Chapter 5: A Total Evidence Cladistic Analysis of the World-Wide Haliotidae.

INTRODUCTION

Two studies concerning the phylogeny of the Haliotidae have been published to date. Brown (1993) took a phenetic approach and used allozyme frequencies for 17 nominal species to infer the phylogeny with Nei's genetic distance. Lee & Vacquier (1995) used the 16kD protein called 'lysin' to infer the phylogeny of 22 abalone species employing a parsimony-based approach. The taxonomic overlap between the two studies is 13 taxa. The two studies are in broad general agreement with one another (Figures 5-1 and 5-2). First, they both show a north Pacific clade, comprised of the west American species and the large Japanese species (*H. discus, H. gigantea*) to the exclusion of the smaller Japanese species such as *H. diversicolor*. The second second shared clade comprises the taxa endemic to Australia, to the exclusion of wide-spread taxa also occurring in Australia (e.g*., H. ovina, H. varia*). For the remainder, the low taxonomic sampling density precludes any further general statements. For instance, of the five species endemic to South Africa (*H. midae, H. parva*, *H. queketti, H. spadicea, H. speciosa*) only *H. midae* was included in these studies. Both papers have a serious deficiency, in that the type species of the genus, *H. asinina*, was not included. Hence, the tentative taxonomic recommendation on the genus level made in Lee & Vacquier (1995) are lacking a sound base (see also Geiger, 1996, 1998a, for discussion).

MATERIALS AND METHODS

Source of specimens

Table 1 gives the sources of all data and specimens used in this study. No animals could be located for the following taxa considered valid (cf. chapter 1): *H. crebrisculpta, H. cyclobates, H. exigua, H. pulcherrima, H. sorenseni,* and *H. squamosa*. These taxa are not listed in the morphological data matrices. Lee & Vacquier (1995) had some tissue of *H. sorenseni* available, hence, only the other 5 of the total of 56 taxa will not be included in any of the data analyses of the family.

SEM preparation

Preparation and viewing of radulae as well as desiccation of epipodia followed the protocols of Geiger (1996) and Stewart & Geiger (1999).

Cladistic methodology

My basic understanding of the cladistic method has been presented in detail in Chapter 3 (= Geiger & Fitzhugh, in review).

For the analyses presented here, the following parameters were used. All characters were unordered and equally weighted. No differential transition/transversion of gap-cost were used, and GAPMODE = NEWSTATE was set. Due to the number of taxa analyzed, heuristic search options of PAUP 3.1 (Swofford, 1991) were used, with random taxon addition in effect and random seed number entered manually. Between 10 and 1,000 heuristic searches were performed depending on how quickly the number of equally parsimonious trees remained stable. Subsequent searches with Tree Bisection and Reconnection (TBR) on non-minimal length trees to escape possible local minima were performed. The maximum length of the subminimal trees was dictated by imple-

Table 5-1: Source of specimens and data. -: no data available. Abbreviation for source of specimens furnishing morphological characters (Morphology column): ANSP: Academy of Natural Sciences, Philadelphia. BMNH: British Museum of Natural History. BO: Buzz Owen Collection, Gualala, California. CASIZ: California Academy of Sciences, Invertebrate Zoology, San Francisco, California. DLG: Daniel L. Geiger collection, Los Angeles. JK: Joan Koven collection, Astrolabe Inc., Washington (DC). LACM: The Natural History Museum of Los Angeles County. S (1998): Simone (1998). SAM: South Australian Museum, Sydney. SBMNH: Santa Barbara Museum of Natural History. USNM: United States National Museum, Washington, (DC). WAM: Western Australian Museum, Perth.
Other abbreviation : W (unpubl.): Wray, unpublished data. L&V (1995): Lee & Vacquier, 1995. B (1993): Brown, 1993.

Table 5-1.

Taxon	Morphology	16S mtDNA	lysin	allozyme
asinina	DLG			
	LACM 85-2	W (unpubl.)	-	-
aurantium	S (1998)	-	-	-
australis	DLG	W (unpubl.)	L&V (1995)	B (1993)
brazieri	SAM C32701-05			
	SAM C149014	-	-	-
clathrata	DLG			
	USNM 795269	-	-	-
coccoradiata	BMNH 1887.2.9.168-168a			
	SAM C328730-32	-	-	B (1993)
corrugata	DLG			
	ANSP 50503	W (unpubl.)	L&V (1995)	B (1993)
cracherodii	LACM 54780			
	LACM 79-23	-	L&V (1995)	B (1993)
cyclobates	-	W (unpubl.)	L&V (1995)	B (1993)
dalli	ANSP 315293	-	-	-
discus	LACM 82-12			
	CASIZ 44960	W (unpubl.)	L&V (1995)	B (1993)
dissona	JK	-	-	-
diversicolor	LACM 82-20	W (unpubl.)	L&V (1995)	-
dohrniana	DLG	-	-	-
elegans	USNM 360940			
	WAM S1003-4	-	-	-
fatui	USNM 486708	-	-	-
fulgens	LACM 146819	W (unpubl.)	L&V (1995)	B (1993)
gigantea	BMNH 1878.3.26.2	-	-	-
glabra	BMNH 1854.2.19.117			
	ANSP 230038	W (unpubl.)	-	-
hargravesi	SAM C327854-57			
	SAM C149015	-	-	-
iris	DLG	W (unpubl.)	L&V (1995)	B (1993)
jacnensis	LACM 113210	-	-	-
kamtschatkana	CASIZ 21339	W (unpubl.)	L&V (1995)	B (1993)
laevigata	LACM 87-97	W (unpubl.)	L&V (1995)	B (1993)
madaka	BMNH	W (unpubl.)	L&V (1995)	-
mariae	DLG	-	-	-
marmorata	BMNH no #	-	-	-
midae	DLG			
	CASIZ 20810	W (unpubl.)	L&V (1995)	B (1993)
ovina	LACM 85-3	W (unpubl.)	L&V (1995)	-
parva	DLG	-	-	-
planata	DLG	-	-	-
pourtalesii	USNM 856447	W (unpubl.)	-	-
pulcherrima	-	-	-	-
pustulata	DLG	-	L&V (1995)	-
queketti	DLG	-	-	-

Table 5-1: continued

roberti	SBMNH no#	W (unpubl.)	-	-
roei	LACM	W (unpubl.)	L&V (1995)	B (1993)
rubiginosa	SAM C328696-700			
	SAM C352383	-	-	-
rubra	LACM 71-495	W (unpubl.)	L&V (1995)	B (1993)
rufescens	LACM 1813-49	-	L&V (1995)	B (1993)
rugosa	DLG	W (unpubl.)	-	-
scalaris	LACM 87-100			
	WAM S1007-8	W (unpubl.)	L&V (1995)	B (1993)
semiplicata	WAM S1005-6	-	-	-
sorenseni	-	W (unpubl.)	L&V (1995)	-
spadicea	DLG	-	-	-
speciosa	DLG	-	-	-
squamata	SBMNH no#	-	-	-
	SAM C328733-41	-	-	-
stomatiaeformis	BO	-	-	-
tuberculata	DLG	W (unpubl.)	L&V (1995)	-
unilateralis	DLG	-	-	-
varia	DLG			
	LACM 85-5	W (unpubl.)	L&V (1995)	-
virginea	CASIZ 44954	-	-	B (1993)
walallensis	LACM 1813-49	W (unpubl.)	L&V (1995)	-

mentation limits in PAUP 3.1, not memory (RAM) limitations. PAUP 3.1 can not store more than approximately 30,000 trees internally; the precise number of maximally storable trees depends on the number of taxa as well as the number of characters. In cases of multiple most parsimonious resolutions, a strict consensus tree is shown. Tree statistics refer to any of the fully resolved topologies, not the strict consensus tree.

Abbreviations used are CI (consistency index), RI (retention index), RC (rescaled consistency index), and MPR (maximum parsimonious resolution).

Cladograms were prepared in MacClade 3.04 (Maddison & Maddison, 1992) and further modified in McDraw™ II 1.1 (Claris, 1987).

Outgroup comparison

Characters were polarized through outgroup comparison where possible. Outgroups were chosen from all member of Vetigastropoda, because the relationships amongst Vetigastropoda are currently unresolved (cf. Baten, 1975; Haszprunar, 1988b; Tillier *et al*., 1994; Hickman, 1996; Harasewych *et al*., 1997). The character states were derived from both original observations (Pleurotomariidae: AMNH 187302, AMNH 3483, AMNH 19855. Trochidae: LACM 46-32.12. Lepetodrilidae: DLG. Fissurellidae: LACM 45-1.5) as well as from the literature (Hickman, 1984; Haszprunar, 1988c; McLean, 1984b, 1988; Franz, 1989; Hickman & McLean, 1990; Sasaki, 1998). If multiple taxa were available, the most basal taxon was chosen according to available phylogenetic hypotheses (Trochidae, Turbinidae: Hickman, 1996. Pleurotomariidae: Harasewych *et al*., 1997).

ALLOZYME FREQUENCIES

The use of the allozyme frequency data provided by Brown (1993) provides a particular challenge, because the original frequency data had to be transformed into character-state data. In its most basic form character coding is a problem of belief representation. This problem can be subdivided into three areas:

- What is the cut-off level at which a low frequency occurrence of an observation is still considered to represent a character state found in a species?
- How are the electromorphs of proteins stained with the same stain treated? Do these different alleles belong to the same locus, or are they from different loci? As different character states in one character resulting in massively polymorphic taxa, or using presence / absence coding for each electromorph resulting in a clean, binary data matrix?
- A third but minor problem is when are two electromophs given the same identifier? This is a similar problem to when we call two legs the same length.

For the first problem the general statistical significance level of 0.05 can be used for the inclusion/exclusion argument. For sample sizes of up to 20 individuals every electromorph found was included in the binary matrix, whereas single instances of particular electromorphs in sample sizes between 21 and 40, and double instances in sample sizes between 41 and 60 were excluded. No larger sample sizes were used by Brown (1993). Assuming that the specimens used were a representative sample from the entire population, this treatment compares favorably with practices in morphology and is more thorough than what typical sample sizes in studies with sequence data allow. For morphological studies usually far fewer than 20 specimens are studied, and for sequence

data rarely more than a single specimen is assessed. Hence, the 0.05 inclusion/exclusion argument is rather conservative when compared to other types of data.

The second problem is the most difficult. Stains are considered to be specific for a particular enzyme, however, it is not neccessarily known whether that enzyme is produced by a single- or a multi-copy gene. Addtionally, epigenetic bands may have been identified as discrete allozymes; without the original gels, the nature of the band (epigenetic, actual allozyme) can not be resolved. Consequently, the question of homology or orthology versus paralogy becomes central. A second problem concerns the the staining method, which is a substrate modification to produce a color in the presence of a particular enzyme. It is assumed that the substrate reaction is highly specific for the enzyme in question. One can certainly entertain the possibility that one substrate may be utilized by more than one target protein, hence, staining protein may not be a reliable indicator of similarity. I have opted here to take the position of the *advocatus diaboli*, meaning that I assumed multi-copy genes and multiple use of substrates due to absence of evidence to the contrary. As a necessary consequence, all electromorphs are coded in a presence/absence matrix. Such a practice may strike some as inappropriate, but such are my subjective beliefs, which are open to reinterpretation, of course, by subsequent workers.

On the other hand I believe that bands of the same electric motility and same staining properties provide good grounds to postulate their similarity and, hence, to explain them by means of primary homology, further assuming that Brown (1993) did not identify epigenetic bands as discrete allozymes. The similarity is based on topographical identity being the same motility and character-state identity being the same staining property. For a more detailed discussion of belief formation and its inherent subjectivity see Chapter 3 (= Geiger & Fitzhugh, in review).

Table 4-2 gives the recoded allozyme frequency data of the 16 abalone taxa used by Brown (1993). Note that *H. conicopora* has been synonymyzed under *H. rubra* by Geiger (1998a), therefore, the data from these two nominal taxa were combined. Data from multiple populations of the same species were averaged.

Outgroup comparison

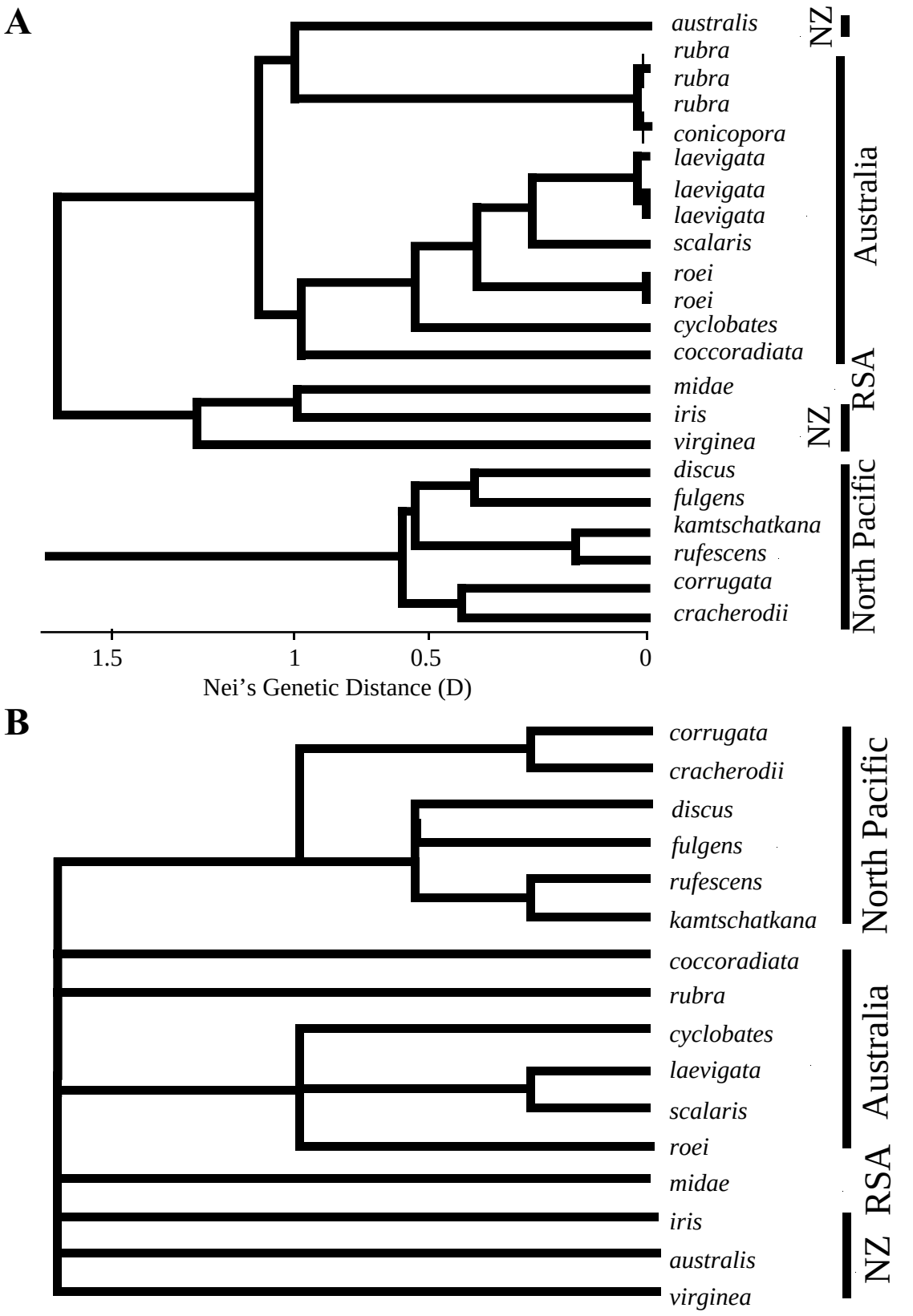
As the identification of the electromorphs is relative to the other taxa studied, addition of data to the published ones is virtually impossible without repeating most of the original study. Consequently, the data-set for the ingroup is limited to that of the original ingroup, and outgroup comparison is impossible. The cladistic analysis of that data-set, hence, will show unrooted cladograms, or networks. These networks may be rooted anywhere within the tree, for which reason it is inappropriate to call groups of taxa clades. I will, therefore, rather use the informal word 'grouping' to highlight some of the particular finds from the analysis of that data-set.

Reanalysis of allozyme data with comparison to the results of Brown (1993)

An initial search with TBR on minimal trees only produced eight equally most parsimonious trees of 117 steps (CI = 0.479, RI = 0.555, RC = 0.266). A search with TBR on non-minimal length trees of up to 120 steps found four additional topologies of 117 steps. This case illustrates clearly the importance of branch swapping on non-minimal length trees in heuristic searches. A branch-and-bound search with upper bound = 117 steps found the same 12 trees.

Brown's (1993) tree is fully resolved (Figure 5-1: A), whereas the reanalysis of the allozyme frequencies as character state data using parsimony is characterized by a basal polytomy (Figure 5-1: B). Was a lot of information lost during the transformation of

Figure 5-1. Trees generated from allozyme frequency data (Brown, 1993). A) Neighbor joining tree of Brown (1993) redrawn, with geographic distribution added. B) Strict consensus tree of 12 equally most parsimonious resolutions from the allozyme data only. 56 informative characters, 117 steps, CI = 0.479, RI = 0.555, RC = 0.266. NZ: New Zealand. RSA: Republic of South Africa. *Haliotis conicopora* is synonymized under *H. rubra*. Note the North Pacific group, and the majority of endemic Australian species form one group. The three New Zealand species, however, are either dispersed over the tree (A), or are are united in a basal polytomy (B).


A
australis
rubra
rubra
rubra
conicopora
laevigata
laevigata
laevigata
scalaris
roei
roei
cyclobates
coccoradiata
midae
iris
virginea
discus
fulgens
kamtschatkana
rufescens
corrugata
cracherodii
NZ
Australia
RSA
NZ
North Pacific
1.5
1
0.5
0
Nei's Genetic Distance (D)
B
corrugata
cracherodii
discus
fulgens
rufescens
kamtschatkana
coccoradiata
rubra
cyclobates
laevigata
scalaris
roei
midae
iris
australis
virginea
North Pacific
Australia
RSA
NZ

frequency data to character state data? For this question a particularity of distance methods such as neighbor joining using Nei's genetic distance has far reaching consequences. The distance algorithm ignores other potential resolutions and is particularly sensitive to the initial, most similar taxa. Although the higher resolution of the topologies is touted by advocates of phenetic methologies as a clear advantage (see Wenzel, 1997, for discussion), the irrelevance of the analytical procedure to evolutionary questions has been lost from sight. As a consequence, the patterns resulting from the parsimony analysis are the only meaningful ones, because they are based on the Darwinian notion of descent with modification, and hence also on a cause - effect relation of speciation and observed distributional patterns of character states among taxa. Resolution is secondary, as polytomies may be real, the so called 'hard' polytomies (Wenzel, 1997).

The distance topology and the recoded parsimony topology do not squarely contradict one another, despite the relatively low resolution of the parsimony tree. The common elements are as follows. The north Pacific species and most of the species endemic to Australian are grouped together. The three New Zealand species as well as the single south African species are united in the large basal polytomy, hence, no statement about their precise relationships can be made. To appease advocates of compatibility analysis prior to combination of data-sets, the present analysis does not contra-indicate combination with the other data (see below). However, the compatibility argument is vacuous, because the primary quest of a phylogenetic analysis is explanation of observations; inclusion or exclusion of data in a total evidence analysis depends solely on the relevance of the data to the question addressed (see Chapter 3 = Geiger & Fitzhugh, in review). The presence or absence of allozymes as properties of organisms is relevant beyond any doubt.

DNA CODING STRATEGIES FOR 16KD CDNA

The DNA-sequence data on a sperm acrosomal protein (16kD lysin) published by Lee & Vacquier (1992, 1995) was incorporated as follows. All positions were included in the present analysis, though the original alignment and data representation was somewhat altered as will be explained below. Character number refers to the recoded matrix. The new coding strategies are those introduced in Chapter 3 (= Geiger & Fitzhugh, in review), hence, I will only briefly summarize the procedures here. For a detailed justification see Chapter 3.

'Data contraction' reduces a stretch of bases or amino acids of questionable alignment to a single position. It uses a higher level of generality for the observations at hand. It was used in both the translated region (= open reading frame. ORF: positions 21, 153) as well as the 3' untranslated (UTR: positions 5, 6, 20, 75, 132) region.

'Stretch coding' introduces (aut-)apomorphic character states for stretches of questionably aligned sequence in a few taxa only. It is directly comparable to practices in morphology, where the introduction of new character states is used for observations that cannot be classified in existing states. It has been used in positions 45-46, 101-104, and 111-114 in the UTR of the lysin cDNA.

Table 4-3 gives the recoded data from Lee & Vacquier (1995), with explicit notes on the coding strategies employed.

In order to obtain the most information from the data, additional recoding strategies were explored. It can be argued that the use of the amino-acid sequence for coding regions of genes such as the ORF of lysin offers more phylogenetic meaningful information, because problems due to the redundancy of the genetic code and the limited number of character states—four bases as opposed to twenty amino acids—are avoided. This approach can be taken further, noting that coding regions form functional entities.

Hence, it may not be the actual amino acids, which show the greatest amount of information, but the functional groups. The amino-acid sequence was recoded according to the seven functional groups (Stryer, 1995) and re-analyzed. The strict consensus tree of 174 MPRs was characterized by a large polytomy (tree not shown). Accordingly, this approach was abandoned and the amino-acid sequence of the lysin's ORF was used for further analysis.

Outgroup comparison

Outgroup comparison is unfortunately impossible. The lysin protein of a trochid has been sequenced but is only 50% of the length of the abalone lysin (Hellberg, 1998). Alignment is impossible and even the question of whether these functional proteins are homologous, or in molecular parlance 'orthologous', arises (cf. Hellberg & Vacquier, 1999). However, as lysis is under positive Darwinian selection (Lee & Vacquier, 1992; Vacquier & Lee, 1993; Lee *et al*., 1995; Swanson & Vacquier, 1995), such high dissimilarities of putatively homologous genes between families are not surprising, but do not help with outgroup comparison either. For other Vetigastropoda, no lysin data are available. Hence, I searched for unrooted trees.

Reanalysis of the lysin data

Lee & Vacquier (1995) partitioned the data into the ORF and the UTR. My initial treatment will follow their data partitioning to show that the effect of recoding the data is distinct from the effect of combining the data-sets. For the same reason, the taxa were the same as used by Lee & Vacquier (1995). Differences between their study and the present study include the above mentioned alignment and coding variation, but also extend to the treatment of gaps. Whereas Lee & Vacquier (1995) treated gaps as miss-

Figure 5-2. Trees generated from data of Lee & Vacquier (1995). EM-RS: European-Mediterranean and Red Sea. IP: wide-spread Indo-Pacific. NZ: New Zealand. RSA: Republic of South Africa.

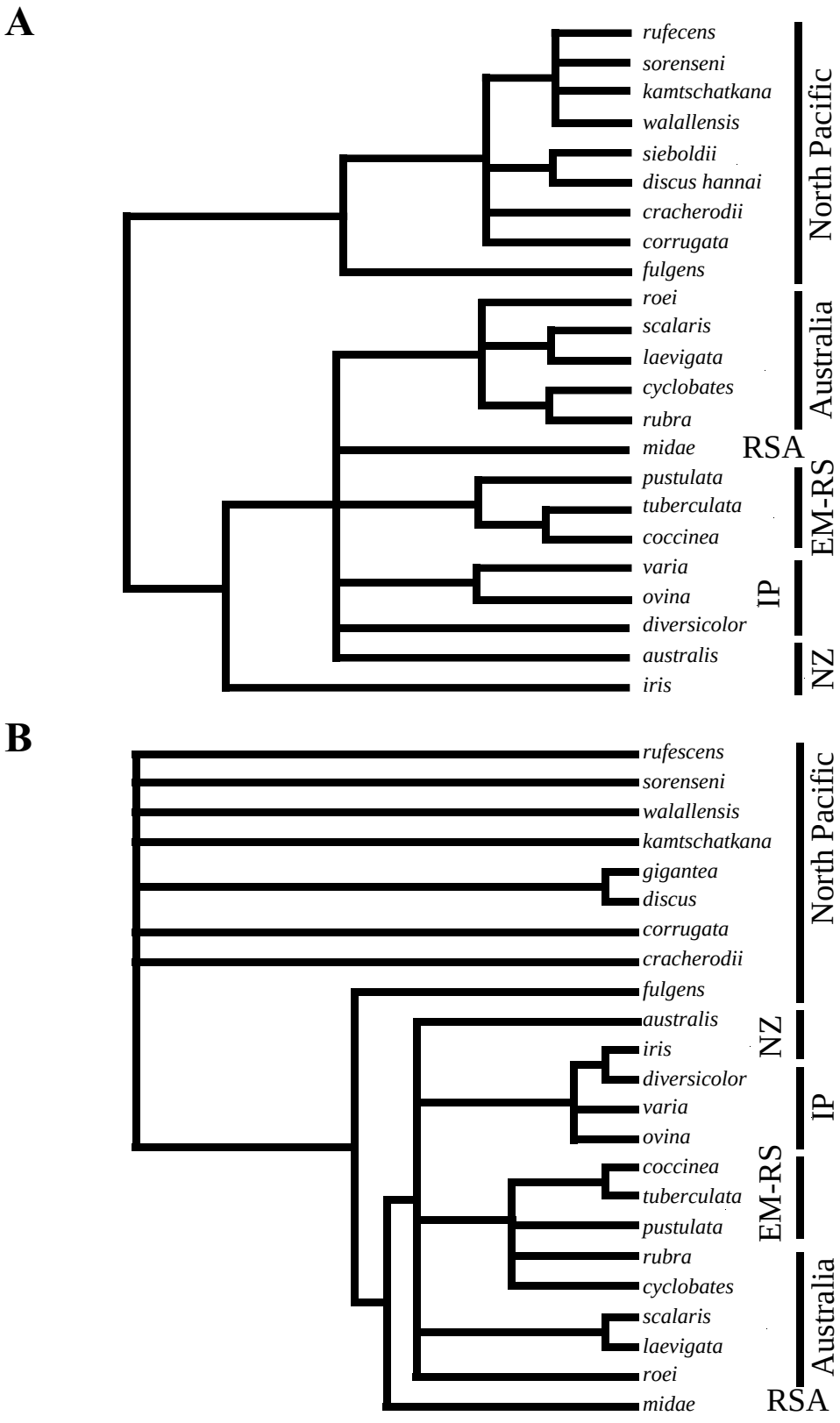
A) ORF tree redrawn from Lee & Vacquier (1995: fig. 4a) with geographic distribution superimposed.

B) ORF strict consensus tree from recoded data. 316 steps, CI = 0.560, RI = 0.673, RC = 0.377.

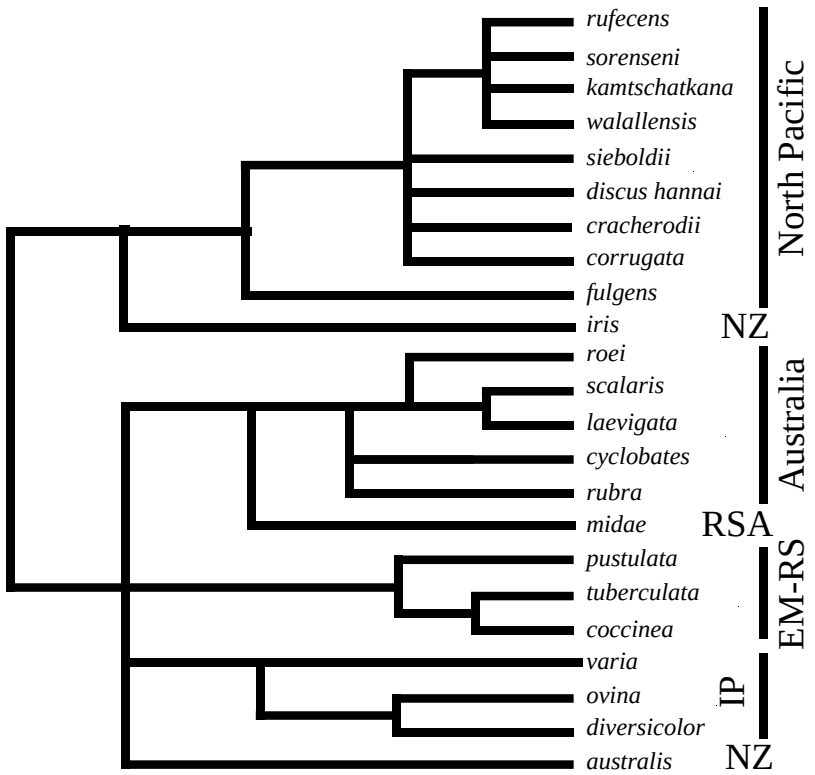
C) UTR tree redrawn from Lee & Vacquier (1995: fig. 5a) with geographic distribution superimposed.

D). UTR strict consensus tree from recoded data. 423 steps, CI = 0.709, RI = 0.858, RC = 0.609.

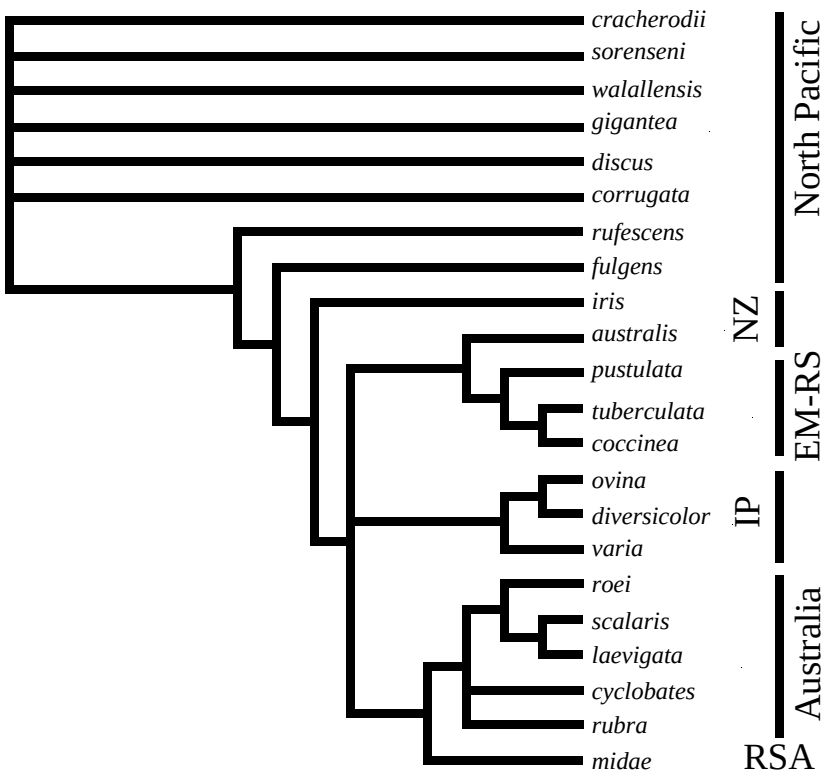
E) Strict consensus tree of ORF and UTR combined. 781 steps, CI = 0.647, RI = 0.787, RC = 0.509.


A
rufecens
sorenseni
kamtschatkana
walallensis
sieboldii
discus hannai
cracherodii
corrugata
fulgens
roei
scalaris
laevigata
cyclobates
rubra
midae
pustulata
tuberculata
coccinea
varia
ovina
diversicolor
australis
iris
North Pacific
Australia
RSA
EM-RS
IP
NZ
B
rufescens
sorenseni
walallensis
kamtschatkana
gigantea
discus
corrugata
cracherodii
fulgens
australis
iris
diversicolor
varia
ovina
coccinea
tuberculata
pustulata
rubra
cyclobates
scalaris
laevigata
roei
midae
North Pacific
NZ
IP
EM-RS
Australia
RSA

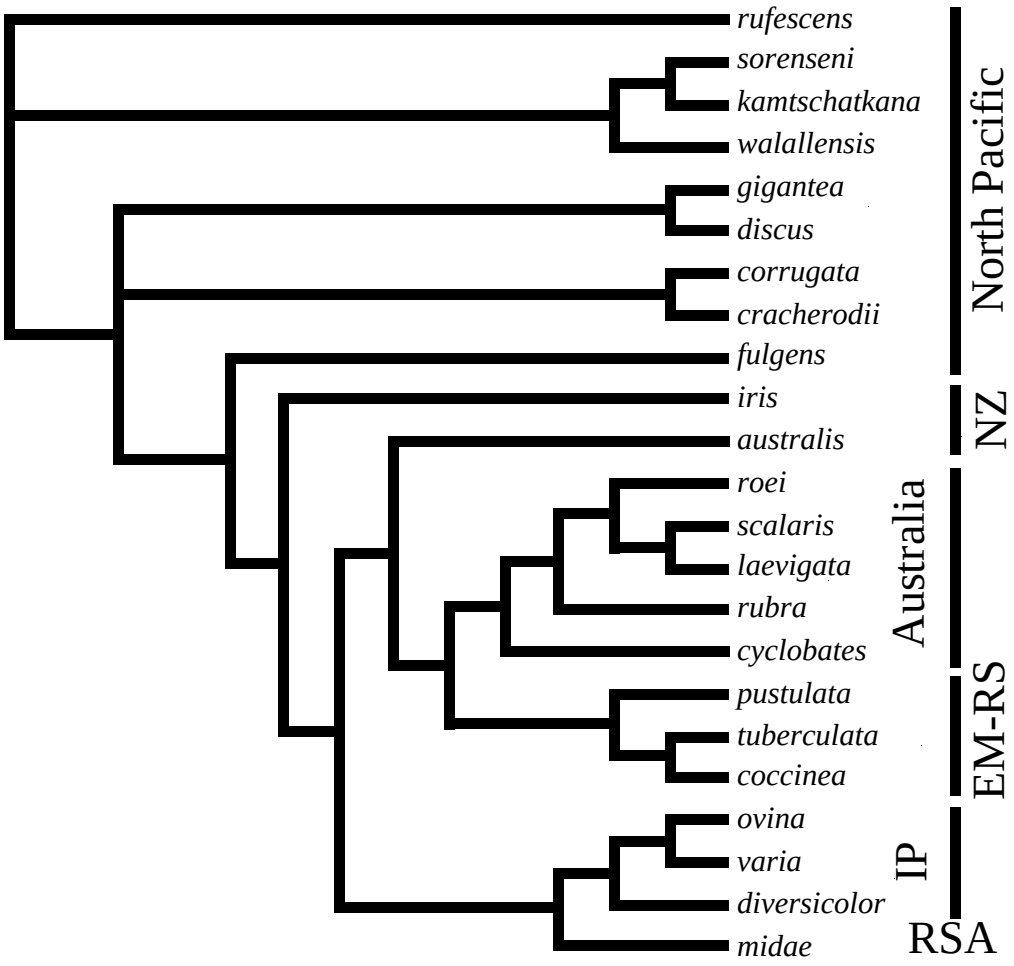
C


rufecens
sorenseni
kamtschatkana
walallensis
sieboldii
discus hannai
cracherodii
corrugata
fulgens
iris
roei
scalaris
laevigata
cyclobates
rubra
midae
pustulata
tuberculata
coccinea
varia
ovina
diversicolor
australis
North Pacific
NZ
Australia
RSA
EM-RS
IP
NZ

D


cracherodii
sorenseni
walallensis
gigantea
discus
corrugata
rufescens
fulgens
iris
australis
pustulata
tuberculata
coccinea
ovina
diversicolor
varia
roei
scalaris
laevigata
cyclobates
rubra
midae
North Pacific
NZ
EM-RS
IP
Australia
RSA

E


rufescens
sorenseni
kamtschatkana
walallensis
gigantea
discus
corrugata
cracherodii
fulgens
iris
australis
roei
scalaris
laevigata
rubra
cyclobates
pustulata
tuberculata
coccinea
ovina
varia
diversicolor
midae
North Pacific
NZ
Australia
EM-RS
IP
RSA

ing information (gapmode = missing), Geiger & Fitzhugh (in review = Chapter 3) argue for a fifth character state (gapmode = newstate), which is followed here.

ORF (Figures 5-2: A and 5-2: B): a heuristic search with TBR on non-minimal length trees of up to a tree length of 318 using 1,000 replications revealed 120 equally most parsimonious trees of 316 steps (CI = 0.560, RI = 0.673, RC = 0.377). Figure 5-2: B shows the strict consensus tree. When compared to the tree published by Lee & Vacquier (1995: fig. 4; Figure 5-2: A) the two correspond fairly well in that they both show a North Pacific group. In Lee & Vacquier (1995) the larger remaining taxa are joined in a large polytomy containing, however, groupings that unite the species endemic to Australian, the widespread species of the Indo-Pacific, as well as the European and Red Sea species. The two New Zealand species are separated, though. This rather surprising pattern of multiple colonizations of New Zealand implied by this analysis agrees with finds from other groups of organisms such as beeches (*Notofagus*), African violets (Gesneriaceae), a caddis fly genus (*Hydrobiosella*), and platycercine parrots discussed by Humphries & Parenti (1986). My re-analysis shows rather similar patterns in that the North Pacific species form one group (although less resolved), and the remaining taxa are joined in a large polytomy (Figure 5-2: B). Note that the two Japanese species (*H. discus*, *H. gigantea*) are now located in this large polytomy, whereas they were previously found nested within the west American abalone species (*H. fulgens* and up). It seems then that the position of the two Japanese species in Lee & Vacquier's original analysis stems from arbitrary alignment. Within the latter polytomy, however, the species endemic to Australian are not in a single group, the European and Red Sea species are closely associated with one of the species endemic to Australian, and one of the New Zealand species (*H. iris*) is nested within the widespread Indo-Pacific grouping. It is clear that changes in coding does have a significant impact on the phylogenetic

hypotheses generated. Re-evaluation particularly of the homoplastic character state changes as implied from my coding strategies did not reveal any coding problems. Therefore, I interpret the inferred homoplasies that remained despite the careful recoding carried out to be best explained by multiple changes at these sites, and not due to additional coding problems.

UTR (Figures 5-2:C and 5-2:D): thirty-eight equally most parsimonious trees of length 423 were found with branch swapping on non-minimal length trees of up to 425 steps and 1,000 replications (CI = 0.709, RI = 0.858, RC = 0.609). This re-analysis also grouped the North Pacific species, the species endemic to Australian, the Red Sea - European species, and the wide-spread Indo-Pacific species , with the two New Zealand species (*H. australis, H. iris*) being separated by two nodes. The difference between my (Figure 5-2: D) and Lee & Vacquier's (1995: fig. 5a; Figure 5-2: C) analyses lies in the resolution within some of these groups. The north Pacific group shows three nodes in Lee & Vacquier's work, whereas two nodes separate *H. fulgens* and *H. rufescens*, respectively, from a large polytomy in the re-analysis. The position of the Japanese species undergoes the same changes as discussed for the ORF above. Additionally, the position of the two New Zealand species differs between the two analyses, though the New Zealand species represent individual speciation events, and not a small radiation on this remote island group.

ORF combined with UTR (Figure 5-2:E): A heuristic search with TBR and 10 replications found two equally parsimonious topologies of 781 steps. Subsequent searches with TBR on non-minimal length trees (≤ 785 steps, 100 replications, CI = 0.647, RI = 0.787, RC = 0.509) found the same two equally most parsimonious trees. A branch-and-bound search with upper bound of 781 was aborted after almost 100 hours of search on a 604/120MHz processor that did not get passed in approximately 22% of all topologies

to be evaluated. Branch-and-bound searches become very problematic with more than 20 taxa, unless the data-matrix is extraordinarily clean (D. Kizirian, pers. comm.). Molecular data matrices in general do not satisfy such a condition. The strict consensus tree of these two trees is shown in Figure 5-2: E.

As the consensus tree is unrooted, I will identify groups rather than clades as the rooting may be placed anywhere on the tree resulting in non-monophyly of apparent clades. The first group, comprising the north Pacific species, contains the only two trichotomies found. Note that the Japanese species are again nested *within* the west American species. The second group contains the species endemic to Australian. Third, a small wide-spread Indo-Pacific group can be identified. Note the dispersed occurrence of the two of the three New Zealand taxa (*H. australis, H. iris*), which are separated by one node implying again separate colonization events.

It is interesting to note that the Red Sea *H. pustulata* forms a group with the European taxa *H. tuberculata* and *H. t. coccinea*. This pattern resounds with one of three biogeographical hypotheses discussed by Geiger & Groves (1999 = Chapter 2) and Geiger (in press = Chapter 4), the Tethys origin. This model is based on published chromosomal count data, which suggests an origin of the abalone in the Tethys, with subsequent eastward dispersal to the Indo-Pacific, and then to the north Pacific. However, due to the limited taxon sampling of this reanalysis, I want to de-emphasize this find rather, than to use it as a corner stone of the Tethys hypothesis. One may also note the apparent derived position of the (*pustulata* (*tuberculata, coccinea*)) group, which would be a counter indication to the required basal position in support of the Tethys model. Additionally, the Red Sea is geologically speaking a relatively young rift-shaped depression of approximately 20 My of age (Reiss & Hottinger, 1984).

PARTIAL 16S MTDNA (CHARLEY WRAY, UNPUBL. DATA)

Twenty-six unpublished 16S mtDNA sequences were kindly made available to me by Charley Wray (see Table 5-1), a former post-doctoral associate at the University of California, Berkeley. The sequences of 540 bp were treated with my new coding strategies; forty-six positions of questionable alignment were excluded (positions 1-18, 257-266, and 337-354). The sequences are not shown here, because they are unpublished data of Charley Wray. Figure 5-3 shows the strict consensus topology of 15 MPRs (Length = 189; CI = 0.434; RI = 0.698; RC = 0.303).

For outgroup comparison, Genbank was searched for alignable vetigastropod sequences, but none could be found. The network (Figure 5-3) shows the two familiar groups being the north Pacific species and the Australian endemic ones. The Indo-Pacific is polyphyletic in contrast to the combined analysis discussed below. The two New Zealand species do not form a group within this data-set either.

The combination of the lysin ORF and UTR, the allozyme frequencies, and the 16S mtDNA data yielded three MPRs (Figure 5-4). The features shared with the already discussed topologies are the North Pacific group including the tropical eastern Pacific *H. roberti* and the Caribbean *H. pourtalesii*, and the species endemic to Australia to the exclusion of the wide-spread taxa also occurring in tropical Australia (*H. ovina*, *H. varia*). The three New Zealand species (*H. australis*, *H. iris*, *H. virginea*) are distributed over the entire tree, which shows that either three different colonization events occurred for this remote group of islands, or that the abalone fauna of New Zealand is a relict fauna from an ancient stock, but is certainly not a discrete radiation on these islands. The wide-spread Indo-Pacific taxa (*H. diversicolor*, *H. ovina*, *H. varia*) and *H. glabra* with a more restricted distribution in the Indo-Malayan archipelago (see Chapter 4 = Geiger, in press) form another group to the exclusion of the enigmatic type species of

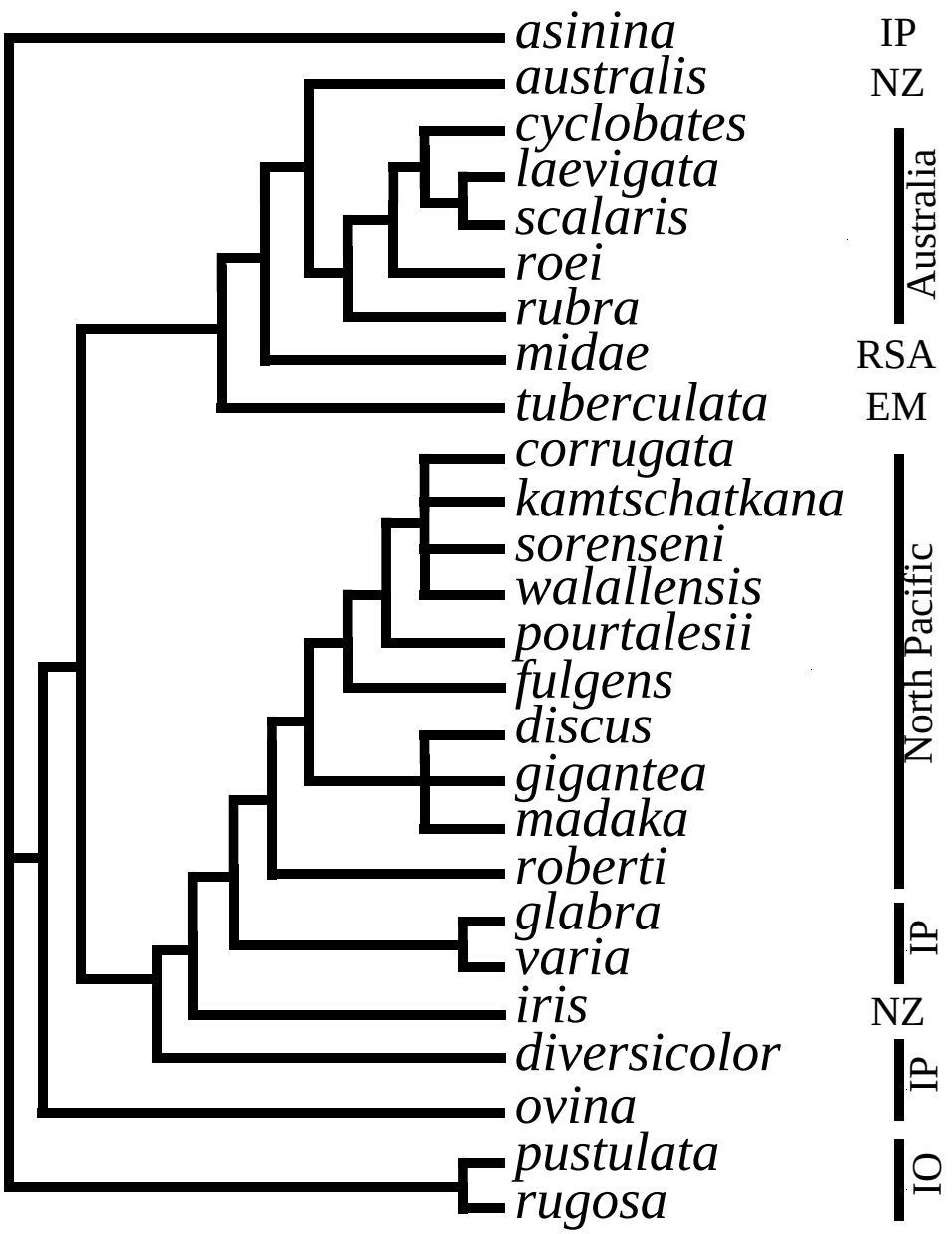


Figure 5-3. Strict consensus of 15 MPRs topology of 16S mtDNA sequences (Wray, unpubl. data). EM: European-Mediterranean. RS: Red-Sea. IO: Indian Ocean. IP: Indo-Pacific. NZ: New Zealand. RSA: South Africa.

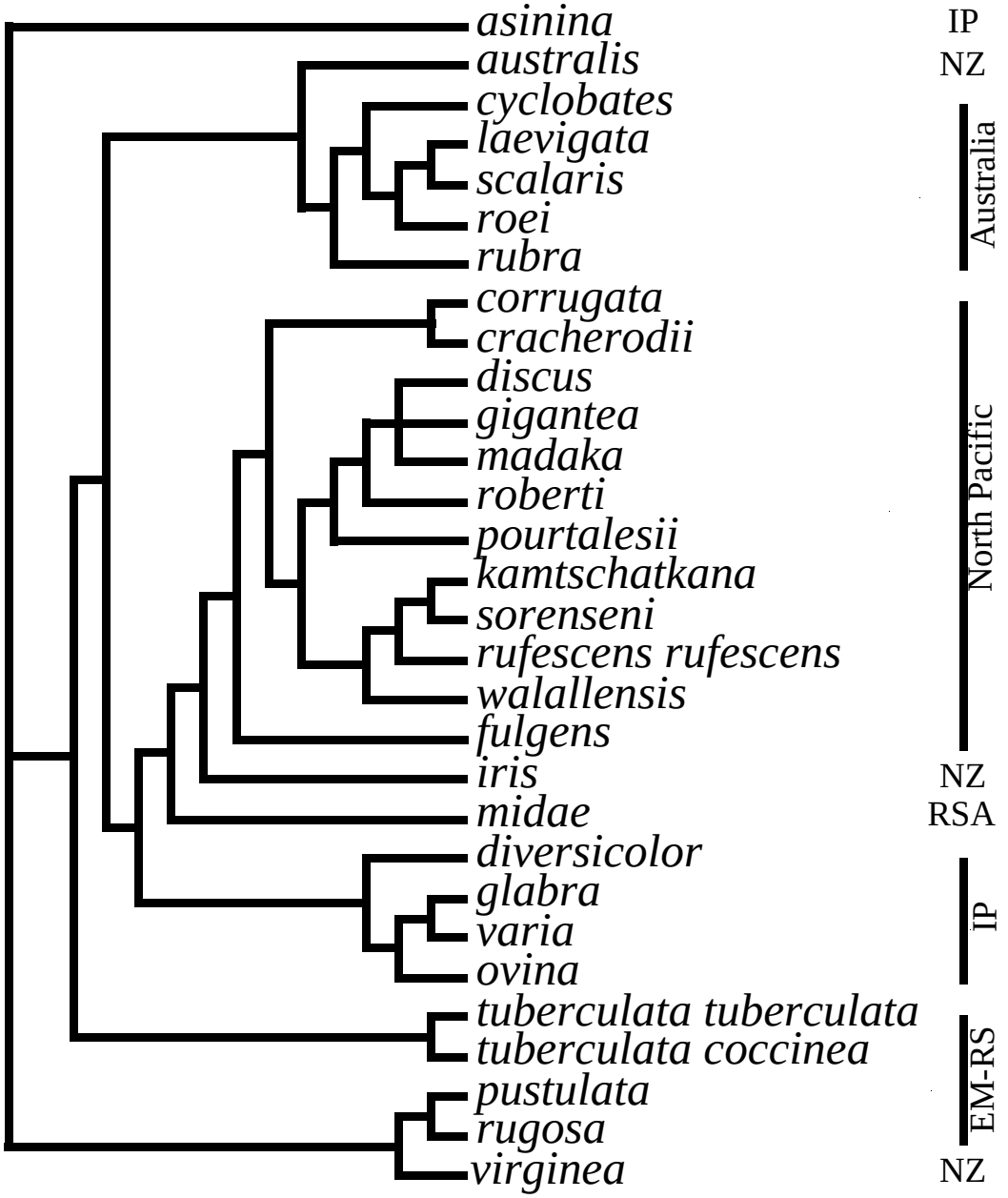


Figure 5-4. Strict consensus tree of 3 MPRs from all allozyme (Brown, 1993), lysin (Lee & Vacquier, 1995), and 16S mtDNA (Wray, unpubl.) data combined. EM-RS: European-Mediterranean-Red-Sea. IP: Indo-Pacific. NZ: New Zealand. RSA: South Africa.

the genus (for discussion see Chapter 1 = Geiger, 1998a). This shows that the very distinct shell morphology of *H. asinina* is also reflected in its position on the phylogenetic tree.

MORPHOLOGICAL CHARACTERS

The radula in abalone

The radula is one of the unique structures of mollusks and has played a very important role in gastropod systematics and classification (e.g., Thiele, 1931). It is a chitinous band with attached chitinous teeth. A radular formula is commonly used, which reads number of rows x marginal teeth + lateral teeth + rachidian (R) + lateral teeth + marginal teeth. Warén (1990) has shown that during ontogeny the marginal teeth are derived from lateral teeth and that the distinction of laterals and marginals is, therefore, artificial. However, because this distinction is so widely accepted, and has descriptive utility, it is used here as well. The arrangement and fine structure of the teeth has been important for the recognition of taxa at virtually every level of gastropod classification. However, the utility of the radula to address species level problems in the family Haliotidae had been called into question (Talmadge, 1956; Barnard, 1963). These early claims have been disproved subsequently (Geiger, 1996; 1998a, 1999a; Simone, 1998; Stewart & Geiger, 1999). Table 5-2 gives the data matrix for the radular characters; the character states are discussed below. Geiger (1996) introduced terminology to identify various parts of the teeth of abalone radula (Figure 5-5) and this terminology has been adopted for this account.

The radular structure and the radular formula change during early growth. The radula is seen first in the post-torsional veliger (Crofts, 1937; Dinamani & McRae, 1986). Competent larvae have at least three rows of radular teeth (Tong & Moss, 1992). At

Pleurotomariidae	0	0	0	0	-	0	-	-	0	0	0
Scissurellidae (*Scissurella*)	1	1	1	1	1	1	0	0	1	0	0
Lepetodrilidae (Lepetodrilus)	1	1	0	0	-	1	0	0	0	1	0
Fissurellidae (*Fissurella*)	1	0	0	0	-	1	0	0	0	0	0
Trochidae (*Tegula*)	0	1	1	0	0	1	0	0	1	1	0
Turbonidae (*Liotia*)	1	1	1	1	1	1	0	0	0	1	0
asinina	0	0	0	0	0	0	0	1	0	0	0
aurantium	0	0	0	0	0	1	0	0	0	0	1
australis	0	0	0	0	0	0	0	0	0	1	0
brazieri	1	0	0	0	1	0	0	0	0	0	0
clathrata	1	1	1	1	0	1	0	0	5	0	0
coccoradiata	1	0	1	0	0	0	1	0	0	1	0
corrugata	0	0	0	1	1	0	0	0	0	1	0
cracherodii	0	1	1	1	1	0	5	0	1	0	0
dalli	1	0	1	1	0	1	0	0	0	0	1
discus	1	0	1	1	1	0	0	0	0	1	0
dissona	1	1	1	1	0	1	0	0	0	0	0
diversicolor	0	0	0	0	0	0	1	0	1	0	0
elegans	0	0	0	0	0	1	0	0	1	0	0
fatui	1	1	0	1	0	1	0	0	0	0	0
fulgens	0	0	1	1	1	0	1	0	1	1	0
gigantea	0	0	0	1	0	0	0	0	0	0	0
glabra	1	1	1	1	1	1	0	0	0	0	0
hargravesi	0	0	0	0	1	1	0	0	0	0	0
iris	1	0	0	1	0	0	1	0	0	0	0
jacnensis	0	1	1	1	1	1	0	0	0	1	0
kamtschatkana	0	1	1	1	1	0	0	0	1	0	0
laevigata	0	0	0	1	0	0	1	0	1	0	0
madaka	0	0	1	1	1	0	0	0	0	0	0
mariae	0	0	0	1	0	1	1	0	0	0	0
marmorata	0	0	1	0	0	0	0	0	0	0	0
midae	0	0	1	1	0	0	0	0	0	0	0
ovina	1	1	0	0	0	0	0	0	1	0	0
parva	0	0	0	0	0	1	1	0	0	0	0
planata	1	1	1	1	0	1	5	0	0	0	0
pourtalesii	0	0	0	0	0	1	0	0	0	0	1
pustulata	1	5	0	1	0	1	5	0	0	0	0
queketti	1	0	0	0	0	1	0	0	1	0	0
roberti	1	0	1	1	0	1	0	0	0	0	1
roei	1	0	0	0	0	0	1	0	1	0	0
rubiginosa	0	0	0	1	1	0	0	0	0	0	0
rubra	0	0	0	0	0	0	1	0	0	0	0
rufescens	0	0	1	1	1	0	0	0	1	0	0
rugosa	1	1	5	1	0	0	0	0	0	0	0
scalaris	0	0	0	0	0	0	5	0	1	0	0
semiplicata	1	0	0	0	0	0	0	0	0	0	0
spadicea	1	1	1	1	1	0	0	0	0	0	0
speciosa	1	0	0	0	0	1	1	0	0	0	0
squamata	1	0	0	0	0	1	0	1	1	0	0
stomatiaeformis	1	0	1	0	0	0	0	0	0	0	0
tuberculata tuberculata	1	0	0	0	0	1	1	0	0	0	0
tuberculata coccinea	0	0	0	0	0	0	?	?	0	0	0
unilateralis	1	5	5	1	0	1	5	0	0	0	0
varia	1	1	1	1	0	1	0	0	0	0	0
virginea	1	1	0	1	1	0	0	0	0	1	0
walallensis	1	0	1	1	1	0	1	1	1	0	0

Table 5-2. Character states for the radular characters. For description of character states see text. 5 = (0,1). - = inapplicable.

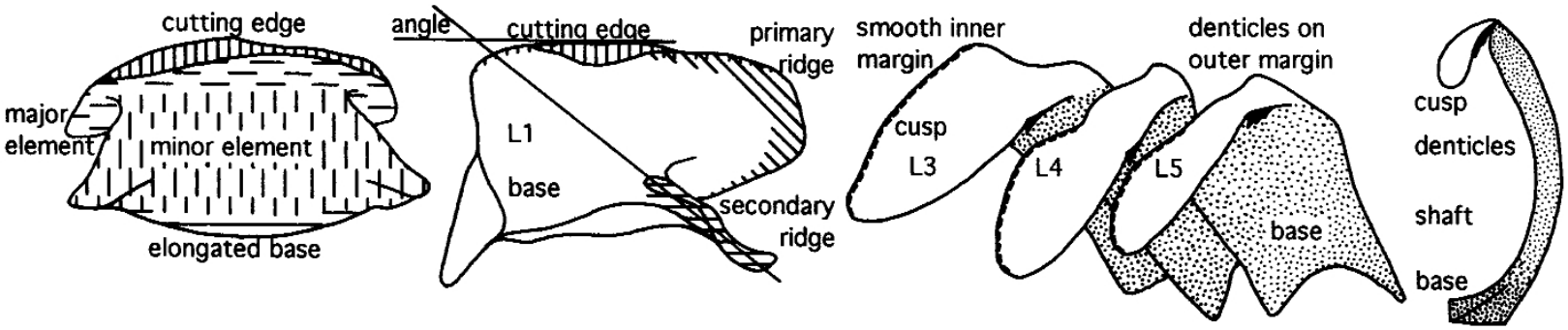


Figure 5-5. Diagram illustrating the terminology used for the radular teeth found in the family Haliotidae. The diagram is taken from Geiger (1996).

metamorphosis, the radula formula is >6-8 x 3 + R + 3 (Tong, 1985; Dinamani & McRae, 1986), and marginal teeth and additional rows are formed during subsequent growth (Dinamani & McRae, 1986). The radular formula for a three-week old juvenile is 5 + 4 + R + 4 + 5 (Crofts, 1929). The rachidian tooth of this stage is pluricusped and similar to those of other small vetigastropods (Pleurotomariidae: Hickman, 1984a. Haliotidae: Garland *et al*., 1985; Tong, 1985; Dinamani & McRae, 1986. Trochoidea: Hickman & McLean, 1990; Warén, 1990. Scissurellidae: Thiele, 1912; Marshall, 1993). The fully developed radular pattern, with its five lateral teeth, is found after two months (Crofts, 1929). The radula formula given by most authors for mature specimens is ∞ + 5 + R + 5 + ∞ (Fischer, 1885; Herbert, 1990 and references therein), where '∞' signifies a large but unspecified number. However, Dai & Wu (1989), Wu & Huang (1989), and Wu (1991) specified the number of marginal teeth between 55 and 65 for the various species studied.

Outgroup comparison

The radulae of Haliotidae differ in various aspects from those of other Vetigastropoda. The rows of the radular teeth are almost symmetrical in abalone, but are distinctly asymmetrical in pleurotomariids. The rachidian tooth is well-formed in Haliotidae, but reduced in Pleurotomariidae. The fine outer marginal teeth in Haliotidae show denticulate cusps (Wu & Huang; 1989; Herbert, 1990; Geiger, 1996; 1999a; Stewart & Geiger, 1999); in Pleurotomariidae, however, a fan of articulated bristles is found (Hickman, 1984a; Harasewych & Askew, 1993; Anseeuw & Goto, 1996). A comparison of Pleurotomariidae and Haliotidae to Scissurellidae is not appropriate as discussed above. The independence of radular morphology and feeding ecology has to be questioned due to the extensive morphological plasticity of the radula in response to the

feeding ecology of the respective animals. The radulae of Pleurotomariidae and Haliotidae will not help to resolve their phylogenetic relationship, because the former is that of a specialized spongivore, the latter that of a strict macroalgal herbivore.

The coding of paedo- and peramorphic structures adds further problems, as in the case of the radular characters of Scissurellidae and Haliotidae. One could consider stage-specific structures, which would overstate the degree of differentiation. In this case, the serrate rachidian of mature Scissurellidae is coded differently from the rachdian with a smooth cutting edge in Haliotidae. Alternatively, heterochronic processes are taken into account, creating characters with inapplicable character states. In this second case, the rachidian of Scissurellidae and Haliotidae < 5 mm is coded as serrate for both families, but the character state of the rachidian for animals > 10 mm is inapplicable to Scissurellidae. The use of the radula to resolve family level relationships within Vetigastropoda is, therefore, questionable (see also Haszprunar, 1993).

I have coded the radulae of the outgroups according to apparent shared similarities. It could also be argued that most character states should be coded as inapplicables, but I consider a clearly specified shared similarity statement—though admittedly tentative—more meaningful than the escape route of inapplicables.

Characters and character states

The characters and their states are illustrated in three figures. Figure 5-6 shows the modifications of the central field (rachidian tooth and lateral tooth 1). Figure 5-7 shows modifications on the lateral teeth 3-5 and Figure 5-8 illustrates the variation in the marginal teeth.

1: Posterio-basal projection on rachidian tooth (Figure 5-6). Absent: 0; Present: 1.

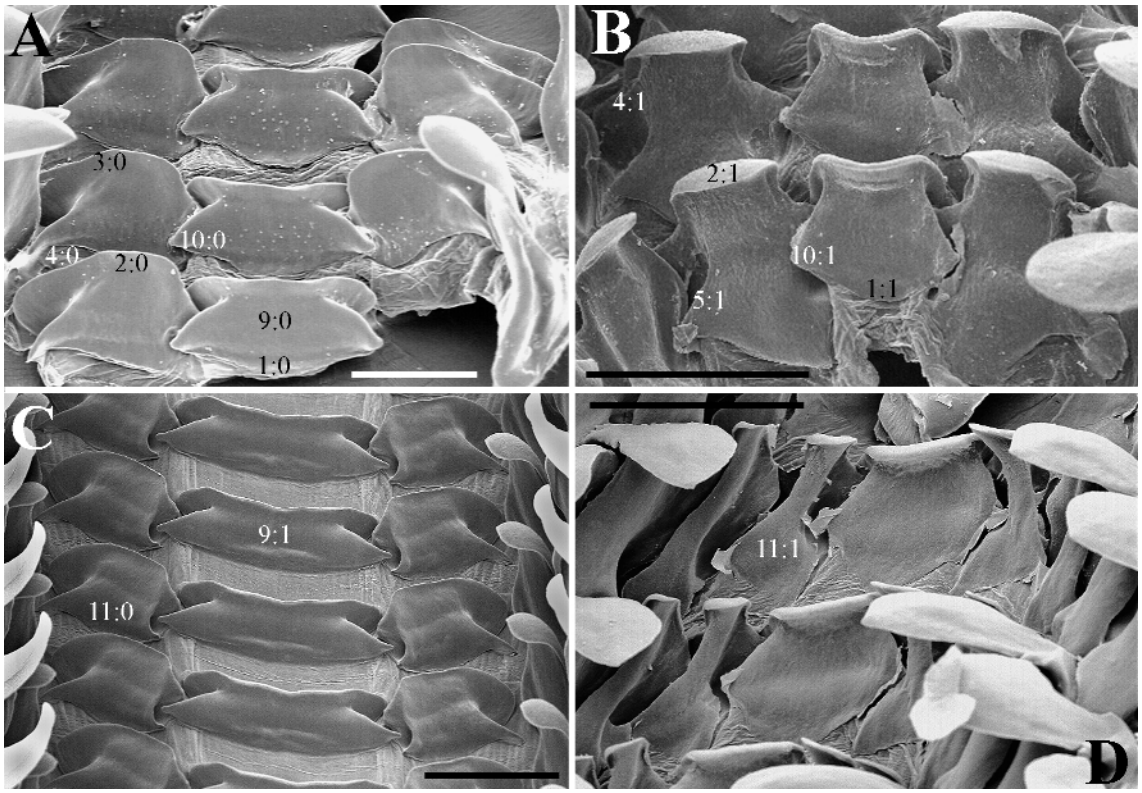


Figure 5-6. Illustration of radular character states for central field. Number on images refer to character number. A. *Haliotis. stomatiaeformis*. Scale bar = 100 µm. B. *Haliotis glabra*. Scale bar = 200 µm. C. *Haliotis elegans*. Scale bar = 200 µm. D. *Haliotis roberti*. Scale bar = 100 µm.

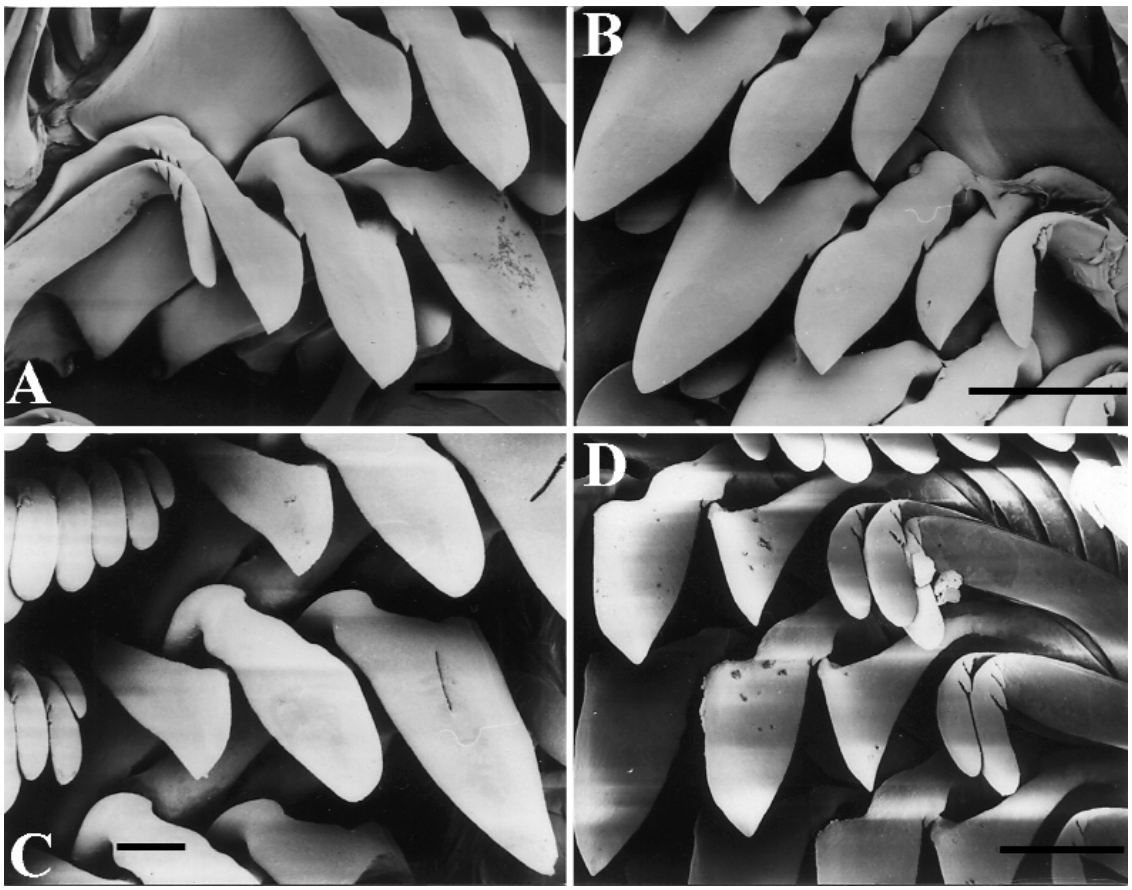


Figure 5-7. Illustration of radular character states for lateral teeth 3-5. Note denticles on outer edge of teeth in A, B, and D, in contrast to C, for which the outer edge is smooth. A: *Haliotis pustulata*. B: *Haliotis tuberculata*. C: *Haliotis asinina*. D. *Haliotis varia*. Scale bar = 100 µm.

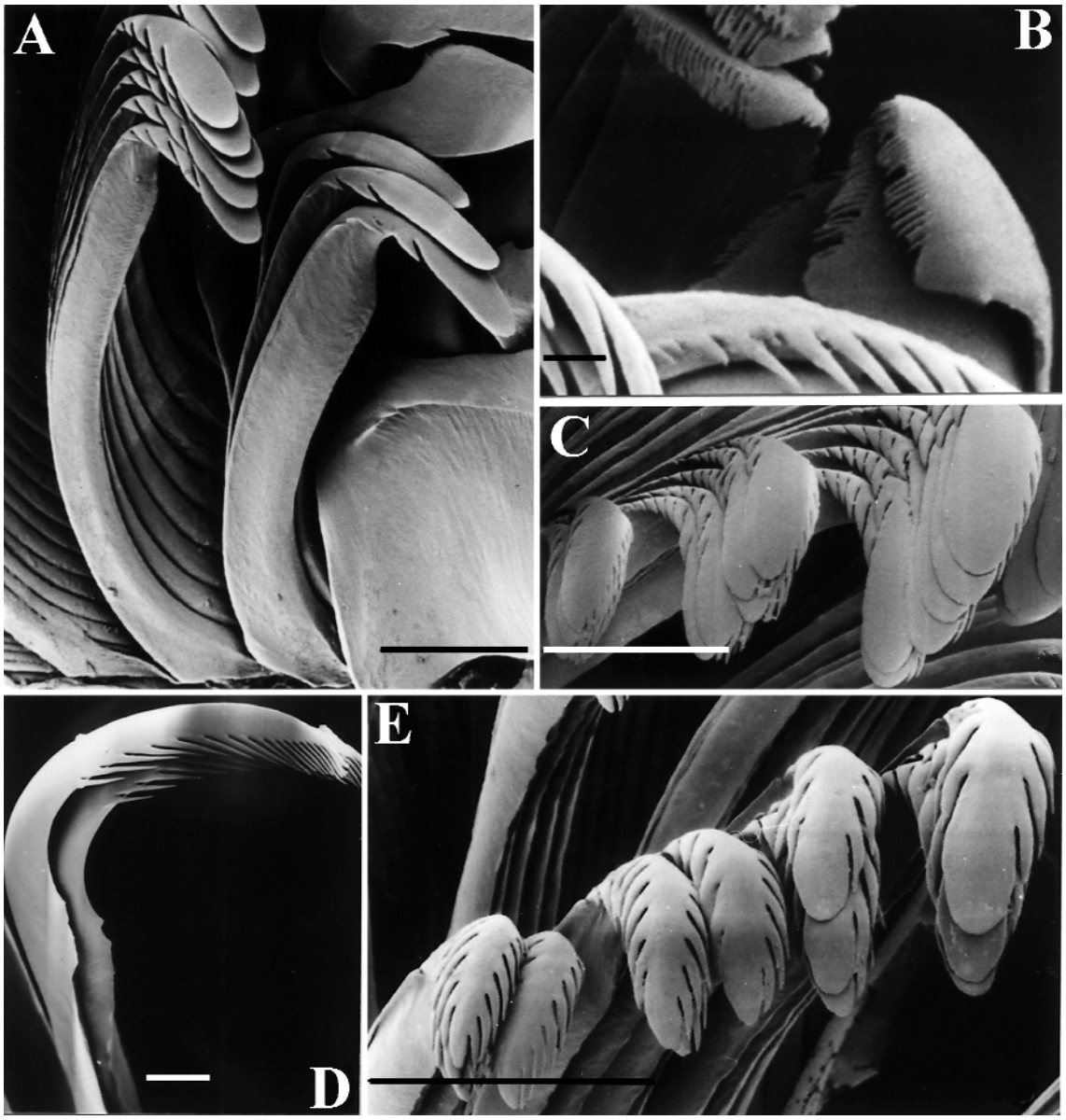


Figure 5-8. Illustration of radular character states for marginal teeth. The symmetrical condition (0) is shown in A. Condition (1) for character 7 is seen in E. Condition (1) for character 8 is shown in C. A: *Haliotis pustulata*. Scale bar = 100 µm. B. Outer marginals of *H. asinina*. Scale bar = 10 µm. C. Inner marginals of *H. asinina*. Scale bar = 100 µm. D. Lateral view of middle marginals of *H. tuberculata*. Scale bar = 10 µm. D. Inner marginals of *H. varia*. Scale bar = 100 µm.

2: Thickened cutting edge on rachidian tooth and lateral tooth 1 (Figure 5-6). Absent: 0; Present: 1.

3: Lateral tooth 1 with acute angle formed by cutting edge and primary ridge (Figure 5-6). Absent: 0; Present: 1.

4: Concave primary ridge on lateral tooth 1 (Figure 5-6). Absent: 0; Present: 1.

5: Size of primary ridge (Figure 5-6): ≤ 10% the size of the secondary ridge: 0; > 10% larger than secondary ridge: 1.

6: Denticles on outer margin of lateral teeth 3-5 (Figure 5-7). Absent: 0; present: 1.

7: Denticles on inner marginal teeth at least by one denticle asymmetrical (Figure 5-8). No: 0; yes: 1.

8: Denticles on inner marginal teeth at least by three denticles asymmetrical (Figure 5-8). No: 0; yes: 1.

9: Rachidian tooth broader than lateral tooth 1, and lateral tooth 1 approximately same size as lateral tooth 3 (Figure 5-6). No: 0; Yes: 1.

10: Rachidian tooth posteriolateral wings with knobs/joints (Figure 5-6)? Absent: 0; present: 1.

11: Lateral tooth 1 much narrower than rachidian tooth and lateral tooth 3 (Figure 5-6)? No: 0; Yes: 1.

THE EPIPODIUM

The epipodium is a fleshy girdle with various projections situated on the dorsal portion of the foot below the margin of the shell (Crofts, 1929; Cox, 1962). It is best developed in Haliotidae, but is also recognized in other vetigastropod taxa (Pleurotomariidae: Anseeuw & Goto, 1996; Scissurellidae: Haszprunar, 1988a, 1988c; Trochidae: Hickman & McLean, 1990). The epipodium is known to have a species-specific morphology

Pleurotomariidae	0	0	0	0	0	0	0	-	1	0	1	0
Scissurellidae (*Sciissurella*)	0	0	0	5	5	0	0	-	0	0	0	0
Lepetodrilidae (*Lepetodrilus*)	0	0	0	5	5	0	0	-	0	0	0	0
Fissurellidae (*Fissurella*)	1	0	1	0	0	0	0	-	1	0	1	0
Trochidae (*Tegula*)	0	0	0	5	5	0	0	-	1	0	0	0
Turbonidae (*Liotia*)	0	0	0	5	5	0	1	-	0	0	0	0
asinina	1	0	1	1	1	0	0	-	2	1	0	0
aurantium	2	0	1	1	1	0	0	-	0	0	1	0
australis	2	1	0	1	1	1	0	0	0	-	0	0
brazieri	0	0	1	1	1	0	0	-	1	0	1	0
clathrata	5	1	1	1	1	0	0	0	2	0	0	0
coccoradiata	0	0	1	1	1	0	0	-	0	1	1	0
corrugata	2	0	1	0	1	0	0	-	1	1	1	0
cracherodii	2	0	1	0	1	1	1	-	0	0	0	0
dalli	0	0	1	1	1	0	0	-	1	0	1	0
discus	2	0	1	0	1	1	0	-	1	0	1	0
dissona (dried body)	?	?	?	?	?	?	?	?	?	?	?	?
diversicolor	1	0	1	1	1	0	0	-	2	1	1	0
elegans	1	0	0	1	1	1	0	-	2	1	1	0
fatui	0	0	1	0	0	0	0	-	1	1	1	0
fulgens	2	0	1	0	1	0	0	-	2	0	0	0
gigantea	?	?	?	?	?	?	?	?	?	?	?	?
glabra	0	0	1	1	1	0	0	-	1	0	1	0
hargravesi	0	0	1	1	0	0	0	-	1	0	1	0
iris	2	0	1	0	1	0	0	-	0	0	0	1
jacnensis (dried body)	?	?	?	?	?	?	?	?	?	?	?	?
kamtschatkana	2	0	1	0	1	0	0	-	1	0	1	0
laevigata	0	0	1	1	1	0	0	-	2	0	0	0
madaka	2	0	1	0	1	0	0	-	0	0	1	0
mariae	0	1	0	0	0	1	0	0	0	-	0	0
marmorata	1	0	1	1	1	0	0	-	1	0	0	0
midae	2	0	1	1	0	0	0	-	2	0	1	0
ovina	0	0	1	1	1	0	0	-	0	1	0	0
parva	0	1	1	1	0	0	0	0	1	0	1	0
pourtalesii	2	0	1	0	1	0	0	-	1	0	1	0
planata	0	0	1	1	1	0	1	-	1	1	0	0
pustulata	2	1	0	1	1	1	0	0	0	0	0	0
queketti	2	0	1	1	1	0	1	-	1	0	0	0
roberti	2	0	1	1	1	0	0	-	0	0	1	0
roei	2	0	0	1	1	1	0	-	1	0	0	0
rubiginosa	0	0	1	1	0	0	0	-	1	1	1	0
rubra	2	0	1	1	1	1	0	-	0	1	1	0
rufescens	2	0	1	0	1	0	1	-	1	0	0	0
rugosa	2	1	0	1	1	1	0	0	0	0	0	0
scalaris	1	0	1	1	1	0	0	-	1	1	0	0
semiplicata	0	0	1	1	1	0	0	-	1	1	1	0
spadicea	1	0	1	1	0	0	1	-	1	0	0	0
speciosa	2	1	1	1	1	0	0	0	2	0	1	0
squamata	1	0	1	1	1	0	0	-	2	0	1	0
stomatiaeformis	1	0	0	1	0	0	0	-	1	0	1	1
tuberculata tuberculata	1	1	1	1	1	1	0	1	1	1	1	1
tuberculata coccinea	1	1	1	1	1	1	0	1	1	0	1	1
unilateralis	0	1	1	1	1	0	0	1	1	0	0	1
varia	0	0	1	1	1	0	0	-	1	1	0	0
virginea	1	0	1	0	1	1	0	-	0	0	0	0
walallensis	2	0	1	1	1	1	0	-	1	0	0	0

Table 5-3: Character states for epipodial characters. For description of character states see text. 5 = (0, 1). - = inapplicable.

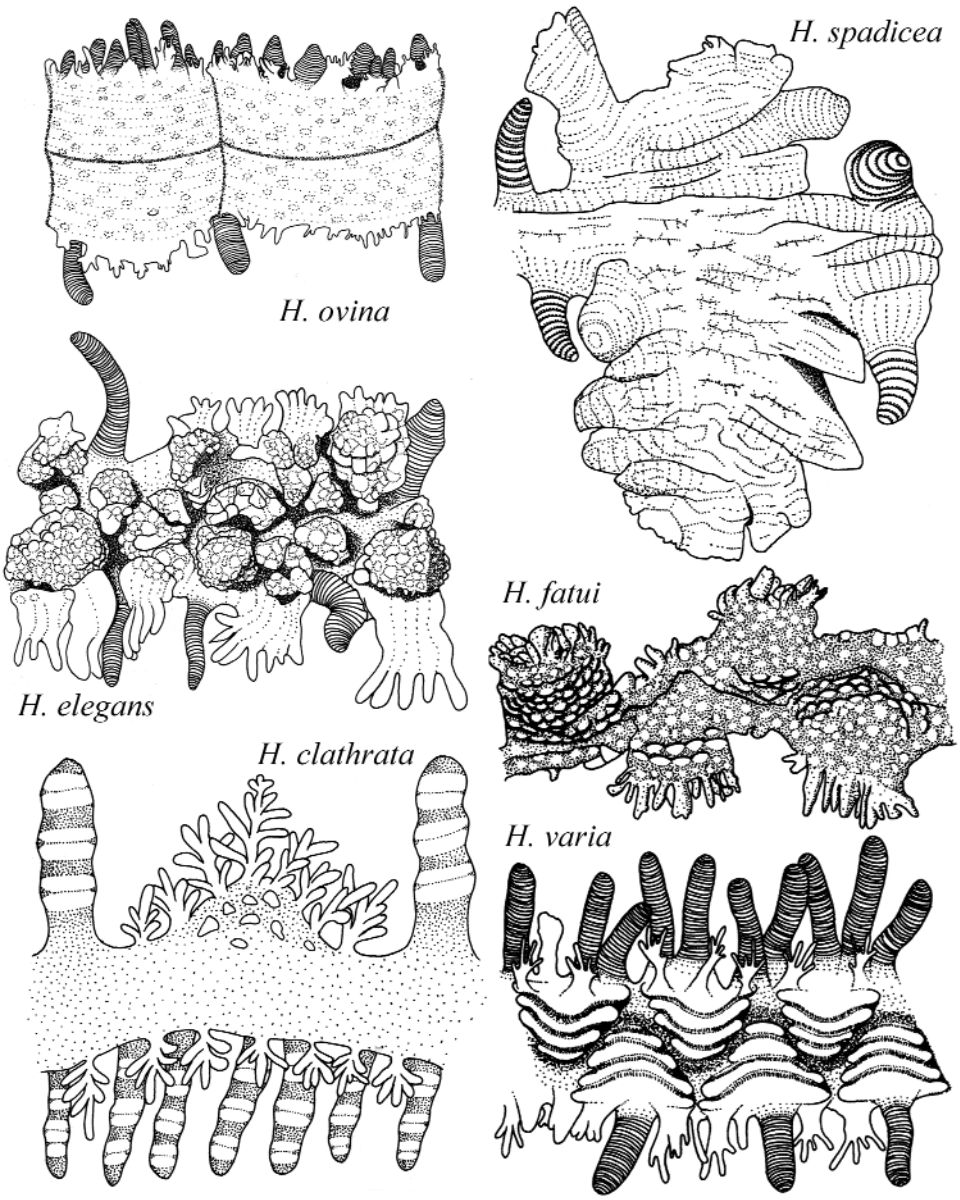


Figure 5-9. Sample epipodia of *Haliotis* spp. Semischematic reconstructions from camera lucida drawings. *Haliotis ovina* (LACM 85-3). *Haliotis spadicea* (DLG). *Haliotis elegans* (USNM 360940). *Haliotis fatui* (USNM 486708); from Geiger (1999). *Haliotis clathrata* (DLG no #); from Stewart & Geiger (1999). *Haliotis varia* (DLG no #); from Geiger (1999). Dorsal and ventral tentacles in *H. spadicea*. Undulating midepipodial fold in *H. fatui* and *H. varia*. Cauliflower-shaped projections in *H. elegans*. Palm-shaped projections in *H. clathrata*.

(Owen *et al*., 1971), but has received little attention since then with some notable exceptions (Geiger, 1996, 1999a; Simone, 1998; Stewart & Geiger, 1999). The data matrix for the epipodial characters is given in Table 5-3 and Figure 5-9 shows some sample epipodia with the various character states.

Outgroup comparison

The epipodium of Haliotidae is particularly well developed and can be termed hypertrophied as compared to those of the remaining Vetigastropoda. The comparison to other vetigastropods proved to be difficult, due to the extent to which the haliotid epipodium is hypertrophied. The postulation of shared similarities for character states could not be satisfactorily resolved. Comparison between non-haliotid vetigastropods proved to be more conducive than the comparison of haliotid and other vetigastropod epipodia. As a consequence, the establishment of character-state polarity through outgroup comparison only succeeded at a rudimentary level.

Characters and character states

1: Symmetry of epipodium. Symmetrical: 0; asymmetrical, dorsal side more developed: 1; asymmetrical, ventral side more developed: 2.

2: Midepipodial tentacles. Absent: 0; present: 1.

3: Midepipodial fold. Present: 0; absent: 1.

4: Dorsal fringe tentacles. Absent: 0; present: 1.

5: Ventral fringe tentacles. Absent: 0; present: 1.

6: Cauliform projections (e. g., *H. elegans*). Absent: 0; present: 1.

7: Bare lateral facing surface (face)? No: 0; Yes: 1.

8: Arrangement of midepipodial tentacles. Scattered: 0; in line: 1.

9: Palmate structures? Absent 0; present without secondary palms on surface: 1; present with secondary palms on surface: 2.

10: Epipodial fold. Straight: 0; undulating: 1.

11: Tuberculate cones on face. Absent: 0; present: 1.

12: Sheeth surrounding tentacles. Absent: 0; present: 1.

Hypobranchial gland

The hypobranchial gland is a highly folded tissue located in the roof of the mantle cavity. The character states are given in Table 5-4. A sample illustration of a hypobranchial gland is shown in Figure 5-10. The glandular tissue shows a few modifications between species of abalone, which were coded as follows:

1: A few, irregular side branches from main lamellae? No: 0; At the tip: 1; Only in posterior part: 2.

2: Does posterior part contain lamellae, which intersect anterior lamellae? No :0; Yes: 1.

3: Length of hypobranchial gland. Less than six times as long as wide: 0; more than six times as long as wide: 1.

4: Number of lamellae on hypobranchial gland. Less than 20: 0; More than 20: 1.

5: At tip many regularly spaced, secondary lamellae? Absent: 0; present: 1.

6: Height of lamellae on right side of hypobranchial gland (*in situ*). As high as lamellae on left side: 0; higher than on left side: 1.

Pleurotomariidae	0	1	0	1	0	0	*laevigata*	0	0	1	1	0	0
Scissurellidae (*Scissurella*)	?	?	?	?	?	?	*madaka*	?	?	?	?	?	?
Lepetodrilidae (*Lepetodrilus*)	0	0	0	1	0	0	*mariae*	2	1	0	1	0	1
Fissurellidae (*Fissurella*)	-	-	-	-	-	-	*marmorata*	2	1	0	1	0	1
Trochidae (*Tegula*)	-	-	-	-	-	-	*midae*	1	1	1	1	0	0
Turbonidae (*Turbo*)	0	0	?	?	0	0	*ovina*	2	1	0	5	0	1
							parva	0	1	0	1	0	1
asinina	1	1	1	1	1	0	*planata*	2	1	1	1	0	1
assimilis	0	1	0	5	1	0	*pourtalesii*	0	0	0	1	0	1
aurantium	0	0	0	0	0	0	*pustulata*	1	1	1	1	0	0
australis	0	0	1	1	0	1	*queketti*	0	1	0	1	0	0
brazieri	0	0	0	1	0	0	*roberti*	0	0	0	1	0	0
clathrata	0	0	0	1	0	1	*roei*	0	1	0	1	0	0
coccoradiata	1	1	0	1	0	1	*rubiginosa*	0	0	0	1	0	0
corrugata	0	1	0	1	1	0	*rubra*	2	1	0	1	0	0
cracherodii	0	5	0	0	1	0	*rufescens*	0	1	0	5	1	0
dalli (missing)	?	?	?	?	?	?	*rugosa*	0	0	0	1	0	1
discus	0	0	1	1	0	0	*scalaris*	2	1	0	1	0	0
dissona (dried body)	?	?	?	?	?	?	*semiplicata*	2	0	1	1	1	1
diversicolor	0	1	1	1	0	0	*spadicea*	0	1	0	0	5	0
elegans	0	0	1	1	0	1	*speciosa*	?	?	?	?	?	?
fatui	2	1	1	1	0	1	*squamata*	0	1	1	1	0	1
fulgens	0	1	0	0	1	0	*stomatiaeformis*	0	0	1	1	0	1
gigantea	1	1	1	1	1	0	*tuberculata tuberculata*	1	0	1	1	0	1
glabra	1	1	0	1	0	0	*tuberculata coccinea*	0	0	0	1	0	0
hargravesi	0	0	0	1	0	0	*unilateralis*	?	?	?	?	?	?
iris	1	1	1	1	1	0	*varia*	2	1	0	0	1	0
jacnensis (dried body)	?	?	?	?	?	?	*virginea*	1	0	0	1	0	1
kamtschatkana	0	1	0	5	1	0	*walallensis*	1	0	0	1	0	1

Table 5-4. Character states for hypobranchial gland characters. For description of character states see text. 5 = (1, 0). - = inaplicable.

Figure 5-10. Example of a hypobranchial gland of *H. fatui*. Holotype USNM 486708. r = rectum. Scale bar = 5 mm. From Geiger (1999).

Potential characters not used

I have not used some characters and observable features of abalone. The underlying reason for the exclusion of this data is uncertain belief formation, i.e., I strongly doubt *prima facie* that the observable shared similarities can be explained by common ancestry. Such considerations are somewhat subjective as pointed out in Chapter 3, hence, I will offer arguments to support my position.

Shell morphology

Abalone shells are extremely plastic in virtually every aspect of their morphology, which includes both qualitative and quantitative features. Striking examples for qualitative characters include the following: lamellosity of the shell (Figures 4-4, 4-5); height of the spire (Figures 4-12, 4-13); rotundity of the shell, which is under environmental influence (cf. Ino, 1952) and changes during ontogeny (Stewart & Geiger, 1999); shell coloration, although useful in other groups of organisms (e.g., Westerneat, 1993; Swenson & Bremer, 1997), is extremely variable in abalone as discussed in Chapter 4 (= Geiger, in press). The most prominent quantitative character is the number of open holes in abalone. I have discussed the variability of this character already in Chapter 1 (= Geiger, 1998a) and Chapter 2 (= Geiger & Groves, 1999).

Shell mineralogy

Shell mineralogy was shown to differ between abalone species in three distinct patterns (Dauphin *et al*., 1989; Dauphin & Denis, 1995). These results were already discussed in Chapter 2. Additional problems with mineralogical characters arise when evaluating studies regarding the mineralogy of Recent organisms into account. The connection between sea water temperature and mineralogical composition of shells, com-

bined with the observation that the supposedly aragonite-specific Feigl stain also binds to high magnesium calcite (C. Hedegaard, pers. comm.), are particularly significant. The influence of sea temperature on the magnesium content in calcium has been demonstrated in various invertebrates (e.g., Weinbauer & Velirimov, 1995; Hastings *et al.*, 1998). As Haliotidae occur over a substantial temperature range from below 10°C in the case of the pinto abalone (*H. kamtschatkana*) to greater than 20°C for many tropical species, it is highly questionable that the shell mineralogy will harbor any phylogenetic signal.

ANALYSES WITH MORPHOLOGICAL CHARACTERS

An outgroup-rooted analysis of all 32 ingroup taxa for which at least some molecular data was available, with morphological data and the six outgroups added, was performed. The heuristic search was constrained so that the ingroup was monophyletic [constraints outgroups=((1-6)(7-58))], because the monophyly of the family Haliotidae is not in question (Figure 5-11). The analysis found 680 MPRs of length 1290 (CI = 0.519, RI = 0.699, RC = 0.363). The strict consensus tree showed hardly any resolution. The only topologies recovered were: (Scissurellidae (Lepetodrilidae, Turbinidae)); ((*pustulata*, *rugosa*) *virginea*) *asinina*) (*tuberculata tuberculata*, *tuberculata coccinea*) *australis*); ((*laevigata*, *scalaris*) *roei*); and (*pourtalesii*, *walallensis*). I suspect that the large amount of missing data in the six outgroups was largely responsible for the many MPRs found. The low information content of the morphological data is also evident from the relationships within the outgroup taxa. Although the between-family relationships within Vetigastropoda is currently uncertain (cf. Baten, 1975; Haszprunar, 1988b; Tillier *et al.*, 1994; Hickman, 1996; Harasewych *et al*., 1997), the monophyly of the Trochoidea (= Trochidae and Turbinidae) is undoubted (Hickman & McLean, 1990;

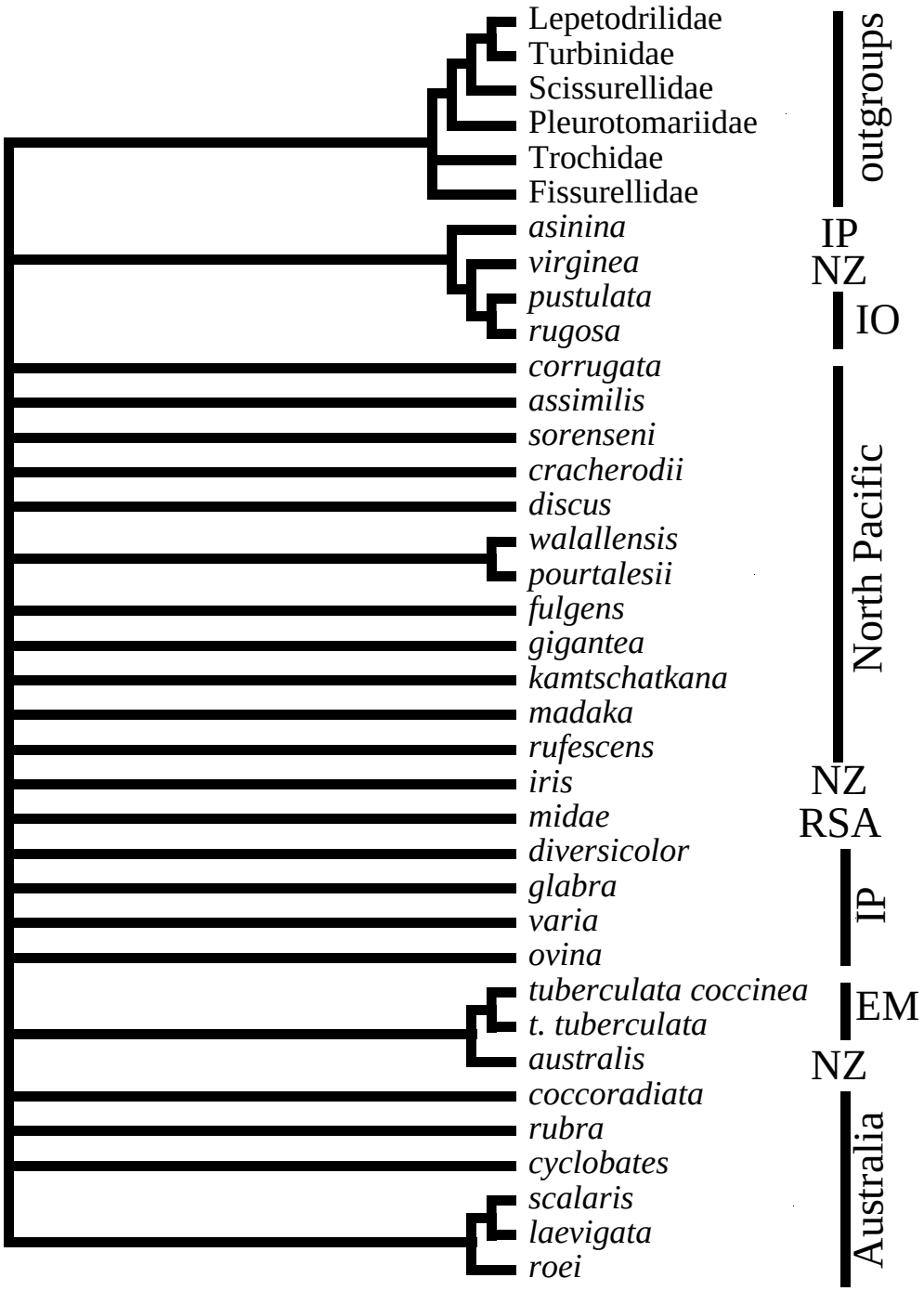


Figure 5-11. Strict consensus tree of 680 MPRs from combined analysis of allozyme (Brown, 1993), lysin (Lee & Vacquier, 1995), 16S mtDNA (Wray, unpubl.), and morphology (this study), with six vetigastropod outgroups. The low resolution is due to excessive missing data.

Hickman, 1996). Figure 5-11, however, showed these two families separated by three nodes.

An unrooted analysis for the same 32 ingroup taxa as above was carried out with all available data, but without the outgroups. This analysis was performed in order to assess whether the large amount of missing data in the outgroup taxa was responsible for the low resolution in Figure 5-11. The analysis found 38 MPRs of length 1256 (CI = 0.525, RI = 0.703, RC = 0.369. Figure 5-12). This sharp drop in number of equally parsimonious trees clearly demonstrates that the amount of missing data in the outgroups causes the parsimony algorithm to find many equally parsimonious topologies. The polytmy for the north-Pacific group, which was resolved in the previous analyses (see Figure5-4) without the morphological data can not be explained by missing data alone. It seems that the morphological data actually reduced the resolution in that group. Note that the between-outgroups relationships are peculiar in that the two top shell families Trochidae and Trubinidae, which are not doubted to be sistergroups, are separated by three nodes. This fact points to the questionable nature of the morphological data set, although one can also argue that these morphological characters are not useful at the betweeen families level, but may be well suitable to elucidate within-family relationships.

An analysis without the allozyme data was performed to determine the relative contribution to the number of MPRs of missing data compared to the low information content in the morphological data. The allozyme data-set is the one with the least taxa, hence, with most missing data. This analysis included 30 taxa. *Haliotis coccoradiata* and *H. virginea* were omitted, because the only molecular data available were allozyme frequencies. The analysis of lysin, 16S mtDNA, and the morphological data produced 106 MPRs of 1123 steps (CI = 0.537, RI = 0.719, RC = 0.386. Figure 5-13). The number of MPRs was approximately 2.5 times as large in the analysis made without the

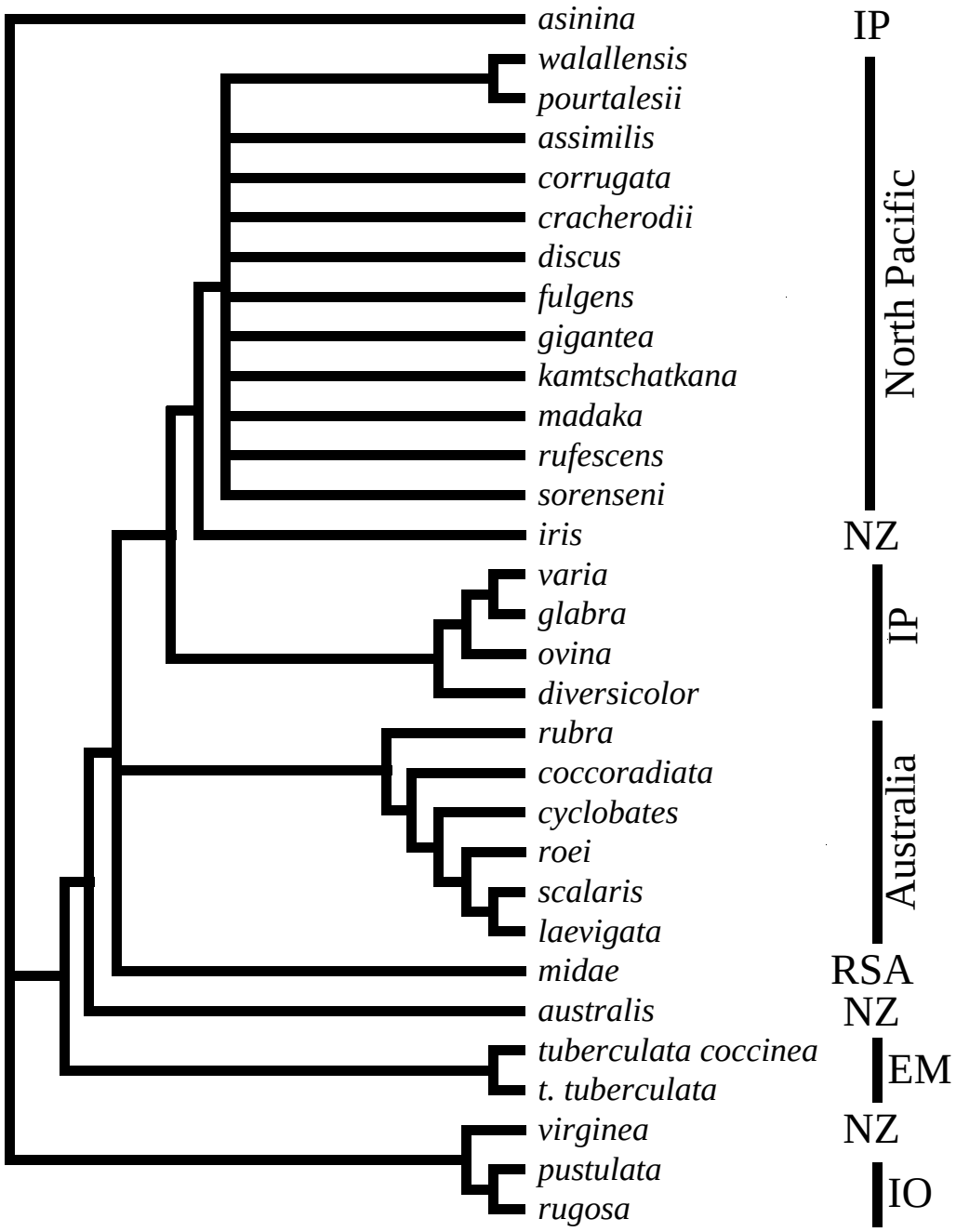


Figure 5-12: Strict consensus tree of 32 ingroup taxa from combined analysis of allozyme (Brown, 1993), lysin (Lee & Vacquier, 1995), 16S mtDNA (Wray, unpubl.), and morphology (this study), without outgroups.

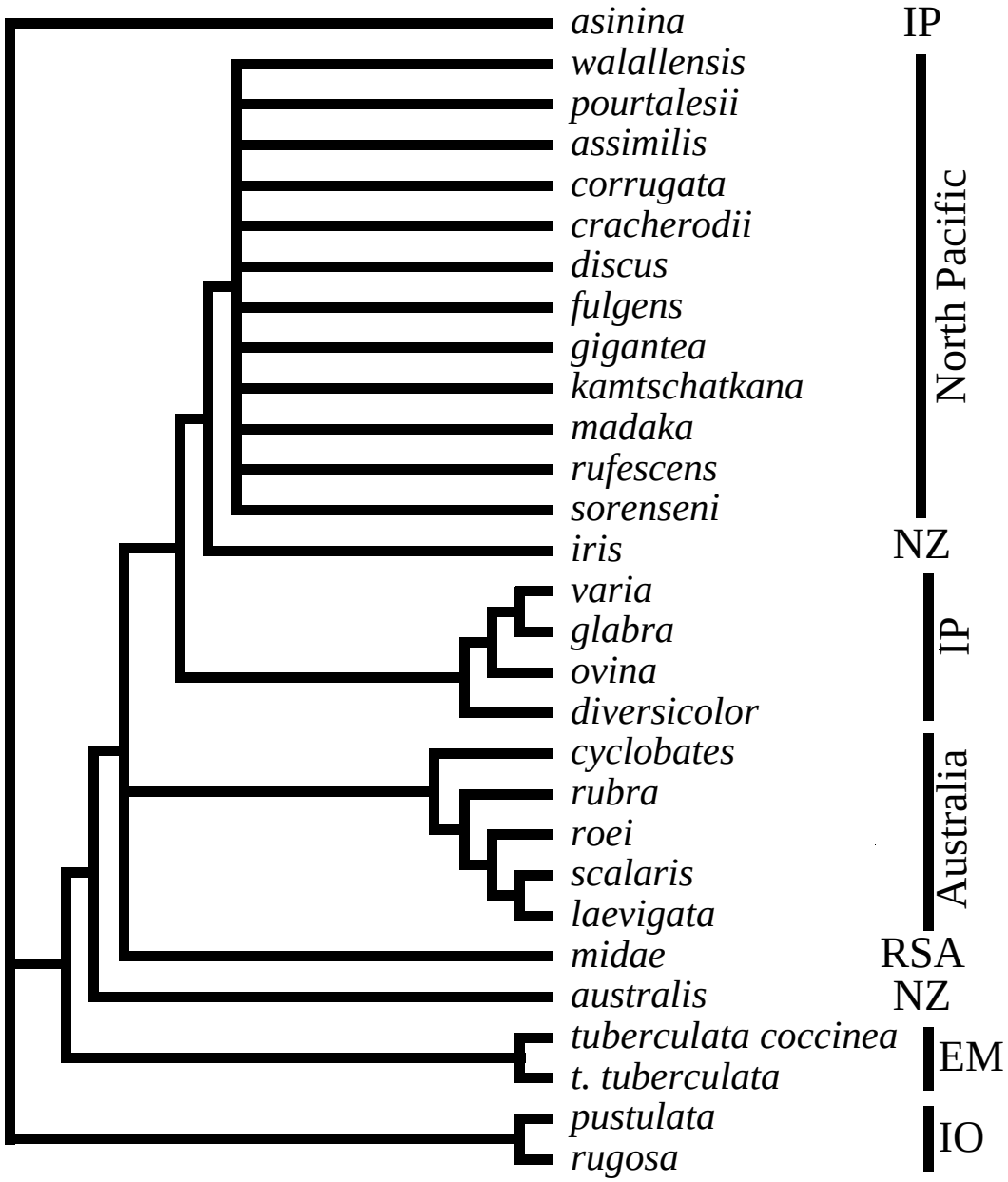


Figure 5-13. Strict consensus tree for combined analysis of lysin (Lee & Vacquier, 1995), 16S mtDNA (Wray, unpubl.), and morphology (this study),, without allozyme (Brown, 1993) data.

allozyme data. The only area of the tree that is affected is the north Pacific taxa. When including allozymes, the (*pourtalesii*, *walallensis*) clade is recovered, byt without allozymes these two taxa are found in the overall polytomy of the north Pacific taxa. Additionally a small rearrangement within the Australian endemic species is found, but in both cases the Australian endemic species are fully resolved. The consensus tree shows, that in some areas of the tree the morphological data seem to remove resolution due to character conflict. This is seen by the large polytomy of the north Pacific taxa.Iin other areas of the tree the morphological characters do not contradict the molecular data as shown by the fully resolved remainder of the tree.

In conclusion, the use of morphological data in this particular study highlights two problem areas. The first is that when a small morphological data-set is added to a larger data-set of other characters, the amount of missing data may prove to be problematic in a total evidence cladistic analysis. Second, morphological data, like any other data set, may be useful in one area of the tree, but may even obscure evolutionary relationships amongst taxa in other areas of the tree. The latter is clearly shown with the north Pacific taxa and the outgroups.

CLASSIFICATION

The classification of abalone has been unresolved. Fourteen genus-level taxa have been proposed, which were discussed in Chapter 1 (= Geiger, 1998a). Lee & Vacquier (1995) made some taxonomic recommendations, but these are fundamentally flawed in that the type species of *Haliotis* s. s., *H. asinina*, was not been included in their study. Accordingly, Geiger (1996, 1998a) recommended use of a single genus until a comprehensive phylogenetic hypothesis has been formed. The present study has advanced this goal it is appropriate to discuss the classification within Haliotidae. I will highlight

those findings that are consistent between the several different analyses and omit from the discussion all those potential groupings and taxa for which there is no sound basis for an in-depth appraisal.

Haliotis*, *sensus stricto

The type species *H. asinina* occupies an isolated position in the phylogenetic tree. It does not group with any other abalone species. There are two possible ways to deal with this conclusion. Either the genus *Haliotis* is monotypic, which is a little-favored option, particularly amongst phylogeneticists, because a monotypic genus is by definition monophyletic, i.e., it does not help in the understanding of shared similarities among species. The alternative is to treat *Haliotis* as the single genus for the family Haliotidae, i.e., to continue the current practice. To restrict the use of *Haliotis* to all the species that are not assigned to another genus-level taxon (see below) is not advocated, because it would make that taxon polyphyletic.

Nordotis

The most prominent clade contains the north Pacific species on both sides of the Pacific, i.e., the large Japanese species and all west American species, as well as the tropical American species on either side of Panama. For these species either the name *Nordotis* or *Notohaliotis* has been employed (e.g., Kira, 1962). *Notohaliotis* has as its type species the Australian *H. rubra* (see below), therefore, *Notohaliotis* can not be applied for the north Pacific clade. Two genus-level taxa have as type species members of the north Pacific clade: *Nordotis* Habe & Kosuge, 1964 (type = *H. gigantea* Gmelin, 1791), and *Usahaliotis* Habe & Kosuge, 1964 (type = *H. cracherodii* Leach, 1814).

Applying the first reviser's principle, I recommend the name *Nordotis* to be applied to this group.

Nordotis is characterized by one rather weak radular synapomorphy, the concavity in the primary ridge on the lateral tooth 1. This character is also found in in some other abalone taxa, but is a rather consistent feature of the north Pacific species. The epipodium is highly asymmetrical with the ventral side being more strongly developed than the dorsal side. In the UTR portion of the lysin gene all species contain two noticeable gapstretches with an interspersed variable region in positions 1 to 10, an insert sequence in positions 26 to 39, and a consistent gapstretch in positions 145-153 (see Table 4-3).

Notohaliotis

The second consistent group contains the endemic Australian species, to the exclusion of those warm water Australian species, which have a wide Indo-Pacific distribution, such as *H. clathrata*, *H. ovina*, and *H. varia*. A number of genus-level taxa could be applied to this group: *Exohaliotis* Cotton & Godfrey, 1933 (type: *H. cyclobates*), *Marinauris* Iredale, 1927 (type: *H. brazieri*), *Neohaliotis* Cotton & Godfrey, 1933 (type: *H. scalaris*), and *Notohaliotis* Cotton & Godfrey, 1933 (type: *H. rubra*). The most frequently used name is *Notohaliotis*. Note that *Padollus* has often been used for some of the Australian species, particularly *H. scalaris*. The type of *Padollus*, however, is the south African *H. parva*, hence *Padollus* can not be applied to the clade of Australian endemic species. No diagnostic synapomorphies can be given for this weakly supported clade. The species are grouped by rather diffuse characters, such as character state 'd' in position 75 of the lysin UTR.

Sanhaliotis

The last consistent group is the Indo-Pacific species. Membership in this clade is not as well defined as in the other clades but it seems to contain at least *H. diversicolor*, *H. glabra*, *H. ovina*, and *H. varia*. *Ovinotis* Cotton, 1943 (type: *H. ovina*) and *Sanhaliotis* Iredale, 1929 (type: *H. varia*) have been proposed as the generic name for species in this clade. *Sanhaliotis* has priority over *Ovinotis*. No diagnostic synapomorphies can be given for this clade. The species are grouped by rather diffuse characters. It seems that a strong posteriobasal projection on the rachidian tooth and a strongly angulated primary ridge on lateral tooth one are found in some of these taxa (Figure 5-6: B). These characters are also found in some other taxa such as *H. jacnensis*, which I suspect to belong in this clade, but the present data do not justify a firm conclusion.

Use of Genus-level taxa

I recommend the use of only the single genus-level taxon *Haliotis*, because the cladistic hypothesis is not firmly established and is based in some cases on too little data. Additionally, not all species can be assigned to any given genus-level taxon with any degree of certainty. Those who wish to emphasize particular groupings may use subgenera, which are not mandated for unambiguous identification of a species with a binomen. It also avoids the use of the polyphyletic taxon *Haliotis* s.s.

Conclusions

The conclusions of this dissertation are summarized as follows:

CHAPTER 1: RECENT TAXA

There are 56 valid species with 10 valid subspecies out of the over 200 species-level taxa proposed (Geiger, 1998a, 1998b, 1999; Geiger & Stewart, 1998; Stewart & Geiger, 1999). The species count has been reduced from the previously accepted 75 by 20 species. One additional species was described. This demonstrates that a critical evaluation of the ingroup taxa is a crucial first step in any phylogenetic analysis.

CHAPTER 2: FOSSIL TAXA

Abalone are rare in the fossil record. This scarcity is due to a combination of the rocky habitat and the nacreous shell, which makes abalone little likely to be preserved in the fossil record. The first records from the Upper Cretaceous (Maastrichian) does most likely not represent the root of the lineage. The ecology of abalone seems not to have changed over the documented time span. No on-shore/off-shore pattern of vertical distribution could be detected.

CHAPTER 3: DNA SEQUENCE ALIGNMENT

The critical evaluation of the coding related aspects in cladistic analysis has revealed inconsistency, particularly in how DNA sequences are used. As I used a combination of various data, such an evaluation is at the heart of a cladistic analysis. A strict comparison of the treatment of morphological data with that of sequence data, based on philosophy of science (particularly the nature of observation, abductive inference, explanation, classification theory, and relevance) has led to the postulation of a new methodology

for the treatment of sequence data in phylogenetic systematics (Geiger & Fitzhugh, in review; Geiger, 1999b). First, DNA sequence alignment must be limited to the smallest identifiable fragment, what I call *minimum fragment alignment* (relevance issues). Second, gaps must be treated as a fifth character state, because they are postulated during the classification act of observations. Third, any differential weighting scheme is without a basis, because observations are equivalent as such, and because such weighting schemes introduce an explanatory element into the observational phase. Fourth, differences in alignments are due to questionably aligned sequences. Recoding of such areas can only be accomplished with data reduction strategies such as exclusion and data contraction. Flexible coding strategies such as elision, polymorphic coding, missing data coding, presence/absence coding, and case sensitive coding, introduce internal conflicts and inconsistencies. For few but very dissimilar taxa, new coding strategies of *block coding* and *stretch coding* have been introduced, which have been applied in Chapters 4 and 5.

CHAPTER 4: BIOGEOGRAPHY

The distribution of all 56 species is documented, based on specimen data from museum collections (Geiger, in press). There are no abalone with a global distribution. The rather vague distributional indications for some species have been specified more exactly. The biogeographic analysis using Brooks parsimony revealed a general Indo-Pacific origin of the family. The precise origin can, however, not be pinpointed more precisely with the data at hand. Two of the biogeographical hypotheses (Tethys origin, Pacific Rim origin) are considered less likely, but the Indo-Pacific hypothesis is not without inconsistencies, most likely due to the potentially long evolutionary history of the family, reaching back at least to the middle Mesozoic.

CHAPTER 5: TOTAL EVIDENCE CLADISTIC ANALYSIS

All data analyses show the following patterns:

- a north Pacific group including the Japanese as well as the west American species;
- a group of species endemic to Australia;
- a group of widespread Indo-Pacific species;
- the three New Zealand species are dispersed over the entire tree.

Some molecular synapomorphies can be identified for some of the above groups, whereas the morphological characters do not show clear evolutionary patterns. The combination of molecular data sets contributes to the resolution of the combined topology, despite some missing data. The addition of morphological characters, although constituting a complete data-set, allows multiple character state optimizations leading to a large polytomy among the north Pacific taxa.

Outgroup comparison is currently of little value, because of the questionable similarity statements for the morphological characters, which are the only data available for the six vetigastropod outgroups. These similarity statements lead to rather questionable resolutions among the outgroups. The large amount of missing data due to no molecular data being available for the outgroups accounts for the nearly unresolved strict consensus tree.

Only four of the 17 proposed genus-level taxa may be used at this time. The only genus *Haliotis*, *s. l.*, for all abalone is suggested. Additionally, four subgenera may be used: *Haliotis*, *s. s*., for its Indo-Pacific type species *H. asinina*; *Nordotis*, for the north Pacific species; *Notohaliotis*, for the species endemic to Australia; and *Sanhaliotis*, for some of the wide-spread Indo-Pacific taxa. All other genus-level taxa do not warrant recognition at this time.