) [A-C] = 1.0673 R(b) [A-G] = 5.3205 R(c) [A-T] = 2.0630 R(d) [C-G] = 0.4678 R(e) [C-T] = 5.7894 R(f) [G-T] = 1.0000
COI Position 1	same as dataset 0					
COI Position 2	same as dataset 0					
COI Position 3	same as dataset 0					
Dataset II (ambiguous alignment positions were determined with the software Aliscore and excluded from further analyses) Outgroup = *Bathymargarites symplector* (Vetigastropoda)						
18S rDNA complete	52	GTR+I+G	0.2624	0.6120	freqA = 0.2648 freqC = 0.2141 freqG = 0.2818 freqT = 0.2393	R(a) [A-C] = 1.0843 R(b) [A-G] = 2.3710 R(c) [A-T] = 0.7846 R(d) [C-G] = 1.0280 R(e) [C-T] = 4.5929 R(f) [G-T] = 1.0000
28S rDNA partial (D1-D3)	52	GTR+I+G	0.1797	0.6649	freqA = 0.1886 freqC = 0.2929 freqG = 0.3399 freqT = 0.1786	R(a) [A-C] = 0.8507 R(b) [A-G] = 2.2883 R(c) [A-T] = 1.5812 R(d) [C-G] = 0.5264 R(e) [C-T] = 5.6147 R(f) [G-T] = 1.0000
16S rDNA partial	47	GTR+I+G	0.3466	0.7178	freqA = 0.3012 freqC = 0.1455 freqG = 0.2181 freqT = 0.3352	R(a) [A-C] = 2.0588 R(b) [A-G] = 6.2902 R(c) [A-T] = 2.3338 R(d) [C-G] = 0.2031 R(e) [C-T] = 6.4473 R(f) [G-T] = 10000
COI Position 1	51	GTR+I+G	0.2489	0.7955	freqA = 0.2867 freqC = 0.1326 freqG = 0.2213 freqT = 0.3594	R(a) [A-C] = 2.4516 R(b) [A-G] = 3.2453 R(c) [A-T] = 1.9847 R(d) [C-G] = 1.0668 R(e) [C-T] = 25.4845 R(f) [G-T] = 1.0000
COI Position 2	same as taxon set "all position"					
COI Position 3	51	HKY+G	0	2.2008	freqA = 0.2152 freqC = 0.1599 freqG = 0.1024 freqT = 0.5224	Ti/tv ratio = 2.3411
Dataset III (secondary structure 18S) Outgroup = *Bathymargarites symplector* (Vetigastropoda)						
18S rDNA complete	45	GTR+I+G	0.1519	0.3861	freqA = 0.2120 freqC = 0.2569 freqG = 0.2916 freqT = 0.2395	R(a) [A-C] = 0.9825 R(b) [A-G] = 1.6636 R(c) [A-T] = 0.8875 R(d) [C-G] = 0.8755 R(e) [C-T] = 2.9367 R(f) [G-T] = 1.0000
Dataset IV (secondary structure 28S) Outgroup= *Nordotis discus* (Vetigastropoda)						
28S rDNA complete	22	GTR+I+G	0.3616	0.4475	freqA = 0.2071 freqC = 0.2764 freqG = 0.3236 freqT = 0.1930	R(a) [A-C] = 0.7132 R(b) [A-G] = 1.5505 R(c) [A-T] = 1.1976 R(d) [C-G] = 0.8776 R(e) [C-T] = 3.9398 R(f) [G-T] = 10000

Tab. A6: References of taxa images

Taxa	Source
Vetigastropoda	www.cienciatk.csic.es
Caenogastropoda	www.sydneycichlid.com
"Lower Heterobranchia"	
Valvatoidea	Angela Dinapoli
Architeconicoidea	Angela Dinapoli
Omalogyroidea	Angela Dinapoli
Orbitestellidae	Angela Dinapoli
Cimidae	Angela Dinapoli
Rissoelloidea	Angela Dinapoli
Murchisonellidae	Angela Dinapoli
Pyramidellidae	Angela Dinapoli
Glacidorbidae	Angela Dinapoli
Acteonidae	Annette Klussmann-Kolb
Opisthobranchia	
Nudipleura	www.seaslugforum.net
Sacoglossa	www.seaslugforum.net
Akeroidea	www.seaslugforum.net
Cephalaspidae	www.seaslugforum.net
Umbraculoidea	Angela Dinapoli
Pteropoda	www.seaslugforum.net
Pulmonata	
Otinoidea	Angela Dinapoli
Amphiboloidea	www.roboastra.com
Hygrophila	www.fugleognatur.dk
Siphonarioidea	www.conchology.be
Stylommatophora	www.gardensafari.net
Systellomatophora	Annette Klussmann-Kolb
Ellobioidea	www.gastropods.com
Uncertain systematic rank	
Aclididae	Angela Dinapoli

Tab. A7: Character matrix of 18S rRNA of domain 43 and domain E23, 2 & 5

Taxa	Domain 43	Domain E23, 2 & 5
Vetigastropoda		
Bathymargarites symplector	0	0
Diodora graeca	0	0
Caenogastropoda		
Littorina littorea	?	1
Pomacea bridgesii	1	1
Aperostoma palmeri	1	1
Balcis eburnea	1	1
Aclididae		
Larochella alta	1	?
Graphis sp.	1	1
"Lower Heterobranchia"		
Valvatoidea		
Valvata piscinalis	?	?
Cornirostra pellucida	?	?
Architeconicoidea		
Architectonica perspectiva	?	?
Omalogyroidea		
Omalogyra fusca	1	?
Omalogyra sp.	1	?
Orbitestellidae		
Orbitestella vera	1	1
Orbitestella sp.	1	1
Cimidae		
Cima sp.	1	1
Rissoelloidea		
Rissoella elongatospira	1	2
Rissoella rissoaformis	?	?
Murchisonellidae		
Murchisonella sp.	2	1
Ebala sp.	?	?
Pyramidellidae		
Odostomia sp.	3	2
Turbonilla sp.	3	2
Eulimella ventricosa	3	2
Glacidorbidae		
Glacidorbis rusticus	3	3
Acteonidae		
Pupa solidula	?	1
Rictaxis punctocaelatus	?	?
Opisthobranchia		
Nudipleura		
Tomthompsonia antarctica	1	2
Bathyberthella antarctica	1	2
Chromodoris krohni	?	2
Goniodoris nodosa	1	2
Sacoglossa		
Elysia viridis	3	2
Oxynoe antillarum	5	2
Cyerce nigricans	3	2
Akeroidea		
Akera bullata	3	2

Cephalaspidea		
Haminoea hydatis	3	2
Toledonia globosa	3	2
Umbraculoidea		
Umbraculum umbraculum	3	2
Tylodina perversa	?	2
Thecosomata		
Hyalocylis striata	3	2
Gymnosomata		
Spongiobranchaea australis	3	2
Pulmonata		
Otinoidea		
Smeagol phillipensis	3	2
Otina ovata	3	2
Amphiboloidea		
Phallomedusa solida	6	2
Hygrophila		
Latia neritoides	3	2
Acroloxus lacustris	4	2
Siphonarioidea		
Siphonaria alternata	3	2
Siphonaria capensis	?	?
Stylommatophora		
Deroceras reticulatum	3	2
Discus rotundatus	3	2
Systellomatophora		
Onchidella floridana	3	2
Onchidium verruculatum	3	2
Ellobioidea		
Ophicardelus ornatus	3	2

Tab. A8: Character matrix of 28S rRNA of domain E11 and domain G5_1

Taxa	Domain E11	Domain G5_1
Vetigastropoda		
Nordotis discus	0	0
Lepetodrilus elevatus	0	0
Gibbula magnus	0	0
Caenogastropoda		
Ilyanassa obsoleta	1	0
Aclididae		
Graphis sp.	1	1
"Lower Heterobranchia"		
Valvatoidea		
Valvata piscinalis	1	1
Cornirostra pellucida	1	1
Cimidae		
Cima sp.	1	1
Rissoelloidea		
Rissoella rissoaformis	1	1
Pyramidellidae		
Boonea seminuda	1	0
Glacidorbidae		
Glacidorbis rusticus	2	2
Acteonidae		
Rictaxis punctocaelatus	1	0
Opisthobranchia		
Nudipleura		
Diaulula sandiegensis	1	0
Sacoglossa		
Oxynoe antillarum	2	1
Cephalaspidea		
Haminoea solitaria	3	0
Umbraculoidea		
Umbraculum umbraculum	2	0
Aplysioidea		
Aplysia californica	2	0
Pulmonata		
Otinoidea		
Smeagol phillipensis	2	0
Hygrophila		
Latia neritoides	2	0
Siphonarioidea		
Siphonaria capensis	2	0
Stylommatophora		
Deroceras reticulatum	2	0
Arion silvaticus	2	0

INTERNATIONAL STRATIGRAPHIC CHART

International Commission on Stratigraphy

ICS

IUGS

Eonothem / Eon	Erathem / Era	System / Period	Series / Epoch	Stage / Age	Age Ma
Phanerozoic	Cenozoic	Quaternary *	Holocene		0.0117
			Pleistocene	Upper	0.126
				"Ionian"	0.781
				Calabrian	1.806
				Gelasian	2.588
		Neogene	Pliocene	Piacenzian	3.600
				Zanclean	5.332
			Miocene	Messinian	7.246
				Tortonian	11.608
				Serravallian	13.82
				Langhian	15.97
				Burdigalian	20.43
				Aquitanian	23.03
		Paleogene	Oligocene	Chattian	28.4 ±0.1
				Rupelian	33.9 ±0.1
			Eocene	Priabonian	37.2 ±0.1
				Bartonian	40.4 ±0.2
				Lutetian	48.6 ±0.2
				Ypresian	55.8 ±0.2
			Paleocene	Thanetian	58.7 ±0.2
				Selandian	~ 61.1
				Danian	65.5 ±0.3
	Mesozoic	Cretaceous	Upper	Maastrichtian	70.6 ±0.6
				Campanian	83.5 ±0.7
				Santonian	85.8 ±0.7
				Coniacian	~ 88.6
				Turonian	93.6 ±0.8
				Cenomanian	99.6 ±0.9
			Lower	Albian	112.0 ±1.0
				Aptian	125.0 ±1.0
				Barremian	130.0 ±1.5
				Hauterivian	~ 133.9
				Valanginian	140.2 ±3.0
				Berriasian	145.5 ±4.0
		Jurassic	Upper	Tithonian	150.8 ±4.0
				Kimmeridgian	~ 155.6
				Oxfordian	161.2 ±4.0
			Middle	Callovian	164.7 ±4.0
				Bathonian	167.7 ±3.5
				Bajocian	171.6 ±3.0
				Aalenian	175.6 ±2.0
			Lower	Toarcian	183.0 ±1.5
				Pliensbachian	189.6 ±1.5
				Sinemurian	196.5 ±1.0
				Hettangian	199.6 ±0.6
		Triassic	Upper	Rhaetian	203.6 ±1.5
				Norian	216.5 ±2.0
				Carnian	~ 228.7
			Middle	Ladinian	237.0 ±2.0
				Anisian	~ 245.9
			Lower	Olenekian	~ 249.5
				Induan	251.0 ±0.4
	Paleozoic	Permian	Lopingian	Changhsingian	253.8 ±0.7
				Wuchiapingian	260.4 ±0.7
			Guadalupian	Capitanian	265.8 ±0.7
				Wordian	268.0 ±0.7
				Roadian	270.6 ±0.7
			Cisuralian	Kungurian	275.6 ±0.7
				Artinskian	284.4 ±0.7
				Sakmarian	294.6 ±0.8
				Asselian	299.0 ±0.8
		Carboniferous – Pennsylvanian	Upper	Gzhelian	303.4 ±0.9
				Kasimovian	307.2 ±1.0
			Middle	Moscovian	311.7 ±1.1
			Lower	Bashkirian	318.1 ±1.3
		Carboniferous – Mississippian	Upper	Serpukhovian	328.3 ±1.6
			Middle	Visean	345.3 ±2.1
			Lower	Tournaisian	359.2 ±2.5
		Devonian	Upper	Famennian	374.5 ±2.6
				Frasnian	385.3 ±2.6
			Middle	Givetian	391.8 ±2.7
				Eifelian	397.5 ±2.7
			Lower	Emsian	407.0 ±2.8
				Pragian	411.2 ±2.8
				Lochkovian	416.0 ±2.8
		Silurian	Pridoli		418.7 ±2.7
			Ludlow	Ludfordian	421.3 ±2.6
				Gorstian	422.9 ±2.5
			Wenlock	Homerian	426.2 ±2.4
				Sheinwoodian	428.2 ±2.3
			Llandovery	Telychian	436.0 ±1.9
				Aeronian	439.0 ±1.8
				Rhuddanian	443.7 ±1.5
		Ordovician	Upper	Hirnantian	445.6 ±1.5
				Katian	455.8 ±1.6
				Sandbian	460.9 ±1.6
			Middle	Darriwilian	468.1 ±1.6
				Dapingian	471.8 ±1.6
			Lower	Floian	478.6 ±1.7
				Tremadocian	488.3 ±1.7
		Cambrian	Furongian	*Stage 10*	~ 492 *
				Stage 9	~ 496 *
				Paibian	~ 499
			Series 3	Guzhangian	~ 503
				Drumian	~ 506.5
				Stage 5	~ 510 *
			Series 2	*Stage 4*	~ 515 *
				Stage 3	~ 521 *
			Terreneuvian	*Stage 2*	~ 528 *
				Fortunian	542.0 ±1.0

Precambrian – Eonothem / Eon	Erathem / Era	System / Period	Age Ma
Proterozoic	Neoproterozoic	Ediacaran	~635
		Cryogenian	850
		Tonian	1000
	Mesoproterozoic	Stenian	1200
		Ectasian	1400
		Calymmian	1600
	Paleoproterozoic	Statherian	1800
		Orosirian	2050
		Rhyacian	2300
		Siderian	2500
Archean	Neoarchean		2800
	Mesoarchean		3200
	Paleoarchean		3600
	Eoarchean		4000
Hadean (informal)			~4600

(Top of Precambrian: 542)

* The status of the Quaternary is not yet decided. Its base may be assigned as the base of the Gelasian and extend the base of the Pleistocene to 2.6 Ma. The "Tertiary" comprises the Paleogene and Neogene and has no official rank.

This chart was drafted by Gabi Ogg. Intra Cambrian unit ages with * are informal, and awaiting ratified definitions.

Copyright © 2008 International Commission on Stratigraphy

Subdivisions of the global geologic record are formally defined by their lower boundary. Each unit of the Phanerozoic (~542 Ma to Present) and the base of Ediacaran are defined by a basal Global Standard Section and Point (GSSP), whereas Precambrian units are formally subdivided by absolute age (Global Standard Stratigraphic Age, GSSA). Details of each GSSP are posted on the ICS website (*www.stratigraphy.org*).

Numerical ages of the unit boundaries in the Phanerozoic are subject to revision. Some stages within the Cambrian will be formally named upon international agreement on their GSSP limits. Most sub-Series boundaries (e.g., Middle and Upper Aptian) are not formally defined.

Colors are according to the Commission for the Geological Map of the World (*www.cgmw.org*).

The listed numerical ages are from 'A Geologic Time Scale 2004', by F.M. Gradstein, J.G. Ogg, A.G. Smith, et al. (2004; Cambridge University Press) and "The Concise Geologic Time Scale" by J.G. Ogg, G. Ogg and F.M. Gradstein (in press)

Fig. A1: International stratigraphic chart; International Commission on Stratigraphy (2008).

Angela Dinapoli
R7, 1-2, 68161 Mannheim, Germany
Phone (private): +49 (0)621 3266323
E-Mail: a.dinapoli@bio.uni-frankfurt.de

Curriculum vitae

Personal details

Date of birth	24 July 1975
Place of birth	Mannheim
Nationality	german
Marital status	married

School education

1981 – 1985	Rheinauer Grundschule (elementary School), Mannheim, Germany
1985 – 1991	Konrad-Duden-Realschule (secondary School), Mannheim, Germany
1991 – 1993	Apprenticeship as Industrial Business Management Assistant, Friatec AG, Mannheim, Germany
1993 – 1996	Friedrich-List-Gymnasium (primary School), Mannheim, Germany Degree: Abitur (equivalent of A-level)

Scientific education and degrees

1996 – 1997	Studies of hispanistics at the University of Mannheim, Germany
1997 – 2004	Studies of biology at the Ruprecht-Karls-University, Heidelberg, Germany; Subjects: zoology, ecology and palaeontology
2004	Diploma thesis (MSc) conducted at the Zoological Institute, Ruprecht-Karls-University, Heidelberg, Germany and the Research Institute & Natural History Museum Senckenberg, Section for Malacology, Frankfurt am Main, Germany Topic: *Polyplacophora (Mollusca) from Socotra Island – Systematics and Biogeography* Supervisors: Prof. Dr. Dr. hc. V. Storch, Prof. Dr. Th. Braunbeck and Dr. Ronald Janssen Degree: *Diploma* (MSc) in biology by the Ruprecht-Karls-University, Heidelberg, Germany
Since 2005	Doctoral thesis conducted at the Institute for Ecology, Diversity and Evolution, Goethe-University of Frankfurt am Main, Germany Topic: *Phylogeny and Evolution of the Heterobranchia (Mollusca, Gastropoda)* Supervisor: Prof. Dr. Annette Klussmann-Kolb Funded by the German Science Foundation (KL 1303/4-1, 4-2)

Publications

KLUSSMANN-KOLB, A., **DINAPOLI, A.**, KUHN, K., STREIT, B. & ALBRECHT, C. (2008)
From sea to land and beyond - new insights into the evolution of euthyneuran Gastropoda (Mollusca)
BMC Evolutionary Biology 8 (57)

DINAPOLI, A., FRANSSEN, S. NADUVILEZHATH, L., TAMER, A. & KLUSSMMANN-KOLB, A. (2006)
Utility of H3-Genesequences for phylogenetic reconstruction – a case study of heterobranch Gastropoda
Bonner zoologische Beiträge 55 (3/4): 191-202.

KLUSSMANN-KOLB, A. & **DINAPOLI, A.** (2006)
Systematic position of the pelagic Thecosomata and Gymnosomata within Opisthobranchia (Mollusca, Gastropoda) – revival of the Pteropoda.
Journal of Zoological Systematics and Evolutionary Research 44: 118-129.

DINAPOLI, A. & JANSSEN, R. (in press)
Polyplacophora (Mollusca) of Socotra Island – Systematics and Biogeography
Fauna of Arabia

Talks

Dinapoli, A. & KLUSSMMANN-KOLB, A. (2008)
The Heterobranchia (Mollusca, Gastropoda) - Molecular phylogeny and evolution.
101th *annual meeting of the Deutschen Zoologischen Gesellschaft (DZG); Friedrich Schiller University, Jena, Germany* (in english)

Dinapoli, A., WEIGAND, A. & KLUSSMMANN-KOLB, A. (2008)
Comparative analysis of 18S rRNA secondary structures of Heterobranchia (Mollusca, Gastropoda): phylogenetic signal and structural variability.
5th congress of the European Malacological Societies (CEMS); Ponta Delgada, São Miguel Island, Azores, Portugal (in english)

Dinapoli, A. & KLUSSMMANN-KOLB, A. (2007)
Toward a molecular phylogeny of the Heterobranchia (Mollusca, Gastropoda)
9th annual meeting of the Gesellschaft für Biologische Systematik (GfBS); Naturhistorisches Museum, Wien, Austria (in english)

Poster

WEIGAND, A., **Dinapoli, A.** & KLUSSMMANN-KOLB, A. (2009)
New insights into an old gene:
Prospects and Problems of 18S rDNA for Phylogenetic Analyses:
Insights from Secondary Structures of Heterobranchia (Gastropoda)
Celebrating Darwin: From The Origin of Species to Deep Metazoan Phylogeny, Berlin, Germany

Dinapoli, A. & KLUSSMMANN-KOLB, A. (2008)
Minute and Minute and misunderstood:
The systematic position of *Cima* CHASTER*, 1896, Graphis* JEFFREYS*,* 1867 and *Larochella* POWELL*, 1927* (Mollusca, Gastropoda)
10th Jährliches Treffen der Gesellschaft für Biologische Systematik, Göttingen, Germany

DINAPOLI, A., FRANSSEN, S. NADUVILEZHATH, L., TAMER, A. & KLUSMMANN-KOLB, A. (2006)
Utility of H3-Genesequences for phylogenetic reconstruction – a case study of heterobranch Gastropoda
2nd International Workshop on Opisthobranchia; Zoologisches Forschungsmuseum Alexander Koenig, Bonn, Germany

Academic Teachers

Prof. Dr. C. R. Bartram
Prof. Dr. R Beiderbeck
Prof. Dr. P. Bengtson
Prof. Dr. P. Bock
Prof. Dr. Th. Braunbeck
Prof. Dr. P. Comba
Prof. Dr. C. Erber
Prof. Dr. M. Hassel
Dr. H. Hollert
Prof. Dr. R. Herrmann
Prof. Dr. H. Irngartinger
Prof. Dr. M. Lanzer
Prof. Dr. H. F. Moeller
Prof. Dr. H. Schneider
Prof. Dr. V. Schweizer
Prof. Dr. Dr. hc.. V. Storch
HD Dr. G. Vogt
Dr. U. Weinhold
Dr. K. Wünnemann
Prof. Dr. R. Zwilling

Frankfurt am Main, 26.04.2009

E R K L Ä R U N G

Ich erkläre hiermit, dass ich mich bisher keiner Doktorprüfung unterzogen habe.

Frankfurt am Main, den .....

.....

(Unterschrift)

Eidesstattliche Versicherung

Ich erkläre hiermit an Eides Statt, dass ich die vorgelegte Dissertation über

Phylogeny and Evolution of the Heterobranchia (Mollusca, Gastropoda)

selbständig angefertigt und mich anderer Hilfsmittel als der in ihr angegebenen nicht bedient habe, insbesondere, dass aus Schriften Entlehnungen, soweit sie in der Dissertation nicht ausdrücklich als solche mit Angabe der betreffenden Schrift bezeichnet sind, nicht stattgefunden haben.

Frankfurt am Main, den ....

(Unterschrift)