Molecular phylogenetics of Caenogastropoda (Gastropoda: Mollusca)

D.J. Colgan a,*, W.F. Ponder b, E. Beacham a, J. Macaranas a

a Evolutionary Biology Unit, The Australian Museum, 6 College Street, Sydney 2010, NSW, Australia
b Malacology Section, The Australian Museum, 6 College Street, Sydney 2010, NSW, Australia

Received 21 December 2005; revised 5 September 2006; accepted 9 October 2006
Available online 17 October 2006

Abstract

Caenogastropoda is the dominant group of marine gastropods in terms of species numbers, diversity of habit and habitat and ecological importance. This paper reports the first comprehensive multi-gene phylogenetic study of the group. Data were collected from up to six genes comprising parts of 18S rRNA, 28S rRNA (five segments), 12S rRNA, cytochrome c oxidase subunit I, histone H3 and elongation factor 1a. The alignment has a combined length of 3995 base positions for 36 taxa, comprising 29 Caenogastropoda representing all of its major lineages and seven outgroups. Maximum parsimony, maximum likelihood and Bayesian analyses were conducted. The results generally support monophyly of Caenogastropoda and Hypsogastropoda (Caenogastropoda excepting Architaenioglossa, Cerithioidea and Campanilioidea). Within Hypsogastropoda, maximum likelihood and Bayesian analyses identified a near basal clade of nine or 10 families lacking an anterior inhalant siphon, and Cerithiopsidae s.l. (representing Triphoroidea), where the siphon is probably derived independently from other Hypsogastropoda. The asiphonate family Eatoniellidae was usually included in the clade but was removed in one Bayesian analysis. Of the two other studied families lacking a siphon, the limpet-shaped Calyptraeidae was associated with this group in some analyses, but the tent-shaped Xenophoridae was generally associated with the siphonate Strombidae. The other studied hypsogastropods with an anterior inhalant siphon include nine families, six of which are Neogastropoda, the only traditional caenogastropod group above the superfamily-level with strong morphological support. The hypotheses that Neogastropoda are monophyletic and that the group occupies a derived position within Hypsogastropoda are both contradicted, but weakly, by the molecular analyses. Despite the addition of large amounts of new molecular data, many caenogastropod lineages remain poorly resolved or unresolved in the present analyses, possibly due to a rapid radiation of the Hypsogastropoda following the Permian–Triassic extinction during the early Mesozoic.

Keywords: Architaenioglossa; Bayesian analysis; Cladistics; Heteropoda; Hypsogastropoda; Maximum likelihood; Multi-gene phylogeny; Neogastropoda; Ptenoglossa

1. Introduction

Among the living snails, the largest and most diverse group is Caenogastropoda. The group includes a large number of ecologically and commercially important marine families, as well as several groups that have independently achieved major freshwater and terrestrial radiations. Its members display a wide array of often-convergent shell morphologies (coiled, uncoiled, elongate, globose, limpet-shaped, etc.), and some species have the shell reduced or (rarely) lost. They occupy a wide range of habitats and have diverse habits (benthic epifaunal or burrowers, pelagic drifters or active swimmers; detritus or sedentary suspension feeders, herbivores or grazing or active carnivores, ectoparasites or shell-less internal parasites).

Caenogastropoda consists of about 136 extant and 65 extinct families and thousands of genera currently arranged in 41 superfamilies (Bouchet and Rocroi, 2005). The relationships between the families and superfamilies remain largely unresolved phylogenetically with most named higher taxa probably paraphyletic or even polyphyletic. Available hypotheses about the relationships are largely based on a few key shell characters and anatomical details.
Heteropoda comprises only the pelagic Pterotracheoida (= Carinarioidea) and has been sometimes used as a high-rank taxon (e.g., Bandel and Hemleben, 1987; Ponder and Warén, 1988). Pterotracheoida is included in the Littorinimorpha by Bouchet and Rocroi (2005) and is given no higher rank, agreeing with Thiele (1929–31) and Wenz (1938–44).

Ptenoglossa (= Ctenoglossa), a probably polyphyletic (Ponder and Lindberg, 1997) grouping of Eulimoidea, Janthinioidea and Triphoroidea.

Several new higher taxon names have been introduced by Bandel and his colleagues (e.g., Bandel, 1991, 1993) largely to accommodate fossils. These include: Palao-Caenogastropoda—taxa with their first occurrence in the Palaeozoic; Meta-Mesogastropoda—taxa with Mesozoic origins; Neo-Mesogastropoda—late Mesozoic taxa that have an “expanded ontogeny” (does not include Neogastropoda); and Scaphoconchoidea—taxa with modified veliger stages known as echinospira or limacosphaera larvae. Bandel and Riedel (1994) and Riedel (2000) introduced Latrogastropoda to include Neomesogastropoda and Neogastropoda. Within this grouping, Pleurembolina includes Troschelina (Lauberinoidea and Calyptraeacea) and Vermivora (Ficoidae, Tonnaidea and Neogastropoda).

Current understanding of the systematics of Caenogastropoda broadly follows the findings of the morphological study of overall gastropod phylogeny by Ponder and Lindberg (1997). This study resolved Architaenioglossa, Sorbeconcha and Hypsogastropoda but revealed little structure within the latter group. Ponder and Lindberg (1997) included 11 species of Hypsogastropoda but only one group, the Neogastropoda, was resolved. Bouchet and Rocroi (2005) differ from Ponder and Lindberg (1997) mainly in the recognition of Littorinimorpha (see above).

The membership of Caenogastropoda has not recently been widely questioned except for Architaenioglossa and Campanilidae (Haszprunar, 1988). Architaenioglossa is an enigmatic, entirely non-marine group. While generally regarded as a member of “Mesogastropoda” (Thiele, 1929–31; Wenz, 1938–44) or a caenogastropod group (Ponder and Warén, 1988), Haszprunar (1988) considered it to be the sister group of a clade comprising Caenogastropoda and Heterobranchia. Ponder and Lindberg (1997) found Architaenioglossa (represented by Cyclophoridae and Ampullariidae in their analysis) to be included within Caenogastropoda. They also found it to be monophyletic, albeit with support provided only by four forward homoplasies. Robust support was found for the sister pairing of the two Architaenioglossa (Cyclophoridae and Ampullariidae) included in Strong’s (2003) analysis. Monophyly of the group has, however, been questioned by other morphological studies (Haszprunar, 1988; Simone, 2004) and is contradicted by analyses of DNA sequences including Viviparidae, Ampullariidae and Cyclophoridae (Harasewych et al., 1998) or only Ampullariidae and Cyclophoridae (Colgan et al., 2000, 2003).

(particularly from radulae). The first explicitly cladistic analysis that focussed on the whole group was undertaken by Strong (2003), as discussed below, and Simone (2001, 2004, 2005) has made phylogenetic analyses of several of its substantial components.

The taxon Caenogastropoda was introduced by Cox (1960) to include many elements of the “Prosbobranchia” recognised by Thiele (1929–31) but shown to be paraphyletic by Haszprunar (1988) and Ponder and Lindberg (1997). These were (1) Architaenioglossa, containing Cyclophoroidea, the major group of operculate land snails, and the freshwater families Ampullariidae and Viviparidae, sometimes considered to belong to one superfAMILY but now each recognized as a separate superfAMILY; (2) the remaining “mesogastropods” of Thiele (1929–31) including predominantly marine groups such as the Littorinidae (periwinkles), Cypraeidae (cowries), Cerithiidae (creapers), Calyptraeidae (slipper limpetS), Tonnidae (tun shells), Cassidae (helmet shells), Ranellidae (tritons), Strombidae (strombs), Naticidae (moon snails) and Heteropoda (= Pterotracheoida) and (3) Stenoglossa (= Neogastropoda) an almost exclusively marine and carnivorous group that contains such well-known, diverse and ecologically significant families as Muricidae (rock shells, oyster drills, etc.), Voluitidae (balers, etc.), Mitridae (mitres), Buccinidae (whelks), Turridae (turrids) and Conidae (cones). Thiele’s concept of Mesogastropoda is polyphyletic including some groups now known to be members of the Heterobranchia, the major gastropod clade that also includes the opisthobranchs and pulmonates and that is currently recognised as the sister group of Caenogastropoda (Haszprunar, 1985, 1988; Ponder and Lindberg, 1997).

The family-group taxa studied in this paper are shown in Fig. 1 with many of the higher taxon names used herein indicated. For a comprehensive listing of higher taxa used in Gastropoda see Bouchet and Rocroi (2005). Names other than Architaenioglossa and Neogastropoda in current or recent use for major groups within Caenogastropoda include:

Sorbeconcha introduced by Ponder and Lindberg (1997), includes all caenogastropods other than the Architaenioglossa. Basal members are Cerithioidea and Campaniloidea.

Hypsogastropoda also named by Ponder and Lindberg (1997) includes the great majority of extant caenogastropods, and is defined as all caenogastropods other than Architaenioglossa, Cerithioidea or Campaniloidea.

Neotaenioglossa has been used for the non-architaenioglossan “mesogastropods” (Haszprunar, 1988; Ponder and Warén, 1988). This group is equivalent to Sorbeconcha excluding Neogastropoda. It is now generally acknowledged that Neotaenioglossa is either paraphyletic or polyphyletic (Ponder and Lindberg, 1997; Strong, 2003).

Cerithiimorpha and Littorinimorpha have been used as groupings within Neotaenioglossa. Most recently Littorinimorpha was used to encompass the taeonioglossate Hypsogastropoda (Bouchet and Rocroi, 2005).
Campaniloidea (then comprising only *Campanile*) was regarded by Haszprunar (1988) as the sister group to Heterobranchia. The single living species of Campanilidae, *C. symbolicum* (Houbrick, 1981, 1989) and sperm morphology (Healy, 1986). Subsequently, Healy (1993) transferred the family Plesiotrochidae to the Campaniloidea on the basis of similarity in sperm ultrastructure. The morphological analyses of Ponder and Lindberg (1997) placed Campaniloidea within Caenogastropoda, as anticipated by Ponder and Warén (1988). Molecular studies have also consistently included Campaniloidea in Caenogastropoda (Harasewych et al., 1998; Colgan et al., 2000, 2003; McArthur and Harasewych, 2003).

Only one large-scale cladistic analysis of caenogastropod morphology has been published (Strong, 2003). That study included the two architaenioglossans mentioned earlier, a cerithioidean and 14 hypsogastropods (including six neogastropods) and utilized 64 characters from a wide range of organ systems such as the alimentary, renal/pericardial, nervous, reproductive and the mantle cavity. Unlike Ponder and Lindberg (1997); Strong (2003) did not include ultrastructural characters. Caenogastropoda, Architaenioglossa, Sorbeoconcha and Neogastropoda were monophyletic in Strong’s analyses (2003), but Hypsogastropoda included the studied cerithioidean (family Batillariidae).

This investigation of DNA sequence data was conducted to identify the major lineages within Caenogastropoda and their relationships. We investigate molecular support for monophyly of groups such as Sorbeoconcha, Hypsogastropoda and Neogastropoda that have some morphological support, as well as named groups such as Architaenioglossa,
Neotaenioglossa and Ptenoglossa whose status is doubtful. The data collected for this paper also permitted limited testing, based on two taxa in each case, of the monophyly of some superfamilies that have not previously been tested with a large number of outgroups, although some analyses of those groups have been conducted. These include the Rissooidea (Ponder, 1988; Wilke et al., 2001), Cerithioidea (Houbrick, 1988; Ponder, 1991; Simone, 2001; Lydeard et al., 2002), Tonnioidea (Riedel, 1995), Neogastropoda (Kantor, 1996) and Triphoroidea + Janthinoidea (Nützel, 1998).

To date there has not been a comprehensive molecular study of caenogastropod evolution. Some papers discussed below have included a substantial number of caenogastropod taxa. Most have focussed on a particular subgroup such as a family or group of families (e.g., Harasewych et al., 1997; Oliverio and Mariottini, 2001; Oliverio et al., 2002; Collin, 2003; Meyer, 2003, 2004; Williams et al., 2003; Oliverio and Mariottini, 2001; Oliverio et al., 1997; Riedel, 2000) or on overall gastropod phylogeny (Colgan et al., 2000, 2003; McArthur and Harasewych et al., 1997; Riedel, 2000) or on overall gastropod phylogeny (Colgan et al., 2000, 2003; McArthur and Harasewych, 2003). There is a notable lack of resolution within Caenogastropoda in these overall gastropod phylogenies.

Data for this study were already available for a number of taxa used in our previous studies of gastropod phylogeny. Seventeen caenogastropod taxa were scored for two segments of 28S rRNA (abbreviated as 28S rRNA A and 28S rRNA B) and histone H3 (H3) in Colgan et al. (2000). In this context and below, the term "segment" indicates the product of a single PCR amplification, although most of the multiple segments from 28S rRNA were originally designed for the investigations of particular expansion regions. Three extra sequences were studied for 16 taxa in Colgan et al. (2003). These were an additional segment of 28S ribosomal RNA (abbreviated as 28S rRNA D1), small nuclear RNA U2 (U2 snRNA) and part of cytochrome c oxidase subunit 1 (COI). These extra segments were sequenced here for the taxon (Eulimidae) included in Colgan et al. (2000) but not in Colgan et al. (2003). Data were not collected for U2 snRNA as this sequence is quite short (less than 150 bases).

For this paper, data were collected from five additional gene segments, these being part of the 12S rRNA domain III (abbreviated as 12S rRNA), two segments of the 28S rRNA containing identified expansion regions (abbreviated as 28S rRNA D6 and 28S rRNA D9-10), part of the 18S ribosomal RNA (18S rRNA) and elongation factor 1 alpha (EF1-α). As well as collecting sequences from more segments, data were collected from 12 additional caenogastropods for as many of the 10 studied segments as could be amplified from each specimen.

The nuclear ribosomal genes are regularly used in molecular phylogenetic surveys. Among the coding regions, H3 has been widely used in higher-level phylogenetics (Colgan et al., 1998; Brown et al., 1999; Whiting et al., 2003; Thollesson and Norenburg, 2003; Okusu et al., 2003). EF1-α is also becoming widely used in the phylogenetics of higher level taxa, for example in Polychaeta (McHugh, 1997) and Arthropoda (Regier and Shultz, 1997; Giribet et al., 2001). COI sequences are too variable to resolve major groups in higher-level phylogenetic analyses by themselves (Nylander et al., 1999) but have proven useful in combination with other data for recovering some high rank taxa, e.g. Clitellata in Jördens et al. (2004).

Outgroups were chosen from a range of major gastropod groups. The Heterobranchia is usually found to be the sister group to Caenogastropoda in morphological and molecular analyses, the two clades together comprising the Apogastropoda (sensu Ponder and Lindberg, 1997). More distant outgroups were selected from the Neritimorpha and the Vetigastropoda.

2. Materials and methods

2.1. DNA extraction and sequencing

Specimens were frozen at −70°C or stored in 70–100% ethanol. Species examined, classification and voucher details are listed in Table 1. DNA was isolated from up to 1 g of foot tissue or from the entire animal, if it were small (body size <5 mm length). Ethanol-preserved samples were re-hydrated in sterile double filtered H2O for at least 2 h before extraction. DNA was extracted by the CTAB method of Saghai-Marof et al. (1984) or the AMRESCO RapidGene™ Genomic DNA Purification Kit (Solon, OH, USA) according to manufacturer's instructions. Extractions of rRNA were carried out using TRIZOL (Life Technologies, Rockville, MD, USA) and cDNA synthesis performed using the Superscript™ preamplification system (Life Technologies) according to the manufacturer's protocol Table 2.

DNA dilutions up to 1 in 100 were made to obtain single-banded PCR products. PCRs were generally performed using 0.2–1.0 Units of Red Hot™ thermostable DNA polymerase in a 10-fold dilution of Buffer IV (Advanced Biotechnologies, Columbia, MD, USA: 20 mM (NH4)2SO4, 750 mM Tris–HCl pH 9.0, 0.1% (w/v) Tween), with a final concentration of 0.05 mM dNTPs, 3.5–4.5 mM MgCl2 and 12.5–25 pmol of each primer in a total reaction volume of 50 μl. Negative controls were included in each reaction array. Other DNA polymerases were occasionally used in the particular manufacturer's buffer, with similar concentrations of other reagents. To optimise PCR products, annealing temperatures and times, and MgCl2 concentration were varied. The basic cycling profile was as follows: (95°C for 5 min, annealing for 45 s, 72°C for 1 min) for one cycle, (95°C for 30 s, annealing for 45 s, 72°C for 1 min) for 30–34 cycles and (95°C for 30 s, annealing for 45 s, 72°C for 5 min) for the
Table 1
Classification, provenance and museum registration of the studied specimens

<table>
<thead>
<tr>
<th>Higher taxon</th>
<th>Family</th>
<th>Species</th>
<th>Source</th>
<th>Reg No./Voucher No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vetigastropoda</td>
<td>Fissurelleidae</td>
<td>Montfortula rugosa</td>
<td>Edwards Beach, Balmoral, NSW</td>
<td>C335476</td>
</tr>
<tr>
<td></td>
<td>Pleurotomarioida</td>
<td>Bayerotrochus midas</td>
<td>Johnson Sea Link, Bahamas, 2100ft Depth</td>
<td>USNM888645</td>
</tr>
<tr>
<td>Neritomorpha (= Neritopsina)</td>
<td>Neritidea</td>
<td>Nerita atramentosa</td>
<td>Edwards Beach, Balmoral, NSW</td>
<td>C335471</td>
</tr>
<tr>
<td>Caenogastropoda</td>
<td>Cyclophorioida</td>
<td>Leptopoma perlucida</td>
<td>Coonya Beach, Qld</td>
<td>QM M03142</td>
</tr>
<tr>
<td></td>
<td>Ampullarioida</td>
<td>Pomaceae bridgesii</td>
<td>Brookfield, W of Brisbane, Qld</td>
<td>C333043</td>
</tr>
<tr>
<td></td>
<td>Viviaparoida</td>
<td>Bellaneya heidi guangdungensis</td>
<td>Halfway Pt, Lane Cove River, NSW</td>
<td>C203209</td>
</tr>
<tr>
<td>Sorbeconcha</td>
<td>Cerithioida</td>
<td>Zeacumants subcarinatus</td>
<td>Edwards Beach, Balmoral, NSW</td>
<td>EBU30432</td>
</tr>
<tr>
<td></td>
<td>Campaniloida</td>
<td>Campanile symbolicum</td>
<td>Rottonet Island, WA</td>
<td>C203211</td>
</tr>
<tr>
<td>Hypsogastropoda</td>
<td>Littorinoida</td>
<td>Austrotitorina unifasciata</td>
<td>Edwards Beach, Balmoral, NSW</td>
<td>C335438</td>
</tr>
<tr>
<td></td>
<td>Cingulopsoidea</td>
<td>Crassotiella flammaea</td>
<td>Edwards Beach, Balmoral, NSW</td>
<td>C335411</td>
</tr>
<tr>
<td></td>
<td>Rissooida</td>
<td>Pismina albizona</td>
<td>Edwards Beach, Balmoral, NSW</td>
<td>C335405</td>
</tr>
<tr>
<td></td>
<td>Rissoida</td>
<td>Rissoina fasciata</td>
<td>Long Reef, Collaroy, NSW</td>
<td>C335432</td>
</tr>
<tr>
<td></td>
<td>Stromboidea</td>
<td>Strombus labuanus</td>
<td>Heron Island, Qld</td>
<td>C203214</td>
</tr>
<tr>
<td></td>
<td>Vanikoroidea</td>
<td>Antisabia folacea</td>
<td>Long Reef, Collaroy, NSW</td>
<td>C335477</td>
</tr>
<tr>
<td></td>
<td>Calyptraeoida</td>
<td>Bostryacapalus pritzkeri</td>
<td>Edwards Beach, Balmoral, NSW</td>
<td>C335468</td>
</tr>
<tr>
<td></td>
<td>Xenophoroidea</td>
<td>Xenophora indica</td>
<td>“Kapala” cruise K951805/09 E of Clarence R., NSW</td>
<td>C454048</td>
</tr>
<tr>
<td></td>
<td>Vermetoidea</td>
<td>Serpulbris sp.</td>
<td>Heron Island, Qld</td>
<td>C203213</td>
</tr>
<tr>
<td></td>
<td>Cypraeoida</td>
<td>Cypraea annulus</td>
<td>Heron Island, Qld</td>
<td>C203215</td>
</tr>
<tr>
<td></td>
<td>Naticoidea</td>
<td>Conuber melanostoma</td>
<td>Careel Bay, Pittwater, NSW</td>
<td>EBU30442</td>
</tr>
<tr>
<td></td>
<td>Tonnidea</td>
<td>Cabestana spengleri</td>
<td>Long Reef, Collaroy, NSW</td>
<td>C203216</td>
</tr>
<tr>
<td></td>
<td>Pterotraceoida (&quot;Heteropoda&quot;)</td>
<td>Pterotracea corona</td>
<td>San Pedro Basin, trawled, USA</td>
<td>EBU30445</td>
</tr>
<tr>
<td></td>
<td>Triphorioida</td>
<td>Ataxocerithium sp.</td>
<td>Off Cronulla, NSW</td>
<td>C203217</td>
</tr>
<tr>
<td></td>
<td>Epitonioida</td>
<td>Epitonium jukesianum</td>
<td>Edwards Beach, Balmoral, NSW</td>
<td>C335429</td>
</tr>
<tr>
<td></td>
<td>Eulimoidea</td>
<td>Stilfer sp.</td>
<td>Heron Island, Qld</td>
<td>DNA only</td>
</tr>
<tr>
<td>Neogastropoda</td>
<td>Buccinoidea</td>
<td>Nassarius burchardi</td>
<td>The Entrance, NSW</td>
<td>C203219</td>
</tr>
<tr>
<td></td>
<td>Muricoidea</td>
<td>Dicathais orbita</td>
<td>Edwards Beach, Balmoral, NSW</td>
<td>C335420</td>
</tr>
<tr>
<td></td>
<td>Mitridae</td>
<td>Mitra carbonaria</td>
<td>Long Reef, Collaroy, NSW</td>
<td>C335424</td>
</tr>
<tr>
<td></td>
<td>Volutidae</td>
<td>CymbiolaHunteri</td>
<td>“Kapala” cruise K960401 off Newcastle</td>
<td>C453717</td>
</tr>
<tr>
<td></td>
<td>Cancellarioidea</td>
<td>Cancellaria undulata</td>
<td>Woolgoolga, NSW</td>
<td>C203222</td>
</tr>
<tr>
<td></td>
<td>Conoidea</td>
<td>Conus miles</td>
<td>Heron Island, Qld</td>
<td>C203223</td>
</tr>
<tr>
<td>Heterobranchia</td>
<td>Valvatoidea</td>
<td>Cornirostridum</td>
<td>Fingal Head, Port Stephens, NSW</td>
<td>C203224</td>
</tr>
<tr>
<td></td>
<td>Opisthobranchia</td>
<td>Aplysiida</td>
<td>Aplysia juliana</td>
<td>Long Reef, Collaroy, NSW</td>
</tr>
<tr>
<td></td>
<td>Acteonoida</td>
<td>Aplustrida</td>
<td>Long Reef, Collaroy, NSW</td>
<td>C203227</td>
</tr>
<tr>
<td></td>
<td>Aplysioida</td>
<td>Bullina lineata</td>
<td>Long Reef, Collaroy, NSW</td>
<td></td>
</tr>
<tr>
<td>Pulmonata</td>
<td>Amphiboloidea</td>
<td>Salinator solida</td>
<td>Tillerger Ck, Port Stephens, NSW</td>
<td>C203229</td>
</tr>
</tbody>
</table>

The higher classification follows Bouchet and Rocroi (2005). Registration numbers with the prefix C denote specimens in the Malacology collection of the Australian Museum. EBU numbers denote DNA or whole specimens in the frozen tissue collection of the Australian Museum. QM M numbers indicate specimens from the Queensland Museum and USNM numbers indicate specimens from the National Museum of Natural History.

* Bouchet and Rocroi (2005) use Newtoniellidae as the family name for Ataxocerithium but as the systematics of this group are poorly understood, we choose to use the better known family name Cerithiopsidae (in the broad sense).

final cycle. Annealing temperatures are indicated below. To obtain PCR products from difficult samples, 20 μL of GeneReleaser™ (Bioventures, Murfreesboro, TN, USA) was added to the DNA template and microwaved for 6 min. The remaining PCR mix (with reduced H₂O) was immediately added and cycling commenced.
Primer sequences are listed in Table 1. Several primer pairs (Cox AF and AR, Cox BF and BR, etc.) were used for cytochrome c oxidase subunit I. Annealing temperatures were usually 45 °C for all primer combinations but were reduced to as low as 40°C if this were necessary to obtain products or as high as 52 °C to eliminate secondary bands. Generally, 12SF and 12SR (the universal primers of Kocher et al. (1989) were used for 12S rRNA. For taxa not successfully amplified with this pair, the nested primers 12S2F and 12S2R were used. These were designed from sequences generated using the first primer pair. Annealing temperatures were usually 50 °C for all primer combinations but were reduced to 46 °C if necessary. For the 28S rRNA D1 expansion region, D1F and D1R were used with the addition of 5% DMSO to the PCR reaction. Annealing temperatures were usually 50 °C for all primer combinations but were reduced to 46 °C if necessary. For the 28S rRNA D6 expansion region, D6F and D6R, designed from sequences collected here, were used with D6R and D6F respectively to amplify the 28S rRNA D6 region in two sections. Annealing temperatures were between 47 and 50 °C for all primer combinations. Annealing temperatures for 28S rRNA A, 28S rRNA B and histone H3 ranged from 48–53 °C. The 28S 9-10 region overlaps 28SB at the 3′ end of the former segment. There were no differences in the sequences in the overlap region. EF-1α was amplified with primers EFF and EFB, designed from the gastropod Alvinoconcha hessleri (D14975) (Kojima et al., 1993) with reference to Drosophila melanogaster sequences. Initial amplifications of genomic DNA were generally unsuccessful. Subsequent amplifications were performed using RT-PCR as described above using an annealing temperature of 60 °C.

Reaction products were resolved on 2% agarose gels containing ethidium bromide. Single band products were purified using the QIAquick™ PCR Purification Kit (Qiagen, Venlo, The Netherlands) or by AMPURE magnetic beads (Agencourt, Beverly, MA, USA) processed by a
liquid handling system (CAS-3800, Corbett Engineering, Mortlake, Australia). Where single products were not obtained, the correct sized band was excised from 2% low melting point agarose in TAE buffer and purified using the QIAquick kit. Products were sequenced in both directions with an ABI® 310 DNA Automatic Capillary Sequencer (Applied Biosystems, Foster City CA, USA) using the DyeDeoxy™ Terminator sequencing method (Big Dye™, version 1.0 or 2.0) according to the manufacturer’s protocol except that the amount of Big Dye was generally reduced to 2 μL. Sequencing primers (1 μL) were used at a concentration of 3.2 pM/μL. Reactions were purified by ethanol precipitation or using CleanSeq magnetic beads (Agencourt) on the Corbett liquid handling system.

2.2. Data analysed

Electropherograms of the two sequence directions were checked using Sequence Navigator (Applied Biosystems 1994) and a consensus sequence was generated. Accession numbers are listed in Table 3. Data collected for this paper have accession numbers of DQ916496–DQ916508 for COI, DQ916585–DQ916605 for EFI-α, DQ916445–DQ916454 for H3, DQ916414–DQ916444 for 12S rRNA, DQ916572–DQ916584 for 28S rDNA D1, DQ916509–DQ916518 for 28S rRNA A, DQ916467–DQ916495 for 28S rRNA D6, DQ916543–DQ916571 for 28S rRNA D9-10, DQ916455–DQ916466 for 28S rRNA B and DQ916519–DQ916542 for 18S rRNA. The data include published sequences from this laboratory (Colgan et al., 2000, 2003, 2006): AF033716–AF033794 for 28S rRNA A and 28S rRNA B, AF033675–AF033715 for H3, AY296815–AY996850 for COI, and AY296873–AY296909 for 28S rRNA D1. Additional published sequence data from the following GenBank accessions were included: 28S rRNA D1, Viviparidae (U75863; Viviparus viviparus McArthur, 1996) and Valvatidae (U75862; Valvata sp. McArthur, 1996); 28S rRNA D6, Viviparidae (U82423; Viviparus sp. Tillier et al., 1994), Valvatidae (U78672; Valvata sp., McArthur and Koop, 1999), Aplysiidae (U78644; Aplysia californica, McArthur and Koop, 1999), Ampullariidae (U78643; Ampullaria sp., McArthur and Koop, 1999); 18S rRNA, Pleurotomariidae (L78893; Bayerotrochus midas, Harasewych et al., 1997). A representative range of the 16S data available in GenBank was downloaded and analysed as a separate dataset. The 16S sequences were not included in the overall compilation as there was too little overlap in family representation in the two datasets.

MCCLADE (Maddison and Maddison, 1992) was used for data manipulations such as joining files for the individual segments and specifying character sets. Tree figures were drawn with TGF (Müller and Müller, 2004).

2.3. Phylogenetic analysis

Sequences were aligned using the default values for parameters in CLUSTAL X (Thompson et al., 1997). The “slow-accurate” algorithm was used for pairwise alignment with costs of 10.0 for gap opening and 0.10 for gap extension. For multiple alignments, the cost for gap opening was set at 10.0 and gap extension at 0.20, with a DNA transition weight of 0.50 and a “delay divergent sequences” percentage of 30. Areas of uncertain alignment (found only in the segments from non-coding genes) were omitted from all reported analyses. In reporting the results, we adopt the conventions that (a) a comma separates monophyletic groups within clades specified by parentheses; (b) a + sign indicates that the group before the sign is paraphyletic with respect to the group following the sign; and (c) that the bootstrap or posterior probability is given immediately after the closing parenthesis. Posterior probabilities are written as “support” levels with the actual probability being multiplied by 100 to give an integer value.

Maximum parsimony analyses were conducted using heuristic searches in PAUP* 4.0 version beta 10 (Swofford, 2001). Analyses were performed using the tree-bisection-reconnection (TBR) branch-swapping algorithm for multiple replications of random stepwise addition of taxa. MULPARS was in effect. All characters were unordered and unweighted except where specified. The steepest descent option was not enforced. Zero length branches were collapsed to give polytomies. Gaps were treated as unknown in most analyses but as a fifth state in one series of the “sensitivity” analyses (see below).

Analyses were conducted separately for the combined data and for 16S rRNA using 1000 random sequence addition replicates. For bootstrap pseudo-sampling, heuristic searches were conducted for 200 bootstrap replicates, each with 20 random addition iterations. The searches using the combined data were repeated with the individual imposition of constraints in separate analyses that enforced monophyly of: Sorbeococoncha (Viviparidae plus Ampullariidae); Architaenioglossa (Ampullariidae, Viviparidae and Cyclophoridae); Rissooidea (Rissoidae and Anabathridae); Ptenoglossa (Cerithiopsidae, Eulimidae, Epitoniidae); or Neogastropoda (Muricidae, Volutidae, Nassariidae, Mitridae, Cancellariidae, Conidae).

Analyses were performed for various subsets of the data including the individual segments (and sub-segments within 28S rRNA D6 and CO1 when these were amplified in multiple reactions) to search for possible PCR artefacts. BLAST searches were made for taxa with long branch lengths in these analyses to confirm that the sequences are caenogastropod in origin. Analyses were performed using 200 replications of random stepwise addition of taxa. To prevent filling of the tree buffer by large numbers of equally long trees from a single replicate, no more than 200 trees (PAUP command nchuck = 200) longer than 50 steps (command chucklen = 50) were kept. For bootstrap pseudo-sampling for individual segments, heuristic searches were conducted for 200 bootstrap replicates, each with 20 random addition iterations where no more than 200 trees longer than 50 steps were kept.
Table 3
GenBank Accession Numbers for the dataset compiled here

<table>
<thead>
<tr>
<th>Family</th>
<th>Species</th>
<th>COI</th>
<th>EF1-α</th>
<th>H3</th>
<th>12S rRNA</th>
<th>28S rRNA D1</th>
<th>28S rRNA A</th>
<th>28S rRNA D6</th>
<th>28S rRNA D9-10</th>
<th>28S rRNA D B</th>
<th>18S rRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fissurellidae</td>
<td>Montfortula rugosa</td>
<td>AY296819</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pleurotomariidae</td>
<td>Bayerotrochus mids</td>
<td>AY296820</td>
<td>AY296879</td>
<td></td>
<td>Af033698</td>
<td>Af033762</td>
<td>Dq916483</td>
<td>Dq916569</td>
<td>Af033763</td>
<td>Dq916541</td>
<td></td>
</tr>
<tr>
<td>Neritidae</td>
<td>Nerita atramentosa</td>
<td>AY296824</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyclophoridae</td>
<td>Leptopoma perlucida</td>
<td>AY296826</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ampullariidae</td>
<td>Pomacea bridgesii</td>
<td>DQ916496</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vipaviridae</td>
<td>Bellamyia heidi guangdungenis</td>
<td>AY296827</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batillariidae</td>
<td>Zeacumantus subcarinatus</td>
<td>AY296833</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turritellidae</td>
<td>Maoricolpus roseus</td>
<td>DQ916497</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campanilidae</td>
<td>Campanile symbolicum</td>
<td>AY296828</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Littorinidae</td>
<td>Austrolittorina unifasciata</td>
<td>AY296829</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Etoniellidae</td>
<td>Crassionella flammaea</td>
<td>DQ916498</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anabathriidae</td>
<td>Pisina albizona</td>
<td>AY296830</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rissooidae</td>
<td>Rissoina fasciata</td>
<td>DQ916500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strombiidae</td>
<td>Strombus luahuanus</td>
<td>AY296831</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hipponicidae</td>
<td>Antiasia folacea</td>
<td>DQ916502</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calyptraeidae</td>
<td>Bostrychus pritzkeri</td>
<td>DQ916503</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xenophoridae</td>
<td>Xenophora indica</td>
<td>DQ916501</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vermetidae</td>
<td>Serpulorbis sp</td>
<td>AY296833</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cypraeidae</td>
<td>Cypraea annulata</td>
<td>AY296832</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Naticidae</td>
<td>Comber melanostoma</td>
<td>DQ916504</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ranelidae</td>
<td>Cabestana spengleri</td>
<td>AY296833</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tonniidae</td>
<td>Tonna cervinea</td>
<td>DQ916506</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pterotraceidae</td>
<td>Pterotracea cornucornus</td>
<td>DQ916505</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerithiopsidae</td>
<td>Axatocerithium sp.</td>
<td>AY296835</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eptionidae</td>
<td>Eptionum jukesiarae</td>
<td>AY296836</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eulimididae</td>
<td>Stilifer sp.</td>
<td>Dq916507</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nassariusida</td>
<td>Nassarius barchardi</td>
<td>AY296837</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muricidae</td>
<td>Dicathais orbita</td>
<td>AY296838</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mitridae</td>
<td>Mitra carbonaria</td>
<td>AY296839</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volutidae</td>
<td>Cymbiola spenceri</td>
<td>DQ916508</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cancellariidae</td>
<td>Cancellaria undulata</td>
<td>AY296841</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conidae</td>
<td>Conus miles</td>
<td>AY296840</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aplysiidae</td>
<td>Aplysia juliana</td>
<td>AY296846</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amphibolidae</td>
<td>Salinator solida</td>
<td>AY296845</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Gene segments not included in the data are indicated by blank cells. Accession numbers with the prefix DQ were collected for this study. All other data are from publications from this laboratory except the seven sequences specified in Section 2.2.
An Incongruence Length Difference Test (Farris et al., 1994) was conducted with PAUP (Swofford, 2001) to assess conflict between the character partitions defined by gene segments using maximum parsimony heuristic searches with 100 replicates, with 20 random additions of sequences in each and keeping no more than 200 trees greater than 20 steps in each replicate.

Two series of “sensitivity” analyses sensu Wheeler (1995) were conducted to ascertain if the relationships inferred from parsimony results were robust against changes in the assumed transformation matrix. All consistent combinations assuming that the costs of transitions, transversions and gaps could be 1, 2 or 4 were tested for the case where gaps were treated as a fifth state using 200 random taxon addition replicates and keeping no more than 100 trees greater than 400 steps in each. A similar series using the same search strategy was used for the case where gaps were treated as missing data, allowing the costs of transitions and transversions to take the values 1, 2 or 4.

Maximum likelihood analyses were performed with PAUP* using the parameters suggested by the Akaike Information Criterion test in MODELTEST (Posada and Crandall, 1998) for 30 replicates with random addition sequences for the overall data and 50 for other analyses. For the overall data, the model selected by MODELTEST was GTR + G + I. The parameter settings for the model were as follows. Initial base frequencies were A = 0.2809, C = 0.1594, G = 0.2419 and T = 0.3178). The number of substitution types was 6 with substitution rate matrix (A–C: 3.2339, A–G: 6.6870, A–T: 3.2945; C–G: 4.0560; C–T: 18.6100; and G–T: 1.0000). A discrete gamma distribution was assumed for each of the rate categories and an α parameter of 0.3550 was assumed. The proportion of invariant sites was set at 0.4137. When the third base positions were excluded, the model selected was GTR + G + I with the following settings. Initial base frequencies were A = 0.2440, C = 0.2320, G = 0.2619, and T = 0.2621. The number of substitution types was 6 with substitution rate matrix (A–C: 1.4467, A–G: 3.4592; A–T: 2.5336; C–G: 1.3119; C–T: 6.7463; and G–T: 1.000). A discrete gamma distribution with four rate categories and an α parameter of 0.3849 was assumed. The proportion of invariant sites was 0.4604. For the 16S rDNA data, the model selected was GTR + G + I with the following settings. Initial base frequencies were A: 0.2419, C: 0.3178, G: 0.2419 and T: 0.3178. The number of substitution types was 6 with substitution rate matrix (A–C: 2.1885, A–G: 10.4245; A–T: 3.2945; C–G: 4.0560; C–T: 18.6100; and G–T: 1.0000). A discrete gamma distribution was assumed for each of the rate categories and an α parameter of 0.4518 was assumed. The proportion of invariant sites was estimated as 0.1751.

Bayesian analyses were conducted using MrBayes (Version 3.1) (Huelsenbeck and Ronquist, 2001), Metropolis coupled, Monte Carlo Markov Chains were run for 1,000,000 steps and a number of generations (varying with the particular analysis) were discarded to allow for convergence. The discard cut-off point was determined by the rule that all trees remaining had likelihoods less than 0.2% worse than the final asymptote of the sample. Four differentially heated chains were run simultaneously. Topologies were sampled every 100 generations. Likelihood settings were determined during the run. Base frequencies were estimated, as were the independent rates of the six substitution types. A discrete gamma distribution was assumed for variation in the rate of substitution between nucleotide positions in the alignment and the shape parameter of this distribution was estimated. The estimation of the various likelihood parameters was conducted separately for each gene segment (and each codon position within coding sequences) using a character partition and the “unlink” command in MrBayes.

Abbreviations for the analyses principally discussed in this paper are given in the following list. As noted above, regions of uncertain alignment were excluded from all analyses.

- AMP: Maximum parsimony, combined data;
- AML: Maximum likelihood, combined data;
- ABY: Bayesian, combined data;
- AMP-P3: Maximum parsimony, combined data, excluding third codon positions;
- AML-P3: Maximum likelihood, combined data, excluding third codon positions;
- ABY-P3: Bayesian, combined data, excluding third codon positions;
- 16SMP: Maximum parsimony, 16S data;
- 16SML: Maximum likelihood, 16S data;
- 16SBY: Bayesian, 16S data.

3. Results

The alignments used in these analyses are available from the first author or from an accession in TREEBASE. In total, the alignment contains 3995 bases, of which 344 were excluded as belonging to regions of uncertain alignment. Of the remaining bases, 2164 showed no variation in the present sample of species, 1119 were parsimony-informative and 368 were variable but not parsimony-informative. χ-squared tests of the homogeneity of base composition in the various taxa gave a probability close to 1 for the 16S rRNA dataset and for all gene segments except CO1. When the base position data for the coding regions are considered individually, inhomogeneity was detectable for all three sets of third base position data but not for any other codon positions. The ILD test returned a probability of 0.91 for AMP and 0.65 for AMP-P3, thereby not rejecting the hypothesis of no character conflict between the individual segments.

3.1. Analyses of the combined dataset

The topologies for the AML, AML-P3, ABY and ABY-P3 analyses are shown in Figs. 2–5 respectively. The first 110,000 generations of the simulation were discarded for the calculation of posterior probabilities for ABY and the
first 80,000 generations for ABY-P3. Both Caenogastropoda and Hypsogastropoda were monophyletic in all of these topologies. Sorbeoconcha was contradicted in all, owing to the inclusion of Campanilidae in a clade containing otherwise only Ampullariidae or Cyclophoridae. Cerithioidea (Batillariidae and Turritellidae) and Tonnaidea (Ranellidae and Tonnidae) were well-supported monophyletic clades in all of these analyses. Neogastropoda was never monophyletic in any analyses of the compiled data although there was no strong evidence against its monophyly.

In the AML topology (Fig. 2), a clade of three neogastropods was the sister group to all other Hypsogastropoda. The two clades formed by the second division within Hypsogastropoda were (1) a group comprising the three other Neogastropoda, the Tonnaidea, Cypraeidae, Strombidae and Xenophoridae and (2) a group comprising Littorinidae, Anabathridae, Rissoidae, Eatonellidae, Vermetidae, Hipponicidae, Naticidae, Pterotracheidae, Cerithiopsidae, Eulimiidae, Epitoniidae and Calyptraeidae. These families except Calyptraeidae are referred to informally as the “GC group”, the name deriving from a shared sequence motif detailed below. The “GC group” was seen in the AML-P3 topology where Calyptraeidae was its sister group and in ABY-P3 where it received posterior probability support of 84. The group excepting Eatonellidae was recognisable in ABY with posterior probability support of 73. In ABY, Calyptraeidae was unresolved within Hypsogastropoda and in ABY-P3, it was the sister of the neogastropod family Cancellariidae with posterior probability support of 74.

One maximum parsimony tree was found for AMP (length 8251; consistency index (CI) 0.303). The bootstrap
supported clades in this analysis are shown in Table 4 and Fig. 2. The supported clades for the assumption that gaps are a fifth state are included for comparison with their treatment as missing data. In the AMP maximum parsimony tree, Turritellidae was included in a clade of neogastropods that formed one of two branches forming the basal division within Hypsogastropoda, thereby also contradicting Cerithioidea. The other neogastropod (Volutidae) was the sister group to Naticidae in a clade deriving from the third division in Hypsogastropoda. Trees constrained to show the “GC group” as monophyletic had a length of 8270.

The lengths of AMP trees constrained to show previously named groups as monophyletic were: (Viviparidae, Ampullariidae): two trees, of 8267 steps with a consistency index (CI) of 0.303; Architaenioglossa: one tree, 8271 steps, CI = 0.303; Rissooidea: two trees, 8252 steps CI = 0.303; Sorbeoconcha: one tree, 8284 steps, CI = 0.302; Ptenoglossa: four trees, 8276 steps, CI = 0.303; and Neogastropoda: one tree, 8263 steps, CI = 0.303. None of the constraints produced trees that were significantly longer than the unconstrained trees using either the Kishino–Hasegawa Test (KH Test) (Kishino and Hasegawa, 1989) or the non-parametric Templeton test (Templeton, 1983).

Two trees were found for the AMP-P3 analysis (length 3854; CI 0.411). The bootstrap supported clades found in this analysis are listed in Table 4. In both these trees, the inclusion of Viviparidae as the sister group to Eatoniiellidae and the sister group pairing of Strombidae and Cerithioidea at the base of the clade including all Hypsogastropoda contradicted monophyly of this group.

Pterotracheidae were generally closely associated with Littorinidae (Figs. 2–5). The families formed a sister group in all analyses, the pairing having bootstrap support of 65% in AMP and 56 in AMP-P3 and posterior probability support of 99 in ABY and 95 in ABY-P3. The two Rissooidea

Fig. 3. The majority rule consensus of the trees sampled during the ABY analysis of the combined data. Posterior probability support levels are given above branches. Named higher level taxa are indicated by bars to the right of the topology. Asterisks indicate that the named clade is not monophyletic.
(Rissoidae and Anabathridae) were never monophyletic in the combined analyses. Often Rissoidae (AML, AML-P3, ABY-P3, AMP-P3) formed an unsupported clade with Eulimidae (Figs. 2, 4, 5), sometimes with Epitoniidae as the sister group to this pair (AML, ABY, Figs. 2 and 3). The two tonnoidean families formed a generally well-supported clade in all combined analyses including AMP and AMP-P3 (Figs. 2–5) as did Strombidae and Xenophoridae, although with less support (Figs. 2–5) and excepting AMP and AMP-P3. Ptenoglossa was not monophyletic in any of the combined analyses. Architaenioglossa was never monophyletic. In all combined analyses except AMP where it was the sister group to Batillariidae and AMP-P3 where it was the sister group to Campaniliidae, Cyclophoridae was the sister group to (Campaniliidae, Ampullariidae). This clade was sister group to the rest of the caenogastropods in all likelihood and Bayesian analyses. Viviparidae was the next branching taxon in these analyses, having a sister group comprised of two monophyletic clades, Cerithioidea and Hypsogastropoda (Figs. 2–5).

A total of 18 non-redundant transformation matrices among those with parameter variation as specified in Section 2 were found to be consistent during PAUP processing. These are specified in the form $[a, b, c]$ where “a” refers to the assumed cost of transitions, “b” to the cost of transversions and “c” to the cost of inserting a gap. Topologies resulting from the searches always recovered monophyly of Caenogastropoda. Tonnoidea and (Littoniidae, Pterotracheidae) were monophyletic for all matrices. The sister pair (Xenophoridae, Strombidae) was sometimes observed ($[140], [240], [221], [122]$ and $[142]$ as was Cerithioidea ($[140], [120], [112], [122], [142]$,
Hypsogastropoda was never monophyletic, being shown to include Viviparidae or Turritellidae or to exclude Eatoniiellidae and/or Eulimidae. Monophyly of Architaenioglossa, Sorbeoconcha, the “GC group” or Neogastropoda was never recovered. Notably, however, as many as five of the six neogastropods were often found in a clade. For two matrices ([1 1 2] and [1 4 2]) this did not include other taxa.

### 3.2. Individual segment analyses

Details of the various datasets and the trees from maximum parsimony analyses of them are given in Table 5. Clades receiving bootstrap support of more than 50% in analyses of individual gene segments are specified in Table 4. Individual gene maximum parsimony analyses rarely recovered named higher taxa. Caenogastropoda was seen in the CO1-P3 (excepting Eulimidae), 28S rRNA D6 and 18S rRNA (excluding Eatoniiellidae) MP analyses. Sorbeoconcha was not seen in any individual gene maximum parsimony analysis. Hypsogastropoda was monophyletic only in 28S rRNA D6 but was supported in 18S rRNA except for the exclusion of Eatoniiellidae.

There were only three positions in the 28S rRNA B segment where two or more bases were observed in two or more taxa within Caenogastropoda. These were at position 2985 in the 28S rRNA sequence in GenBank accession AY145411 from *Ilyanassa obsoleta* (Passamaneck et al., 2004) where thymine was found in Turritellidae and all Hypsogastropoda except Eatoniiellidae, Strombididae and Nassariidae. Cytosine was found in these exceptions and in Architaenioglossa, Campanilidae, Batillariidae and outgroup taxa. Membership of the informal “GC group” was defined by the bases present at the other two variable positions. At position 3026 in the *I. obsoleta* sequence in
GenBank accession AY145411, cytosine is the most common form, but is replaced by guanine in Littorinidae, Anabathridae, Rissoidae, Eatonellidae, Vermetidae, Hipponicidae, Naticidae, Pterotracheidae, Cerithiopsidae, Eulimidae and Epitonidae. At position 3035 in the *Ilyanassa obsoleta* sequence, the situation is reversed with guanine, the common form being replaced by cytosine in the families just listed. These are apparently compensatory changes in the stems of a loop system.

Clades with bootstrap or posterior probability support greater than 50 observed in the 16SMP or 16SBY analyses are shown in Fig. 6 that illustrates the optimal 16SML tree. The first 120,000 generations of the simulation were discarded for the calculation of Bayesian posterior probabilities. In the 16SML tree, Caenogastropoda and Hypsogastropoda were recovered but Cerithioidea are paraphyletic. A clade containing all Neogastropoda was recognisable but was paraphyletic as it included a subclade comprised of Triviidae and Velutinidae. The three families of the “GC group” represented in this dataset were found in two monophyletic clades one of which (including Littorinidae and a number of other asiphonate families) was associated by short branches with clades containing predominantly siphonate taxa (Fig. 6), but also including Xenophoridae, Calyptraeidae, Velutinidae and Capulidae. The other clade, including only Rissoidae and Hipponicidae was shown as one of the sister groups formed by the basal division in Hypsogastropoda. Members of the Rissooidea were seen in both of these clades rendering this superfamily diphyletic.
Most analyses reported here, using the combined data, individual segments or the independent 16S rRNA dataset, support monophyly of Caenogastropoda, often with substantial statistical support. Caenogastropod monophyly has generally been supported, or only weakly contradicted, in studies using DNA sequences from a significant number of species (Harasewych et al., 1998; McArthur and Harasewych, 2003; Colgan et al., 2003). Monophyly is strongly supported by morphology (Ponder and Lindberg, 1997). Consequently the following discussion will focus on relationships within the group.

4.1. Sorbeoconcha and Architaenioglossa

Sorbeoconcha, as defined by Ponder and Lindberg (1997), included Campanilidae and Cerithioidea. This group is not observed in any of our analyses. Sorbeoconcha has never previously been supported by a large molecular dataset, generally owing to the exclusion of Campanilidae or to the inclusion of architaenioglossans within the smallest monophyletic clade including all of its supposed members (Harasewych et al., 1998; Harasewych et al., 2003; Colgan et al., 2003). Monophyly is strongly supported by morphology (Ponder and Lindberg, 1997). Consequently the following discussion will focus on relationships within the group.

Harasewych et al. (1998) conducted the first study of relationships at the base of the caenogastropods that included a substantial number of non-hypsogastropods. They sequenced parts of the 18S rRNA gene in 19 caenogastropods including five architaenioglossans (two Cyclophoridae, two Ampullariidae and one Viviparidae), Campanile and three cerithioideans. Cerithioidea, Ampullariidae and Cyclophoridae were monophyletic in several of their analyses (Harasewych et al., 1998), however the frequent exclusion of Viviparidae and the inclusion of Campanilidae contradicted monophyly of Architaenioglossa. Harasewych et al. (1998) found the non-hypsogastropod Caenogastropoda comprised a monophyletic group in maximum likelihood analyses and in parsimony analyses using only transversions. This intriguing clade was not observed in their other analyses. Nor was it obtained in later analyses of 18S rRNA gene. McArthur and Harasewych, (2003; herein). Monophyly of this clade within Hypsogastropoda was observed for two of 18 transformation matrices in the sensitivity analyses (this study). Monophyly requires a non-significant increase of 15 steps when compared to the unconstrained topology in AMP analyses (results not illustrated).

All molecular analyses containing Campanile, with the exception of the present 16S rRNA analyses, suggest that an architaenioglossan is the sister group of this family, a
Fig. 6. The optimal tree found in 16SML analysis of the 16S rRNA data. Brackets to the right of taxon names define taxonomic groups or statistically supported relationships. Taxa are monophyletic except where contradiction is indicated by an asterisk beside the group name. The node linking Caenogastropoda is identified by “C” and that linking Hyspogastropoda by “H”. The three families of the “GC group” represented in this analysis are written in bold. The scale bar is graduated in units of 0.01 substitutions per site. Posterior probability support levels greater than 60 and maximum parsimony bootstrap support values greater than 50 are written near nodes or to the right of linking brackets. Posterior probabilities are written first. Where only one figure is shown, this is a posterior probability support.
possibility first raised by Harasewych et al. (1998). The relationship they observed was with Cyclophoroidea, as also seen in the present analyses only in AMP-P3, rather than with Ampullariidae as suggested by the support for a sister-group pairing of these families in the likelihood and Bayesian analyses. A possible explanation for these affinities is that the taxon sampling of Cerithioidea is better for the 16S rRNA dataset than for the other sets used in this paper or in Harasewych et al. (1998).

4.2. Hypsogastropoda

Sorboconcha was divided into two major clades in Strong’s (2003) morphological analysis. Hypsogastropod monophyly was contradicted by the inclusion of Batillariidae (Cerithioidea) in the first major clade. This also included Vermetidae, Strombidae, Crepidulidae and Bithyniidae. The second major clade included only hypsogastropods (Littorinidae, Naticidae, Cypraeidae, Epitonidae and Neogastropoda). Both of these major clades were supported by only two homoplasous characters (Strong, 2003, Fig. 26). In contrast, hypsogastropod monophyly has often been observed in molecular analyses. It was recovered in all analyses in Harasewych et al. (1998) and McArthur and Harasewych (2003) and in all current analyses of the combined data excepting AMP and AMP-P3, with posterior probability support of 85 in both ABY and ABY-P3. In AMP, Turritellidae is the sister group to Conidae and in AMP-P3, Cerithioidea is a monophyletic sister group of Strombidae, this group of three taxa being sister to the AMP-P3, Cerithioidea is a monophyletic sister group of Cerithioidea being better for the 16S rRNA dataset than for the other sets used in this paper or in Harasewych et al. (1998).

4.3. The “GC Group” and inhalant siphons

Only one major group of taxa within Hypsogastropoda received support in a notable proportion of the present analyses. Further work, including the investigation of other families is required to establish whether this group is robustly monophyletic so it was here named only for convenience of reference. The name “GC group” recognises a shared sequence motif in the 28S rRNA B segment. The group comprises Littorinidae, Eatoniiidae, Rissoidae, Anabathridae, Vermetidae, Hipponicidae, Pterotracheidae, Cerithiopsidae, Epitonidae, Eulimidae and Naticidae. The group includes most studied Hypsogastropoda that lack an anterior siphon and the siphonate Cerithiopsidae. The other asiphonate taxa included in these studies were the limpet-shaped Calyptraeidae that was sometimes associated with the “GC group” and the tent-shaped Xenophoridae that was always excluded. The group was seen in its entirety in ABY-P3 (with posterior probability support of 84) and AML-P3. The group excepting Eatoniiidae was seen in ABY (with posterior probability support of 73) and is also discernible in AML, but with Calyptraeidae included as the sister group of the limpet-shaped Hipponicidae. The inclusion of the Cerithiopsidae in the “GC group” raises the possibility that the inhalant siphon in cerithiopsids (and the related Triphoridae that are not included in this analysis) may not be homologous with that in other hypsogastropods, a possibility supported by the putative stem group, Pseudozygopleuridae, having lacked a siphonal notch (Nützel, 1998).

Although Calyptraeidae are included in, or closely related to, the “GC group” in some analyses, they are not included in the group as hypothesised here as the sequenced representative does not share the characteristic 28S rRNA B motif. The lack of an inhalant siphon in this family may be convergent owing to the limpet form of its shell. Analogous situations are seen in other families that have at least some taxa with limpet-like shells. For example, considering Trichotropis and Capulus, two taxa sometimes included in a single family (Capulidae), the coiled taxon (Trichotropis) has a siphon while the limpet-shaped Capulus does not.

The presence of the siphon is indicated in the shell by the presence of an anterior apertural notch or siphonal canal. The siphon itself is an anterior extension of the mantle edge and is used to direct the inhalant water current. Anterior apertural notches have, however, been developed in a few groups to accommodate other organs; for example some Rissoidae (e.g. Rissoina) have a notch for a pallial tentacle and stromboideans have, in addition to their small siphonal canal, a notch for a stalked eye-bearing tentacle. A great variety of mantle/siphon and anterior shell notch character combinations is seen in Cerithioidea but in those that do bear a mantle edge siphon, it is formed differently from that in hypsogastropods (Simone, 2001). Ponder and Lindberg (1997) proposed that the Sorboconcha were differentiated from the architeneoglossan grade in that they switched from the pleisiomorphic exhalant control of mantle cavity water currents to inhalant control and that the formation of an inhalant siphon was a consequence of this switch. This change conferred significant evolutionary advantages which were enhanced by the, probably multiple, development of a siphon. In particular, it maximized the potential of the chemoresponse facilities of the osphradium enabling more efficient detection of predators, mates and prey (Lindberg and Ponder, 2001).

4.4. Neogastropoda

Neogastropoda was one of the few traditional clades maintained during the revolution in the understanding of
gastropod phylogeny in the 1980s and 1990s. The group is morphologically defined by multiple synapomorphies (Ponder, 1974; Taylor and Morris, 1988; Ponder and Lindberg, 1996, 1997; Kantor, 1996; Strong, 2003). For example, in Strong’s (2003) analysis, seven non-homoplous synapomorphies and seven homoplous synapomorphies supported monophyly and the group had a Bremer support of five.

Neogastropoda has usually been contradicted, albeit weakly, in molecular analyses (Harasewych et al., 1997; Colgan et al., 2000, 2003; Riedel, 2000; McArthur and Harasewych, 2003; this study). Harasewych et al. (1997) included only two other hypogastropods and two architaeognlossans but Neogastropoda was not resolved as monophyletic in their analyses of partial 18S rRNA sequences. There was little structure in their topologies, and even some putative neogastropod genera lacked support. Harasewych et al. (1998) recovered Neogastropoda (3 taxa) in a few of their analyses (MP, ML), but not all. In Colgan et al. (2003), at most two of the five studied neogastropods were included in a monophyletic clade exclusive of other taxa. Riedel’s (2000) analyses of several neogastropod families and a few other caenogastropods using 16S rRNA and 18S rRNA data also failed to recover monophyly with Marginellidae and, in the 16S rRNA dataset, Vexillum (Costellariidae) consistently falling within lower hypogastropods. Ficus (Ficidae) and Bufonaria (Bursidae) were nested within the other neogastropods in 18S rRNA analyses (Riedel, 2000).

The two main hypotheses that have been advanced regarding the origin of neogastropods are that they arose from an “archaeogastropod” or primitive “mesogastropod” (Ponder, 1974), or alternatively that they arose from a “higher” mesogastropod, usually considered to be a tonnoidea or sharing a common ancestor with that group (e.g., Fretter and Graham, 1962; Taylor et al., 1980; Taylor and Morris, 1988; Riedel, 2000). The latter hypothesis has received some support from morphological (Ponder and Lindberg, 1996, 1997) and ultrastructural data (e.g. Haszprunar, 1985; Healy, 1988, 1996). However, Strong’s analyses (2003) suggest other possible affinities for the group, the nearest relatives of Neogastropoda being Epitonidae, Cypraeidae and Naticidae (Tonnoidea were not represented in these analyses). Our analyses find relationships between neogastropod families (either singly or as groups) and Turritellidae, Tonnoidea, Stromboidea or Cypraeidae. The analyses do not suggest that the Neogastropoda are derived Hypogastropoda although there is a close relationship with Tonnoidea. In all analyses of the multi-gene data (except AMP-P3), one of the sister groups in the first division of Hypogastropoda is predominantly or solely composed of neogastropods.

In the fossil record, Neogastropoda are first recognized during the Cretaceous (Tracey et al., 1993; Bandel, 1993) and Kollmann (1982) and Taylor and Morris (1988) suggested that they originated during the early stages of that period. Numerous proposals for the stem neogastropod lineage have been made. These have included the Palaeozoic siphonate Subulitidae (Cox, 1960; Ponder, 1974), the Purpurinidae from the Triassic/Jurassic (Taylor et al., 1980; Kaim, 2004) and Maturifusus (Maturifusidae) from the early Jurassic to Cretaceous (Szabó, 1983; Schröder, 1995; Bandel, 1993; Riedel, 2000; Kaim, 2004). A possible Late Triassic maturifusid has also been reported (Nützel and Erwin, 2004). The lack of support for a late origin of Neogastropoda here emphasises the need for further work in determining the fossil history of the group.

4.5. Other clades within Hypogastropoda

With the exception of Tonnoidea (Riedel, 1995, as Cassoidea), few previously named clades within Hypogastropoda are supported in the present analyses. In agreement with Simone’s (2005) morphological analysis as well as earlier placements, Xenophoridae and Strombidae form a clade in likelihood and Bayesian analyses (Figs. 2–6). Our results strongly indicate the non-monophyly of the Rissooidea, a very large group of mainly small-sized, families that may have little in common other than superficial similarity. The two families (Anabathridae and Rissoidae) included in the main analyses form a clade with Eulimidae in AMP-P3, AML-P3 (Fig. 4) and ABY-P3 with posterior probability support of 69 (Fig. 5). The two Rissoidae included in the 16S analysis are separated from other rissooidean families (Fig. 6) being more closely associated, albeit with weak support, to the Hipponicidae and Cyclophoridae. The remaining rissooideans including Hydrobiidae s.l., Pomatiopsidae and Bithyniidae are predominantly freshwater taxa. They have a clade formed by Littorinidae and Provannidae as their sister group (Fig. 6).

The sister group pair of Rissoidae with Eulimidae that contradicts Ptenoglossa has low posterior probability support of 53 in ABY (Fig. 3) and a moderate bootstrap support of 68 in AMP-P3 but a high posterior probability of 98 in ABY-P3. Epitonidae was generally closely associated with (Rissoidae, Eulimidae) in overall analyses but not when third codon positions were excluded. 16S data were not available for ptenoglossans. Eulimidae differ from other ptenoglossan families in having a concentrated nervous system and a penis and its members lack the distinctive parapodium seen in other ptenoglossans (Healy, 1988). The traditional concept of Ptenoglossa has not been tested in morphological analyses, although Nützel (1998), using fossils, argues for a sister relationship of Janthinoidea (including Epitonidae) and Triphoroidea. The only support for this suggestion among the many explorations of the present data was that (Epitonidae, Cerithiopsidae) was shown in maximum parsimony analyses of all data including the areas of uncertain alignment (results not reported).

One further clade within Hypogastropoda, the sister pairing of Littorinidae and Pterotracheidae (representing “Heteropoda”), is consistently supported by all analyses of the combined data. This clade received bootstrap support of 65% in AMP and 56% in AMP-P3, and posterior
probability support of 99 in ABY and 95 in ABY-P3. The taxa are clearly closely related as also suggested by Strong and Harasewych (2004) and clearly, as suggested in Bouchet and Rocroi (2005), “Heteropoda” does not warrant “subordinal” status.

4.6. Why are relationships within Hypsogastropoda difficult to resolve?

The general pattern of caenogastropod evolution suggested by the present analyses, in the context of the current understanding of morphology-based phylogenies and classifications, is that several basal clades, including Cyclophoroidea, Ampullarioidea, Viviparoida, Campaniloidea and Cerithioidea were early offshoots from the lineage leading to the pre-eminently successful Hypsogastropoda. As the known members of the architaenioglossan groups are all non-marine, their marine ancestors are assumed to be extinct. The hypsogastropods diversified greatly to form multiple major lineages including the morphologically well-supported Neogastropoda and a group proposed here that includes most of its asiphonate members and Cerithiopsidae representing Triphoroidea.

Despite their diversification in the Mesozoic, few hypogastropod lineages are definitely known to be present in the Palaeozoic. Pseudozygopleuroidea, while not monophyletic (Nützel, 1998), includes Palaeozoic members similar in shell features to some pleonoglossans (Bandel, 2002). At least two families that may belong to this group, Palaeozygopleuridae (questionably caenogastropods) and Pseudozygopleuridae, survived across the Permian–Triassic boundary (Bandel, 2002). Two recent families, Abyssochrysidae and Provannidae, are included in the “zygopleurid group” along with the Pseudozygopleuridae by Bouchet and Rocroi (2005) but outside the hypsogastropods. However, the anatomy and, particularly, the sperm morphology (see Healy, 2000 for references and discussion) of both these recent taxa indicate that they are probably hypogastropods. Although molecular data from Abyssochrysidae are not available, Provannidae are shown as the sister group to Littorinidae in 16S rRNA analyses (Fig. 6). If these families are indeed “zygopleurid”, Palaeozoic members of that group are presumably all or part Hypsogastropoda. There appear to be few other possible Palaeozoic hypogastropods (Bandel, 1993, 2002; Nützel and Mapes, 2001), suggesting that the group’s dominance of the marine gastropod fauna began after the Permian–Triassic extinction (Nützel, 2005). At least three extant hypogastropod families (Rissoidae, Carinariidae and Aporrhaidae) were definitely present in the mid-Jurassic (Tracey et al., 1993; Bandel, 2002; Gründel, 1999), with a few others such as “Hydrobiidae”, Lamelliporidae and Epitonidae appearing later in the Jurassic (Tracey et al., 1993).

The difficulty of resolving relationships within Hypsogastropoda could be due to a rapid early radiation of the group. This idea is consistent with the lengths of basal hypsogastropod branches in analyses of the combined data and 16S (particularly MP). These branches are short, as would be expected if the initial divisions within the group occurred quickly. The ecological vacuum following the Permian–Triassic extinction may have presented the opportunity for an explosive radiation of the Hypsogastropoda, although this is not yet supported by the available fossil record.

Acknowledgments

We thank the Australian Research Council and the Australian Museum for financial support of this project and the late Ken and Yasuko Myer for general financial support of the Australian Museum’s Evolutionary Biology Unit. We thank Edwina Rickard and Andrea Feilen for assistance with some of the sequencing and Sue Livingston, Denis O’Meally and Rebecca Johnson for running the reactions on automatic sequencers. We thank Dr. R. Seapy, California State University, Fullerton for providing a sample of Pterotrachea coronata.

References


