PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Comparative Variation of Morphological and Molecular Evolution through Geologic Time: 28S Ribosomal RNA Versus Morphology in Echinoids

Andrew B. Smith, B. Lafay and R. Christen

Phil. Trans. R. Soc. Lond. B 1992 **338**, 365-382
doi: 10.1098/rstb.1992.0155

Comparative variation of morphological and molecular evolution through geologic time: 28*S* ribosomal RNA versus morphology in echinoids

ANDREW B. SMITH[1], B. LAFAY[2] AND R. CHRISTEN[2]

[1]*Department of Palaeontology, The Natural History Museum, Cromwell Road, London SW7 5BD, U.K.*
[2]*U.R.A. 671 du CNRS & Université Paris VI Station Zoologique, Observatoire Océanologique, F06230 Villefranche sur Mer, France*

SUMMARY

The comparatively good fossil record of post-Palaeozoic echinoids allows rates of morphological change to be estimated over the past 260 million years and compared with rates of molecular evolution. Parsimony analysis of morphological data, based predominantly on skeletal characteristics, and parsimony, distance and maximum likelihood analyses of molecular data, from the first 380 bases from the 5′ end of the 28*S* rRNA molecule, for 10 species of echinoid produce congruent phylogenies. The molecular sequence chosen is demonstrably far from saturation and sister groups have divergence times ranging from about 15 to 260 Ma. Parsimony analysis allows the great majority of molecular and morphological apomorphies to be placed in one of 18 independent geological time intervals, providing a direct measure of rates of evolution for periods in the geological past.

Because most molecular fixed point mutations in our sequences cannot be polarized unambiguously by outgroup comparison (making the outgroup states effectively random), distance and parsimony analyses both tend spuriously to root the echinoid tree on the longest internal branch. A topology identical to that derived from morphological data is, however, obtained using Maximum Likelihood and also parsimony analysis where outgroup rooting is restricted to more conserved regions. This is taken as the correct topology for assessing rates of evolution.

Overall, both morphological and molecular changes show a moderately strong correlation with time elapsed, but a weaker correlation with one another. Statistically significant differences in evolutionary rate are found between some, but not all, pair-wise comparisons of sister lineages for both molecular and morphological data. The molecular clock rate for echinaceans is three times faster than that for cidaroids and irregular echinoids. Spearman's rank correlation test, which requires only relative magnitude of changes to be known, suggests that morphological change has a slightly better correlation with time than does molecular change, averaged over all ten species. However, when just echinaceans are considered an extremely good correlation is found between the number of molecular changes and time elapsed, whereas morphological change remains poorly correlated. Thus, molecular rates approximate to a clock-like model within restricted echinoid clades, but vary significantly between clades. Averaging results over all echinoids produces a correlation that is no better than the correlation between morphological change and time elapsed.

1. INTRODUCTION

Ever since Zuckerkandl & Pauling (1962) suggested that molecules can be used as 'documents of evolutionary history', differences in protein or nucleic acid sequences have become increasingly utilized for reconstructing phylogenies of organisms. Such studies cover diverse time-spans, ranging from the study of prokaryote relationships among groups which diverged over 2 billion years ago (Lake 1989, Woese 1989), to the divergence of humans from other primates, less than 10 million years ago (Gonzalez *et al.* 1990, Koop *et al.* 1989, Holmquist *et al.* 1988) and the emergence of human Influenza A virus some 150 years ago (Gammelin *et al.* 1990). Despite the popularity of this approach there has been continuing debate about variation in rates of molecular evolution and the causes of such variation (Cockburn *et al.* 1990, Krajewski 1990, Springer & Kirsch 1989, Koop *et al.* 1989, Li & Tanimura 1987, Gillespie 1986, Britten 1986). There is also debate as to the best method for deriving phylogenies from molecular data (see review by Swofford & Olsen (1990)).

Ribosomal RNA (rRNA) has in recent years been used for establishing phylogenetic relationships at various distances (e.g. Hedges *et al.* 1990, Gonzalez *et al.* 1990, Ratto & Christen 1990, Christen *et al.* 1991, Larson & Wilson 1989, Hillis & Dixon 1989, 1991,

Phil. Trans. R. Soc. Lond. B (1992) **338,** 365–382
Printed in Great Britain

Ghiselin 1989, Lake 1989, Woese 1989, Raff *et al.* 1988, to cite but a few). Although its precise mechanism of evolution remains poorly understood, rRNA is considered to be largely independent of morphological selection pressures and may be expected to conform to a clock-like rate of evolution. However, the secondary structure of rRNA defines regions which are under greater or lesser selective constraint and which therefore vary in the rate at which point mutations become fixed. Additionally, the rate of rRNA evolution in homologous domains clearly varies between major taxonomic groups, since 'fast-clock' and 'slow-clock' branches can be identified (Gonzalez *et al.* 1990) and may prove a significant problem for phylogenetic analysis (Felsenstein 1978, Lake 1989).

There have been few comparative studies combining morphology, molecular sequences and the fossil record, notable exceptions being the study based on palms recently published by Wilson *et al.* (1990) and the study of evolutionary rates in hominoid evolution by Koop *et al.* (1989). Here we investigate the molecular and morphological evolution of echinoids (sea urchins), an invertebrate group with a relatively good fossil record over the past 250 million years, using partial sequences from the 28*S* rRNA molecule.

Echinoids are chosen because they have a complex skeleton that provides a large number of characters for cladistic analysis. As a predominantly shallow water marine group they also have a reasonably good fossil record. The taxonomy of Recent species is almost entirely based on skeletal characters thus there is usually little difficulty in integrating fossil and Recent taxa into a unified classification. This means that fossil species can generally be placed into their appropriate stem groups without ambiguity. Divergence times, although impossible to measure precisely, can be estimated from the first occurrence of one or other derived sister taxa and the stratigraphical distribution of advanced stem group members that can be treated as putative ancestors.

For molecular data we have selected the 5′ end of the 28*S* rRNA molecule, because this region appears to have a rate of evolution appropriate for divergences over a time span of 250–50 million years (Michot & Bachellerie 1987).

Table 1. *Species analysed in this study (* excluded from analysis of evolutionary rates)*

- Echinoids (ingroup)
- Subclass Cidaroidea
  - *Cidaris cidaris* Linnaeus
- Subclass Euechinoidea
  - Cohort Irregularia (irregular echinoids)
    - Order Spatangoida
      - *Echinocardium cordatum* (Pennant)
      - *Spatangus purpureus* Müller
      - *Brissus unicolor* (Leske)
    - Order Clypeasteroida
      - **Echinocyamus pusillus* (Müller)
  - Cohort Echinacea
    - Superorder Stirodonta
        - *Arbacia lixula* (Linnaeus)
    - Superorder Camarodonta
      - Order Echinoida
        - Family Echinidae
          - *Echinus esculentus* Linnaeus
          - *Psammechinus miliaris* Müller
          - *Paracentrotus lividus* (Lamarck)
        - Family Echinometridae
          - *Lytechinus variegatus* (Leske)
          - *Sphaerechinus granularis* (Lamarck)

- Echinoderms from other classes (outgroup)
  - Crinoidea
    - *Antedon bifida* (Pennant)
  - Asteroidea
    - *Asterias rubens* (Linnaeus)
    - *Asterina gibbosa* (Pennant)
    - *Coscinasterias tenuispina* (Lamarck)
    - *Echiaster sepositum* Gray
    - *Henricia sanguinolenta* (Müller)
    - *Marthasterias glacialis* (Linnaeus)
    - *Solaster papposa* (Linck)
  - Ophiuroidea
    - *Amphipholis squamata* (Chiaje)
    - *Ophiothrix fragilis* Ablidgaard
  - Holothuria
    - *Holothuria forskali* Chiaje
    - *Leptosynapta inhaerens* (Müller)

2. MATERIALS AND METHODS

Eleven species of echinoid (table 1) were collected and identified by the staff of the marine stations at Villefranche sur Mer and Roscoff, France, and the Oceanographic Museum at Monaco. As an outgroup, 12 species belonging to the four other classes of echinoderm were also obtained from the same sources (table 1). Partial sequences of the 28*S* rRNA molecule were obtained for all 23 species, comprising the first 380 sites from the 5′ end following the technique described by Qu *et al.* (1983).

RNA was isolated from about 1 g of fresh tissue from each animal, homogenized with a polytron homogenizer in 5 ml of a solution of guanidinium thyocianate (4 M) Tris HCl (50 mM, pH 7.6). EDTA (4 mM), *N*-lauryl-sarkozyl (2%), 2-mercapthoethanol (1%). Total RNA was then separated from protein by phenol extraction, repeated three times and followed by two chloroform washes. Total RNA was then ethanol precipitated, suspended in sterile distilled water for measurement at 260/280 mm, reprecipitated with sodium acetate and ethanol, and finally resuspended in sterile distilled water at 2 μg μl^{-1}. Quality of the RNA extract was examined by ethidium bromide staining of 1% agarose gels.

RNA sequencing was carried out using reverse transcriptase elongation of ^{32}P 5′ end labelled synthetic DNA primers, in the presence of chain terminators. Three synthetic primers complementary to conserved domains located in the 5′ end of the 28*S* rRNA were used, namely 84–106, 278–302 and 382–404. These and all subsequent positions mentioned are numbered by reference to the mouse sequence (Baroin *et al.* 1988).

We initially included all 11 echinoid species in our analysis. However, the position of the irregular clypeasteroid *Echinocyamus* turned out to be highly unexpected, it being grouped not with other irregular echinoids, but as sister to the regular stirodont *Arbacia*. This appears to be a relatively robust pairing, supported by at least five uncontested synapomorphies. However, *Echinocyamus* and *Arbacia* represent the two longest branches in the tree and we suspect that the pairing may be an artifact caused by this, because Felsenstein (1978) and more recently Lake (1989) have demonstrated that taxa with long branch lengths tend to pair together for spurious reasons under parsimony or distance analysis. Because we are primarily interested in exploring rates of evolution in molecular and morphological trees that are congruent, for the moment we exclude *Echinocyamus* from our analyses, pending its re-sequencing and the addition of data from other clypeasteroids and arbaciids.

Sequences were aligned by eye, starting with the most conserved regions and progressively adding more divergent regions. Alignment between the echinoid species posed little problem. However, five sites (249–253, Appendix 2) were ambiguous with a large number of deletions and more than one equally likely alignment possible. Analyses were run both with and without these five variant positions. The first three sites and all positions beyond position 387 were excluded from this analysis, because of missing data within certain taxa. In total, 53 sites were variable among the ten echinoid sequences out of a total of 379 sites (14%), common deletions to echinoids having been removed (Appendix 2).

Unrooted trees were obtained using the parsimony computer program PAUP (Swofford 1990), using the 'exhaustive' or 'branch and bound' options to ensure the most parsimonious results were found. Bootstrapping, using 1000 replicates, was used to test the robustness of the results. Larger data matrices were also analysed using the computer parsimony program 'Hennig 86' (Farris 1988) and produced similar results. Distance matrix trees were calculated using the 'neighbour joining method' program of Saitou & Nei (1987). Finally the Maximum Likelihood program (DNAML) of Felsenstein (1989) was also used to derive trees from the molecular data.

A character matrix for morphological data was drawn up for the 10 remaining species using Mortensen (1928–1951), Hyman (1955), Smith 1981, 1984, 1988) and Jensen (1981) as the principal sources for characters (Appendix 1). There is undoubtedly a bias towards skeletal characters compared with soft tissue characters and all larval characters (even skeletal characters) were excluded. The choice of characters included is obviously much more subjective than when dealing with molecular sequences, but we have tried to be as comprehensive as possible, utilizing all the characters that previous workers have employed. It is certain that with further work or a different perspective, more characters could be included, particularly species-level autapomorphies, but we assume that our sample approximates a reasonably unbiased subset of all possible morphological characters.

The analysis is based on 81 characters listed in Appendix 1. Characters were treated as unordered and the cladogram rooted by reference to an outgroup composed of Palaeozoic echinoids, with particular emphasis on *Archaeocidaris*, as the closest stem group plesion.

3. RESULTS

(*a*) *Phylogenetic relationships derived from morphological data*

Parsimony analysis of morphological data produced two equally parsimonious solutions, length 98 steps, Consistency Index (CI) 0.92, Retention Index (RI) 0.89. These can be reduced to a single topology by treating the three irregular echinoids as a trichotomy (figure 1). One thousand bootstrap replications gave high levels of support for a number of branches but 50% or less support for the *Echinocardium*/*Spatangus* and the *Psammechinus*/*Paracentrotus* pairings. There are 18 trees of length 99 and 45 of length 100. Three trees out of the 18 within one step have irregular echinoids as sister group to all regulars (including *Cidaris*), whereas others either make the three Echinidae paraphyletic or place *Arbacia* as sister taxon to *Lytechinus* plus *Sphaerechinus*.

Apomorphies were allocated to branches by the computer program but were later checked manually, which led to some minor adjustments.

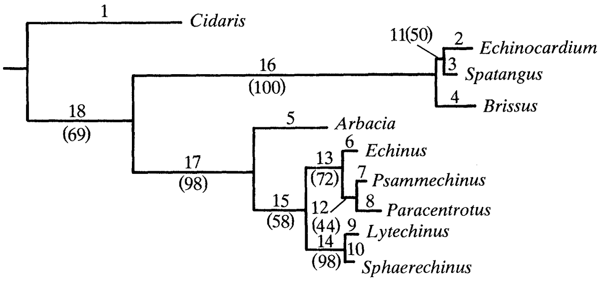

Figure 1. Relationships derived from parsimony analysis of the morphological data (Appendix 1). This is one of two equally parsimonious cladograms, 98 steps long with a consistency index of 0.92, and retention index of 0.89. The alternative tree is identical except that *Echinocardium*, *Spatangus* and *Brissus* form an unresolved trichotomy. Apomorphies supporting branches 1–18 are as follows (numbers refer to characters listed in Appendix 1): **branch 1**, 8b, 13a, 36b, 39b, 40b, 54b, 55b, 66b. **2**, 26b, 29b, 32c, 48b, 81b. **3**, 32b, 45b. **4**, 16b, 28b, 35b, 50b. **5,** 12b, 18b, 20b, 24c, 40b, 53a, 67b, 72b, 73b. **6**, 46b. **7,** –. **8**, 19b, 79a. **9**, 19b. **10**, –. **11**, 31a. **12**, 59a. **13**, 11b, 57b, 60a. **14**, 4b, 15b, 58b, 60b, 64b, 72b. **15**, 18c, 24b, 55b, 56a, 65b, 73c, 79b. **16**, 1b, 2b, 3a, 5b, 7c, 8a, 9b, 10b, 14b, 20b, 21b, 22b, 23b, 25b, 27b, 30b, 33b, 34b, 38c, 41b, 42b, 44b, 47b, 49b, 61b, 62b, 75b, 78c, 80b. **17**, 17b, 36b, 43b, 47b, 69b, 70b, 71b, 72b, 74b. **18**, 3b, 6b, 7b, 8c, 37b, 38b, 51b, 52b, 59b, 63b, 68b, 76b, 77b, 78b.

Percentage support from 1000 bootstrap replications is given in brackets for internal branches. Branch lengths are proportional to the number of changes along each.

(*b*) *Phylogenetic relationships derived from molecular data*

(i) *Unrooted topology*

Very few positions were informative under Lake's evolutionary parsimony test (Lake 1987, 1989) and thus this approach was not applicable. A exhaustive parsimony analysis on all variable positions for echinoids alone found two unrooted trees, length 72, CI 0.86, RI 0.87, which can be expressed as a single tree with the three spatangoids forming a trichotomy (figure 2). One thousand bootstrap replications gave high levels of support for most groupings except the *Echinus*/*Paracentrotus* pairing and the *Sphaerechinus*/*Lytechinus* pairing. However, when the five sites (positions 249–253) whose alignment was ambiguous were omitted, a single unrooted tree, length 62, CI 0.89, RI 0.89, was found with two trichotomies. In addition to the spatangoid trichotomy found previously the three species of Echinidae form an unresolved trichotomy (figure 2). This results from a lack of informative sites rather than because of conflicting data. Distance matrix methods gave identical results.

Maximum Likelihood analysis produced the same topology. Although *Brissus* and *Spatangus* are paired among the three spatangoids, and *Psammechinus* and *Paracentrotus* paired among the three Echinidae, neither of these groupings is statistically significant (all other branches being supported at the $p < 0.05$ level) and both should be interpreted as trichotomies under this method.

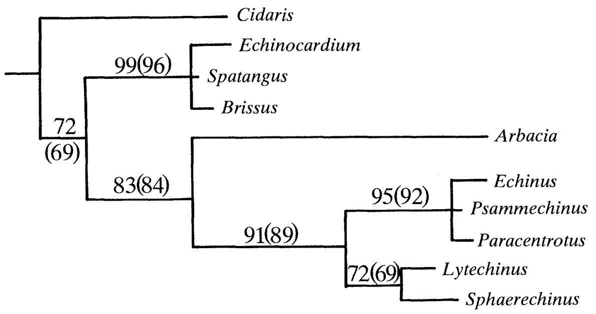

Figure 2. Relationships derived by parsimony, distance and maximum likelihood analyses of the molecular data. The tree shown is based on a parsimony analysis of the molecular data (excluding positions 249–253 in appendix 1 where alignment is ambiguous), and is length 62, CI 0.89, RI 0.89. When these five sites are included *Echinus* and *Paracentrotus* are grouped to the exclusion of *Psammechinus*. Percentage support for internal branches derived from 1000 bootstrap replications is shown; the first number derived from analysis that excludes positions 249–253, the second (in brackets) from an analysis of all 53 variant positions. Maximum Likelihood produced an identical tree with support at the $p < 0.05$ level for all internal branches but could not resolve the *Echinocardium–Spatangus–Brissus* trichotomy or the *Echinus–Paracentrotus–Psammechinus* trichotomy. Distance matrix methods also derived the same topology. Rooting is based on Maximum Likelihood, parsimony analysis of more conserved sites and parsimony analysis of the full data set after reweighting characters according to their consistency index. Branch lengths are proportional to the number of changes along each.

(ii) *Rooting the echinoid molecular tree*

Our initial analyses of echinoids produce trees that can only be rooted by reference to an outgroup. For our outgroup we have sequences from 12 echinoderms belonging to classes other than the Echinoidea (table 1). We used a number of approaches to define the polarity of molecular changes and thus root the echinoid tree.

First, the full sequence was used even though alignment between classes in some parts of the molecule was highly ambiguous. Maximum Likelihood analysis found a single tree, consistent with morphological data and rooted on *Cidaris* (figure 2). Support for internal branches was at the $p < 0.01$ level.

Under parsimony analysis, using pairs of species from each class in turn as outgroup, the internal topology remained more or less stable and consistent with the unrooted topology. However, this approach either rooted the echinoid tree on *Arbacia* (which is also the longest single branch amongst echinoids), or left the cidarid, stirodont, five camarodonts and three spatangoids as an unresolved polychotomy. Distance analysis also maintained the unrooted topology but rooted the tree on *Arbacia*. Only when multiple taxa from all four echinoderm classes were included did we obtain the echinoid tree rooted on *Cidaris* (length 160, CI 0.63, RI 0.73).

Because alignment with outgroup sequences was ambiguous in certain regions and thus highly suspect, we also ran parsimony analyses using only the more conserved stretches of sequence (Appendix 2). First, we used those positions that are invariant in the outgroup or virtually so and whose polarity is in no doubt (16 in total), to construct an outgroup. All other sites were scored as unknown. An exhaustive search found one tree, length 64, CI 0.88, RI 0.88 with the same topology as before and rooted on *Cidaris*. This tree is moderately robust. Although there is one tree one step longer and six trees two steps longer, all but one of these differ only in the topology of the five camarodonts, where there are relatively few characters. It is only in one of the trees two steps longer that the unrooted topology starts to break down and *Arbacia* is placed as sister group to irregular echinoids plus camarodonts. A consensus tree from the eight trees within two steps of the most parsimonious solution places irregular echinoids, camarodonts, *Cidaris* and *Arbacia* in a basal polychotomy. One thousand bootstrap replications supported the branch uniting euechinoids at the 72% level.

A slightly enlarged set of positions, comprising those sites in more conserved regions (25 positions) was used in a second set of analyses (Appendix 2). Although there is no ambiguity in alignment between the ingroup and outgroup in these partial sequences, the outgroup polarity is not always obvious for some of these positions. Ophiuroids, asteroids, holothurians and the single crinoid all individually rooted the echinoid tree on *Cidaris*, although the asteroid outgroup identified a second equally parsimonious tree which grouped *Arbacia* and *Cidaris* as sister taxa. However, when all four classes were used in combination to root the tree, parsimony analysis placed ophiuroids as sister group to *Cidaris* within the echinoid clade. Reweighting characters according to their consistency index on this tree and rerunning the analysis produced the topology of figure 2, with all outgroups united and the echinoid tree rooted on *Cidaris*.

In summary, all approaches identify the same unrooted topology from our molecular data and both Maximum Likelihood and parsimony analysis of conserved regions root this network on *Cidaris* (figure 2).

(*c*) *Comparative rates of evolution*

As morphological and molecular data give compatible cladograms (figures 1 and 2) the phylogenetic relationships of the ten echinoid species is not in dispute. Furthermore, because the fossil record of echinoids is relatively good, we can use the stratigraphical record of first occurrences to give us approximate times of divergence (Appendix 3). This gives 18 independent intervals of time through the past 250 Ma for which we have data on number of molecular mutations fixed and number of apomorphies accrued (figure 3, 1–18; table 2). There are also five major dichotomies (figure 3, I–V) where rates of evolution of sister groups can be compared.

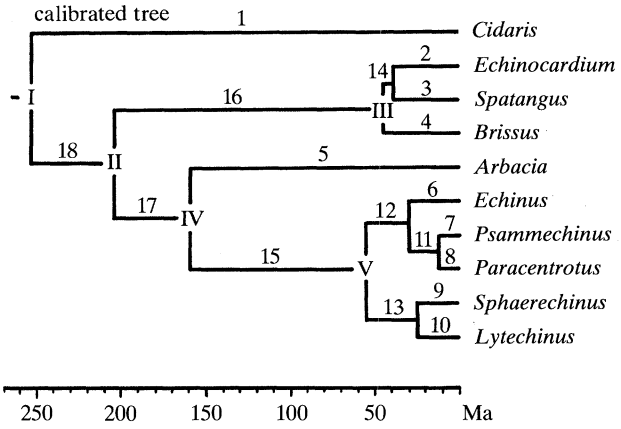

Figure 3. Calibrated tree, with nodes dated by reference to first occurrences in the fossil record (Appendix 3). Dichotomies at which sister groups were tested for equivalence in evolutionary rate are numbered I to V: the 18 independent time intervals for which rates of evolution can be estimated are numbered 1–18.

(i) *Assigning character state changes to specific branches*

Parsimony analysis, unlike distance or maximum likelihood methods, allows inferred nucleotide changes to be assigned to specific branches of the tree. A few sites will, however, remain ambiguous, because they represent differences between the basal sister taxa. Some of these can be polarized by reference to the outgroup. For example, in echinoids the basal dichotomy is between *Cidaris* and other echinoids (the Euechinoidea). There are 11 differences between the *Cidaris* sequence and the reconstructed ancestral euechinoid sequence. Four of these can be polarized by outgroup comparison (e.g. site 308 is U in *Cidaris* and all non-echinoid echinoderms, but A in euechinoids: the mutation is clearly from U to A in the stem group of Euechinoidea). How large a problem unpolarized sites represents will depend on the topology of the tree, the relative 'clock rate' and the number of taxa included within the group. The ancestral sequence for the eleven species of echinoid can be largely resolved, only seven sites (13%) remaining unclear. These seven

Table 2. *Number of molecular and morphological apomorphies for 18 time intervals, based on the first 410 bases from the 5′ end of the 28S rRNA molecule in ten echinoids, and a morphological data matrix of 81 characters*

(Time intervals 1–18 as in figure 3. Durations are given to the nearest 5 Ma. E = epifaunal; I = infaunal; N = predominantly nearshore (0–40 m); O = predominantly offshore (> 40 m).)

time interval	duration (Ma)	number of molecular changes[a] A	B	number of morphological changes[b] A	B	ecology
1	255	7–11	*ca.* 0.93	8	3.1	E, O
2	40	2	1.32	5	12.5	I, N, O
3	40	0	0.00	2	5.0	I, O
4	45	1	0.59	4	8.8	I, O
5	160	18	2.97	9	5.6	E, N
6	30	2	1.76	1	3.3	E, N
7	15	1	1.76	0	0.0	E, N
8	15	0	0.00	2	13.3	E, N
9	25	2	2.11	1	4.0	E, N
10	25	3	3.17	0	0.0	E, N
11	15	1	1.76	1	6.7	E
12	30	5	4.40	3	10.0	E
13	35	5	3.77	6	17.1	E
14	5	0	0.00	1	20.0	I
15	120	8	1.76	7	5.8	E
16	165	5	0.80	29	17.6	I
17	45	6	3.52	9	20.0	E
18	50	2–6	*ca.* 2.11	14	28.0	E

[a] Molecular changes: A = total number in each interval; B = rate of change (number of changes per Ma per site $\times 10^5$).
[b] Morphological changes: A = total number allocated to each interval; B rate of change (number of changes per Ma $\times 10^2$).

sites were evenly allocated between the two basal branches.

We have used no correction factor to accommodate missing substitutions in lineages with fewer branches, as recommended by Fitch & Bruschi (1987) and Fitch & Beintema (1990). This is because we found no correlation between number of branches and number of substitutions. The sparsely branched lineage leading to *Arbacia* shows more change than any of the more densely branched lineages. Furthermore, because the sequence we have chosen to study is far from saturated in echinoids, reversals and multiple hits are not a problem and observed number of fixed mutations must approximate closely to the true number.

(ii) *Comparative rates*

Durations and numbers of molecular and morphological apomorphies for each of the 18 time intervals are listed in table 2. Comparing sister groups, the expectation is that molecular rates should be approximately equivalent in the two branches if evolutionary rates are clock-like. Where multiple routes exist in a sister lineage the maximum and minimum numbers of mutations were calculated. A binomial probability statistic shows that in most cases sister groups do not deviate significantly from our expectation (table 3). However, two paired lineages display significantly different rates of molecular evolution, namely the cidaroid–euechinoid dichotomy and the irregular–echinacean dichotomy. In both cases a marked difference in rates exists, significant at the $p < 0.01$ level, indicating that both cidaroids and irregulars have a slow rate of molecular evolution in comparison to echinaceans.

The same approach can be taken with morphological apomorphies (table 3) and again differences in rates occur in some, but not all, pair-wise comparisons. Although differences in rates are found between the same branches as for molecular data there is no correlation. Morphologically irregular echinoids show the greatest rate of change, whereas for molecular data they have a comparatively slow 'clock-rate'.

To compare rates of morphological and molecular evolution further we investigated the correlation between the three variables, 'duration of geological interval', 'number of molecular changes' and 'number of morphological changes', using the 18 independent time intervals identified in figure 3. Bivariate scatterplots are shown in figure 4 together with their correlation coefficients. The Pearson product moment correlation coefficient, r_p, suggests that the strongest correlation exists between molecular change and time elapsed and the weakest between molecular change and morphological change. The correlation between morphological change and time elapsed is intermediate.

Table 3. *Tests for rate differences between sister taxa, using a binomial probability statistic*

(Molecular and morphological rates (number of mutations per site per Ma) are derived from the number of changes identified along each branch from maximum parsimony analysis. NS = binomial probability statistic with a probability of 0.5 not significant at $p = 0.05$ level. $p < 0.05$, $p < 0.01$ = binomial probability statistic significant at the $p < 0.05$, $p < 0.01$ levels. Roman numerals refer to nodal points in figure 3 at which sister branches are compared. The superscript letter indicates the branch as in figure 3.)

molecular data	
I^{1} – 9 hits; 0.9×10^{-4}	
against	
I^{18-16} – 9–11 hits; $0.9\text{–}1.1 \times 10^{-4}$	NS
I^{18-17} – 24–28 hits; $2.5\text{–}2.9 \times 10^{-4}$	$p < 0.01$
II^{16} – 5–7 hits; $6.4\text{–}9.0 \times 10^{-5}$	
against	
II^{17} – 20–23 hits; $2.6\text{–}3.0 \times 10^{-4}$	$p < 0.01$
III^{14} – 0–2 hits; $< 1.3 \times 10^{-4}$	
against	
III^{4} – 1 hit; 5.8×10^{-5}	NS
IV^{5} – 18 hits; 2.6×10^{-4}	
against	
IV^{15} – 14–17 hits; $2.1\text{–}2.5 \times 10^{-4}$	NS
V^{12} – 6–7 hits; $2.6\text{–}3.1 \times 10^{-4}$	
against	
V^{13} – 7–8 hits; $3.1\text{–}3.5 \times 10^{-4}$	NS
morphological data	
I^{1} – 8 apomorphies	
against	
I^{18-16} – 45–48 apomorphies	$p < 0.001$
I^{18-17} – 33–37 apomorphies	$p < 0.001$
II^{16} – 31–34 apomorphies	
against	
II^{17} – 19–23 apomorphies	$p < 0.05$
III^{14} – 3–5 apomorphies	
against	
III^{4} – 5 apomorphies	NS
IV^{5} – 10 apomorphies	
against	
IV^{15} – 11–13 apomorphies	NS
V^{12} – 4–6 apomorphies	
against	
V^{13} – 6–7 apomorphies	NS

However, the assumptions of this parametric test of correlation may not be valid as they require normal distribution. Furthermore, although molecular change can be quantified in an absolute way, both the divergence time and the number of morphological apomorphies assigned to each branch cannot. To overcome these problems we prefer to use the non-parametric Spearman's rank correlation coefficient, r_s (Tippett 1952). This has the advantage that it compares ranks rather than absolute values and does not assume a linear relationship between populations. Thus we need only the relative magnitudes of time elapsed since divergence and morphological apomorphies accrued, a more realistic goal given the problems of estimating divergence times and morphological change. This approach gives a rather different picture, suggesting that morphological change is slightly better correlated with time than molecular change.

However, the pair-wise comparisons (table 3)

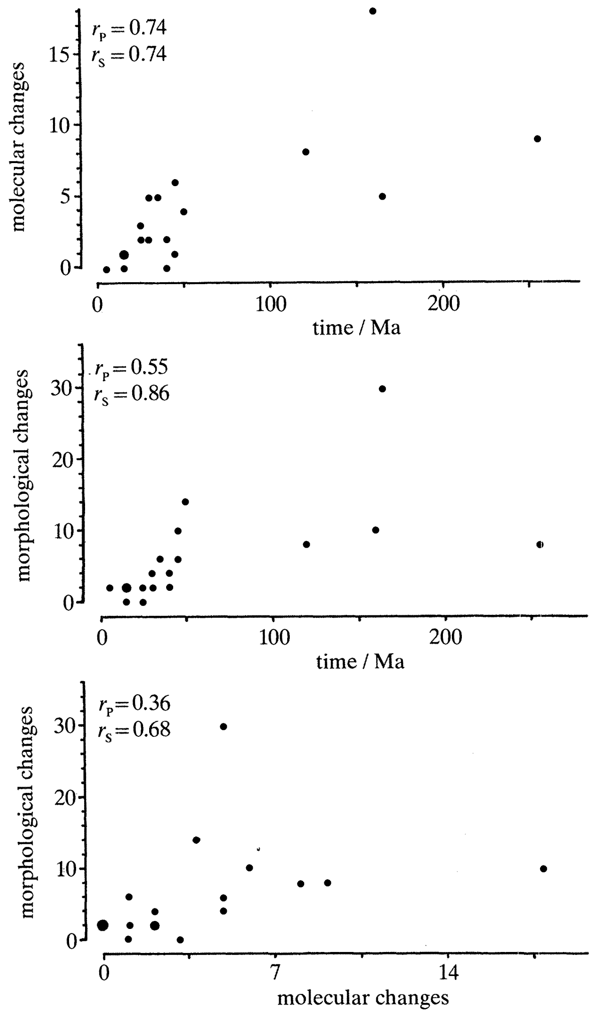

Figure 4. Bivariate scatter plots of morphological and molecular change against time elapsed for the 18 time intervals shown in figure 3. r_P = Pearson's correlation coefficient; r_s = Spearman's rank correlation coefficient.

demonstrate that molecular rates are significantly different between major clades, and thus should not be averaged together. Removing the 'slow-clock' cidaroid and irregular echinoid branches greatly improves the correlation between molecular change and time elapsed (figure 5), whereas the correlation between morphological change and time elapsed remains weak.

4. DISCUSSION

(i) *Choice of sequence and taxa*

In recent years much effort has been devoted to refining mathematic treatments designed to obtain phylogenies from sequence data (Swofford & Olsen 1990, Nei 1987). Clearly this is important but such methods have become so computerized that it is now virtually impossible to check assumptions and calculations by any other way than by comparing the results of the various methods. This has led to an emphasis on the quantity of variant sites being analysed rather than the choice of species or selectivity in the nucleotide sequences analysed (see Ghiselin (1988, 1989) for similar remarks).

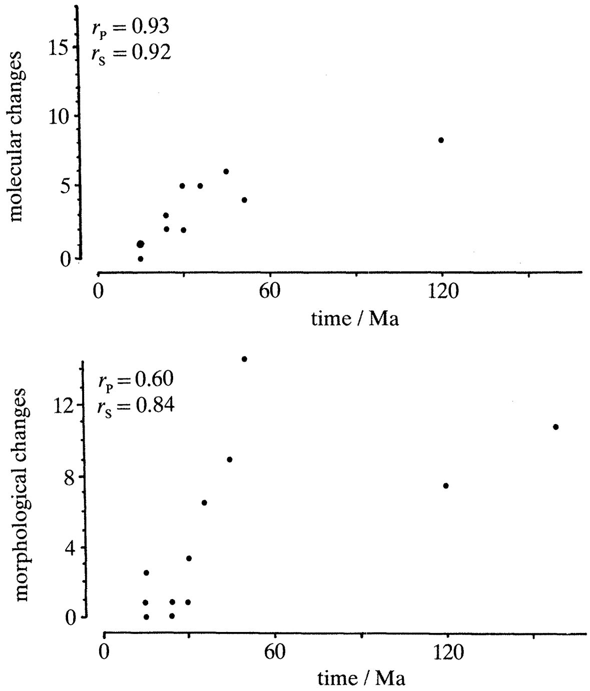

Figure 5. Bivariate scatter plots of morphological and molecular change against time elapsed for just echinacean taxa (branches 5–13, 15, 17, 18 in figure 3). r_P = Pearson's correlation coefficient; r_s = Spearman's rank correlation coefficient.

In our view the choice of the sequence analysed is of paramount importance. Sequences must not be approaching saturation and alignment must be unambiguous, yet there must be sufficient point mutations to be able to generate an unambiguous solution. We therefore chose a molecular region that Michot & Bachellerie (1987) had previously shown to have an appropriate rate of evolution for the phylogenetic question in hand, namely the 5′ end of the 28*S* rRNA molecule.

In our analysis 14% of sites are variant among the ten echinoid species. Alignment is unambiguous except for a very few positions (where additions/deletions are abundant), and the high consistency index and the virtual absence of character reversals demanded by the most parsimonious cladogram provides good evidence that this part of the sequence remains far from saturated among echinoids. Only 12.5% of variable sites have more than two alternative nucleotide bases in the ten echinoid sequences.

Our conservative approach, however, means that some trichotomies cannot be resolved from this domain, because divergences are too recent and/or rates of change too slow. But domains where sequences have diverged sufficiently to provide resolution to these trichotomies would be too variant to align among all echinoids and/or too saturated with double hits and reversals to provide a secure higher level phylogeny (see Lanyon (1988) for a discussion).

Interestingly, the problems encountered in obtaining a fully resolved phylogenetic reconstruction are the same for both morphological and molecular data, and it is the same regions of the cladogram that are the least supported in bootstrap testing. Resolution of polychotomies requires characters that vary among the ingroup but which can also be homologized and polarized with respect to the outgroup.

In molecular sequences, portions of the 28*S* rRNA molecule more variable than the present domain could be used to resolve our two trichotomies. However, there then arises the problem of alignment and polarization. More divergent regions we have examined, such as the D2 domain (located immediately to the 3′ end of the region discussed here), entail significant ambiguity of alignment between major branches of Echinoidea. Thus, although there are a greater number of variable sites to work with, alignment and polarity for these regions, even within echinoids, can be problematic and no resolution of the trichotomies is at the moment possible.

For morphological data exactly the same holds true. There are a number of morphological characters that vary among the three spatangoids, but because these are unique to spatangoids and none is present in any of the other echinoids in this particular study, it is impossible to polarize them, and thus the trichotomy remains unresolved.

The solution in both cases is the addition of a close outgroup (i.e. the Cretaceous stem group for morphological data or Recent hemiasterid spatangoids or holasteroids for molecular data). The polarity of most of these characters could then be determined and the trichotomy resolved for both morphological and molecular data. Similarly, the secure resolution of relationships among the camarodont echinoids requires the addition of other camarodont groups, particularly temnopleurids, to establish polarities.

The taxonomic level at which we have chosen to investigate comparative rates of morphological and molecular evolution is also important. Obviously morphological comparison becomes increasingly difficult between higher taxonomic groups that have been separated for longer periods but so also does molecular comparison, with more and more molecules or domains within molecules reaching saturation levels as divergence time increases. This effectively restricts comparison to only the most highly conserved morphological and molecular data for more distant divergences and greatly reduces the number of characters available, particularly from morphology. Conversely, for relatively recently diverged taxa and fast-evolving domains, the uncertainties in estimating divergence times from the fossil record start to become much more significant and can substantially affect estimates of rate.

(ii) *Rooting molecular trees*

Our analysis of rates of evolution depends first and foremost upon our having rooted the molecular tree correctly, since different rootings can significantly alter the results. A tree rooted on *Arbacia*, as found by using complete sequences from our outgroup, produces rates of change that are much more uniform along the various branches. However, there are good reasons to believe that this rooting is an artifact that arises because the outgroup polarity at most positions is effectively random.

Of the 56 sites that are variant among echinoid species, some (e.g. positions 308 and 321) are invariant in all echinoderm species in our outgroup and thus their polarity is completely unambiguous. Other sites are represented by two or more nucleotides in our outgroup and determining the plesiomorphic state is often impossible. Omitting two positions where the outgroup state was represented predominantly by deletions, 28% of sites are completely invariant in all outgroup sequences, 18.5% are transitions only, 16.5% are transversions between two nucleotides only, 24% are sites that vary among three nucleotides and 13% are sites at which all four nucleotides are found in the outgroup. Two positions that are variant in the outgroup have only a single species where the nucleotide differs and polarity in these cases is also clear. This means that for about 70% of the positions outgroup polarity is ambiguous or unknown and polarity assignment is, in effect, random for most sites.

As Wheeler (1990) has cogently demonstrated, a random outgroup sequence will join to the longest branch of the ingroup and the root position is meaningless. Thus the identification of *Arbacia*, the longest branch in our tree, as root is most probably spurious.

For random outgroup sequences Wheeler (1990) suggested using asymmetry in the character transformations within the ingroup to identify the root. Here, however, we have adopted a simpler approach. When only those positions for which polarity is unambiguous are used to root the tree, the topology matches that found from morphology and maximum likelihood analyses but implies markedly uneven rates of evolution between sister groups.

(iii) *Rates of evolution*

Since both molecular and morphological data produced a similar cladogram for the ten species (after removal of *Echinocyamus*), the two approaches are not in conflict and the two approaches can be compared constructively.

Our echinoid data suggest that molecular changes in 28*S* rRNA correlate strongly with the length of time elapsed but only within specific clades. Comparison between clades indicates that there have been marked changes in rate through geological time and lineages differ by a factor of three in their rates of molecular evolution. At the level of divergence studied here, any errors in estimating divergence time constitute only a relatively small portion of the total time since separation, and certainly cannot account for a threefold difference in mutation rate. Martin *et al.* (1990) have pointed out that in estimating genetic distance the coefficient of variation is large when the number of bases sampled is relatively small (i.e. less than 1000). Thus the variation in molecular rates of change between echinoid clades might disappear were a larger data base available, if our results are no more than a sampling artifact. However, the fact that we can demonstrate a tight correlation between molecular change and time within a well-defined clade (figure 5) suggests that the difference in rate between subgroups is genuine. Similar marked fluctuations in molecular rates have been documented from larger data bases among hominoids (Koop *et al.* 1989), carnivores and primates (Wayne *et al.* 1991), and tetrapods (our own analysis of the data presented by Hedges *et al.* 1990).

It would seem therefore that, among the echinoids studied and for the portion of the rRNA molecule analysed, rates of molecular evolution are relatively clock-like within certain clades, but show significant rate differences between clades. When molecular rates are averaged over a range of echinoid clades, however, the correlation between time elapsed and number of molecular changes becomes distinctly weaker. Bailey *et al.* (1991) have made a similar point with respect to global versus local molecular clock rates in primates. Indeed, when averaged over sufficient taxa, morphological change in echinoids can behave in just as 'clock-like' a fashion as molecular data. Thus for our 10 echinoid species, the acquisition of morphological innovations is as strongly correlated with time as the fixation of point mutations.

Variation in molecular rates has previously been ascribed to differences in generation time, mutation rate or replication and repair mechanisms (e.g. Koop *et al.* 1989, Li & Tanimura 1987, Britten 1986, Gillespie 1986, Wilson *et al.* 1990). Available data on reproduction suggests that most echinoids reach sexual maturity within approximately 1–2 years (Lawrence 1987) and, because none of the species studied here are brooders, all have a comparably high DNA replication rate and produce very large numbers of gametes per year. Nor do irregular echinoids appear to be particularly different from echinaceans in their reproductive biology (e.g. Guillou 1985). Thus it would appear that molecular rates show marked fluctation because of differences in repair and replication processes, or mutation rate.

In the light of these observations a correlation between rate of molecular change and habitat may be significant. It is the three infaunal echinoids and the deep-sea echinoid that show slower molecular change, whereas shallow water epifaunal species (with strong epithelial pigmentation against UV radiation) show significantly higher rates (table 2). Thus there seems to be at least a *prima facie* case for the organism's ecology having a strong effect on its molecular rate of evolution. We can only speculate as to the possible causes for this. From the above correspondence, mutation rates would appear to be higher in shallow-water epifaunal species as opposed to deep-water epifaunal or shallow infaunal species. Work on comparing deep-water and shallow water sister taxa is underway to test this hypothesis.

The thrust of this paper is not to demonstrate that molecular data are superior to morphological data for phylogenetic reconstruction, or vice versa, but to compare the two constructively. We believe that a better understanding of evolutionary processes will only come about through a sensible integration of molecular and morphological data. We should try to focus not on areas where the two approaches are in discord (although these have their own intrinsic interest) but where there is concord, as this then allows broader evolutionary questions to be tackled.

We thank the staff of the marine stations at Villefranche sur Mer and Roscoff and at the Oceanographic Museum at Monaco for supplying and identifying the echinoderms studied here. Colin Patterson, Simon Conway Morris and three anonymous referees gave constructive criticism of an earlier draft, for which we are grateful.

This work was supported by the C.N.R.S. and the Université de Paris VI, and grants from the Association Recherche et Partage as well as the Conseil Général des Alpes Maritimes

REFERENCES

Bailey, W.J., Fitch, D.H.A., Tagle, D.A., Czelusniak, J., Slightom, J.L. & Goodman, M. 1991 Molecular evolution of the In-Globin gene locus: gibbon phylogeny and the hominoid slowdown. *Molec. Biol. Evol.* **8**, 155–184.

Baroin, A., Perasso, R., Qu, L.H., Brugerolle, G., Bachellerie J.-P. & Adoutte, A. 1988 Partial phylogeny of the unicellular eukaryotes based on rapid sequencing of a portion of 28*S* ribosomal RNA. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3474–3478.

Britten, R.J. 1986 Rates of DNA sequence evolution differ between taxonomic groups. *Science, Wash.* **231**, 1393–1398.

Cockburn, A., Mansergh, I.M., Broome, L.S. & Ward S. 1990 Molecular clocks and generation time in burramyid marsupials. *Molec. Biol. Evol.* **7**, 283–285.

Farris, J.S. 1988 *Hennig86, version 1.5* (computer program). Port Jefferson Station, New York.

Felsenstein, J. 1989 *PHYLIP* (computer program). University of Washington, Seattle.

Fitch, W.M. & Beintema, J.J. 1990 Correcting parsimonious trees for unseen nucleotide substitutions: the effect of dense branching as exemplified by Ribonuclease. *Molec. Biol. Evol.* **7**, 438–443.

Fitch, W.M. & Bruschi, M. 1987 The evolution of prokaryotic ferredoxins – with a general method correcting for unobserved substitutions in less branched lineages. *Molec. Biol. Evol.* **4**, 381–394.

Gammelin, M., Altmuller, A., Reinhardt, U., Mandler, J., Harley, V.R., Hudson, P.J., Fitch, W.M. & Scholtissek, C. 1990 Phylogenetic analysis of nucleoproteins suggests that human Influenza A viruses emerged from a 19th-Century avian ancestor. *Molec. Biol. Evol.* **7**, 194–200.

Gillespie, J.H. 1986 Rates of molecular evolution. *A. Rev. Ecol. Syst.* **17**, 637–665.

Ghiselin, M.T. 1988 The origin of molluscs in the light of molecular evidence. *Oxf. Surv. Evol. Biol.* **5**, 66–95.

Ghiselin, M.T. 1989 Summary of our present knowledge of metazoan phylogeny. In *The hierarchy of life. Molecules and morphology in phylogenetic analysis* (ed. B. Fernholm, K. Bremer & H. Jörnvall), pp. 261–272. Amsterdam: Excerpta Medica.

Gonzalez I.L., Sylvester, J.E., Smith, T.F., Stambolian, D. & Schmickel, R.D. 1990 Ribosomal RNA gene sequences and hominoid phylogeny. *Molec. Biol. Evol.* **7**, 203–219.

Guillou, J. 1985 Population dynamics of *Echinocardium cordatum* (Pennant) in the bay of Douarnenez (Brittany). In *Echinodermata: proceedings of the fifth international echinoderm conference, Galway, September, 1984* (ed. B. F. Keegan & B. D. S. O'Connor), pp. 275–280. Rotterdam: A.A. Balkema.

Harland, W.B., Armstrong, R.L., Cox, A.V., Craig, L.E., Smith, A.G. & Smith, D.G. 1989 *A geological time scale, 1989*. Cambridge University Press.

Hedges, S.B., Moberg, K.D. & Maxson, L.R. 1990 Tetrapod phylogeny inferred from 18*S* and 28*S* ribosomal RNA sequences and a review of the evidence for amniote relationships. *Molec. Biol. Evol.* **7**, 607–633.

Hillis, D.M. & Dixon, M.T. 1989 Vertebrate phylogeny: evidence from 28*S* ribosomal DNA sequences. In *The hierarchy of life. Molecules and morphology in phylogenetic analysis* (ed. B. Fernholm, K. Bremer & H. Jörnvall), pp. 355–370. Amsterdam: Excerpta Medica.

Hillis, D.M. & Dixon, M.T. 1991 Ribosomal DNA: molecular evolution and phylogenetic inference. *Q. Rev. Biol.* **66**, 411–453.

Holmquist, R., Miyamoto, M.M. & Goodman, M. 1988 Analysis of higher-primate phylogeny from transversion differences in nuclear and mitochondrial DNA by Lake's method of evolutionary parsimony and operation metrics. *Molec. Biol. Evol.* **5**, 217–236.

Hyman, L.H. 1955 *The invertebrates*, vol. IV (*Echinodermata*). New York: McGraw-Hill.

Jensen M. 1981 Morphology and classification of Euechinoidea Bronn, 1860 – a cladistic analysis. *Vidensk. Meddr. dansk naturh. Foren.* **143**, 7–99.

Jesionek-Szymanska, W. 1970 On a new pygasteroid (Echinoidea) from the Jurassic (Middle Lias) of Nevada, U.S.A. *Acta Palaeont. Polon.* **15**, 411–423.

Kier, P.M. 1984 Fossil spatangoid echinoids of Cuba. *Smithson. Contr. Paleobiol.* **55**, 1–336.

Koop, B.F., Tagle, D.A., Goodman, M. & Slightom, J.L. 1989 A molecular view of primate phylogeny and important systematic and evolutionary questions. *Molec. Biol. Evol.* **6**, 580–612.

Krajewski, C. 1990 Relative rates of single-copy DNA evolution in Cranes. *Molec. Biol. Evol.* **7**, 65–73.

Lake, J.A. 1987 A rate-independant technique for analysis of nucleic acid sequences: evolutionary parsimony. *Molec. Biol. Evol.* **4**, 167–191.

Lake, J.A. 1989 Origin of the eukaryotic nucleus determined by rate-invariant analyses of ribosomal RNA genes. In *The hierarchy of life. Molecules and morphology in phylogenetic analysis* (ed. B. Fernholm, K. Bremer & H. Jörnvall), pp. 87–101. Amsterdam: Excerpta Medica.

Lanyon, S.M. 1988 The stochastic model of molecular evolution: what consequences for systematic investigations? *Auk* **105**, 565–573.

Larson, A. & Wilson, A.L. 1989 Patterns of ribosomal RNA evolution in salamanders. *Molec. Biol. Evol.* **6**, 131–154.

Lawrence, J.M. 1987 *A functional biology of echinoderms*. London: Croom Helm.

Li, W.H. & Tanimura, M. 1987 The molecular clock runs more slowly in man than in apes and monkeys. *Nature, Lond.* **326**, 93–96.

Martin, A.P., Kessing, B.D. & Palumbi, S.R. 1990 Accuracy of estimating genetic distance between species from short sequences of mitochondrial DNA. *Molec. Biol. Evol.* **7**, 485–488.

Michot, B. & Bachellerie, J.-P. 1987 Comparisons of large subunit rRNAs reveal some eukaryote-specific elements of secondary structure. *Biochimie* **69**, 11–23.

Mortensen, T. 1925–1951. *A monograph of the Echinoidea*, vol. I–V. Copenhagen: C.A. Reitzel.

Nei, M. 1987 *Molecular evolutionary genetics*. New York: Columbia University Press.

Pace, N.R., G.J. Olsen & Woese, C.R. 1986 Ribosomal RNA phylogeny and the primary lines of evolutionary descent. *Cell* **45**, 325–326.

Qu, L.H., Michot, B. & Bachellerie, J.-P. 1983 Improved methods for structure probing in large RNAs: a rapid "heterologous" sequencing approach is coupled to the direct mapping of nuclease accessible sites. Application to the 5′ terminal domain of eukaryotic 28*S* RNA. *Nucl. Acids Res.* **11**, 5903–5920.

Raff, R.A., Field, K.G., Ghiselin, M.T., Lane, D.J., Olsen, G.J., Pace, N.R., Parks, A.L., Parr, B.A. & Raff, E.C. 1988 Molecular analysis of distant phylogenetic relationships in echinoderms. In *Echinoderm phylogeny and evolutionary biology* (ed. C. R. C. Paul & A. B. Smith), pp. 29–41. Oxford: Clarendon Press.

Ratto, A. & Christen, R. 1990 Phylogénie moléculaire des échinodermes déduite de séquences partielles des ARN ribosomiques. *C. r. Acad. Sci., Paris* **310**, 169–173.

Sa, R.O. De & Hillis, D.M. 1990 Phylogenetic relationships of the pipid frogs *Xenopus* and *Silurana*: an integration of ribosomal DNA and morphology. *Molec. Biol. Evol.* **7**, 365–376.

Saitou, N. & Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molec. Biol. Evol.* **4**, 406–425.

Smith, A.B. 1981 Implications of lantern morphology for the phylogeny of post-Palaeozoic echinoids. *Palaeontology* **24**, 779–801.

Smith, A.B. 1984 *Echinoid palaeobiology*. London: Allen and Unwin.

Smith, A.B. 1988 Phylogenetic relationships, divergence times and rates of molecular evolution for camarodont sea urchins. *Molec. Biol. Evol.* **5**, 345–365.

Smith, A.B. 1989 RNA sequence data in phylogenetic reconstruction: testing the limits of its resolution. *Cladistics* **5**, 321–344.

Smith, A.B. & Hollingworth, N.T.J. 1990 Tooth structure and phylogeny of the Upper Permian echinoid *Miocidaris keyserlingi*. *Proc. Yorks. geol. Soc.* **48**, 47–60.

Springer, M.S. & Kirsch, J.A.W. 1989 Rates of single-copy DNA evolution in phalangeriform marsupials. *Molec. Biol. Evol.* **6**, 331–341.

Stackebrandt, E. & Woese, C.R. 1981 The evolution of prokaryotes. In *Molecular and cellular aspects of microbial evolution* (ed. M. J. Carlisle, J. R. Collins & B. E. B. Moseley), pp. 1–31. Cambridge University Press.

Swofford, D.L. 1990 *PAUP, phylogenetic analysis using parsimony, version 3.0* (computer program). Illinois Natural History Survey, Champaign, Illinois.

Swofford, D.L. & Olsen, G.J. 1990 Phylogeny reconstruction. In *Molecular systematics* (ed. D. M. Hillis & C. Moritz), pp. 411–501. Sunderland Mass.: Sinauer Assoc. Inc.

Tippett, L.H.C. 1952. *The methods of statistics*, 4th edn. New York: J. Wiley and Sons.

Wayne, R.K., Valkenburgh, B. Van & O'Brien, S.J. 1991 Molecular distance and divergence time in carnivores and primates. *Molec. Biol. Evol.* **8**, 297–319.

Wilson, M.A., Gaut, B. & Clegg, M.T. 1990 Chloroplast DNA evolves slowly in the palm family (Arecaceae). *Molec. Biol. Evol.* **7**, 303–314.

Wheeler, W.C. 1990 Nucleic acid sequence phylogeny and random outgroups. *Cladistics* **6**, 363–368.

Woese, C.R. 1989 Archaeobacteria and the nature of their evolution. In *The hierarchy of life. Molecules and morphology in phylogenetic analysis*. (ed. B. Fernholm, K. Bremer & H. Jörnvall), pp. 119–130. Amsterdam: Excerpta Medica.

Zuckerkandl, E. & Pauling, L. 1962 Molecular disease, evolution, and genic diversity. In *Horizons in biochemistry* (ed. M. Kasha & B. Pullman), pp. 189–225. New York: Academic Press.

Received 2 March 1992; accepted 31 March 1992

APPENDIX 1

Morphological characters used. Apomorphic states that can be polarized unambiguously by reference to stem group echinoids and/or other echinoderm groups are shown in **bold**, unpolarized states are indicated by a [U]. Note, however, that all characters were treated as unordered in our parsimony analyses.

Test morphology

Shape

1. a, radially pentameral (0); b, **bilaterally symetrical** (1).

Mouth

2. a, circular and central (0); b, **anterior and D-shaped** (1).
3. Buccal notches and pharyngeal expansion sacs: a, absent (0); b, **present** (1).
4. Buccal notches when present: a, shallow (0); b, **deep** (1).
5. Mouth: a, large relative to test, primordial plates greatly reduced or lost (0); b, **small relative to test, primordial plates prominent** (1).
6. a, Ambulacral and interambulacral plates migrate from corona to peristome (0); b, **no migration of plates during life** (1).
7. a, multiple pore-bearing plates on peristome (0); b, **10 buccal plates only** (1); c, **no tube feet on peristome** (2).

Perignathic girdle

8. a, absent (0); b, **present – both protractor and retractor muscles attach to interambulacral apophyses** (1); c, **present – retractors attach to ambulacral auricles, protractors to poorly developed interambulacral apophyses** (2).

Apical disc

9. a, Encloses periproct (0); b, **periproct outside apical disc** (1).
10. a, G5 present (0); b, **G5 absent** (1).
11. a, Periproct composed of a mass of small plates with no suranal plate (0); b, **periproct with enlarged 'suranal' plate during development** (1).
12. Periproct: a, not valve-like (0); b, **valve-like, composed of four triangular plates** (1).
13. G5 crystalographic c-axis; a, perpendicular (0), b, tangential (1) [U].
14. a, Apical disc of alternating ocular and genital plates (0); b, apical disc ethmolytic (1) [U].
15. Oculars I,V: a, exsert (0); b, insert (1) [U].
16. Periproct separated from peristome by: a, five interambulacral plates (0); b, four interambulacral plates (1) [U].

Ambulacra

17. a, Simple (0); b, **compound** (1).
18. Compound plating; a, all elements reaching the perradial suture (0); b, with middle element dominant and upper and lower elements occluded (Arbaciid style) (1); c, with lowest element dominant and middle elements occluded (Echinid style) (2) [U].
19. Compound plates: a, trigeminate (0); b, **polygeminate** (1).
20. Aboral pores: a, simple, no respiratory tube feet (0); b, **conjugate, supporting functionally specialized respiratory tube feet** (1).
21. Phyllodes: a, absent (0); b, **spatangoid type** (1).
22. Penicillate tube-feet around mouth: a, absent (0); b, **present** (1).
23. Pores: a, double (0); b, **reduced to single openings below petals** (1).
24. Pores in each ambulacral plate series: a, uniserial (0); b, **in strong arcs throughout** (1); c, **uniserial above, broadening markedly adorally** (2) [U].
25. Ambulacral III pores and tube-feet: a, not differentiated from other ambulacra (0); b, **differentiated** (1).
26. a, Penicillate tube-feet not developed in ambulacrum III adapically (0); b, **penicillate tube-feet developed adapically in ambulacrum III** (1).
27. Petals: a, not developed (0); b, **present** (1).
28. Ambulacra I,II,IV,V: a, flush (0); b, **sunken** (1).
29. Ambulacrum III: a, flush (0); b, **sunken** (1).
30. Subanal penicillate tube-feet: a, absent (0); b, **present** (1).
31. Subanal fasciole encloses: a, 3 ambulacral plates (0); b, 5–6 ambulacral plates (1) [U].
32. Adapically ambulacra (petals): a, parallel-sided (0); b, lanceolate (1); c, expanding adapically (2) [U].

Interambulacra

33. a, Both plates in a column reach peristome edge (0); b, **one plate only reaches peristome** (1).
34. a, All five interambulacra similar (0); b, **sternal plating differentiated** (1).
35. In posterior interambulacrum, first plate (labral plate): a, equidimensional (0); b, broad and short (1) [U].
36. Orientation of crystalographic C-Axis relative to surface of corona: a, perpendicular (0); b, **tangential** (1).

Spines and tuberculation

37. Sphaeridiae: a, absent (0); b, **present** (1).
38. Spines: a, heterogeneous, with a primary spine surrounded by a scrobicule (0); b, with a few large primary interambulacral tubercles and spines on each plate and no scrobicular circles (1); c, with large numbers of small equal-sized tubercles scattered over plate – spines forming hair-like coat (2) [U].
39. Spines: a, hollow (0); b, **solid** (1).
40. Spines: a, lacking cortex (0); b, **with cortex** (1).
41. Spines: a, simple, needle-like (0); b, **small, spatulate** (1).
42. Spines and tubercles: a, undifferentiated around test (0); b, **functionally and structurally differentiated** (1).

43. Tubercles: a, perforate (0); b, **imperforate** (1).
44. Tubercles: a, non-crenulate (0); b, **crenulate** (1).
45. Aboral spines and tubercles: a, uniform, not sunken (0); b, **very heterogeneous, with large, sunken tubercles bearing large spines scattered about** (1).
46. Primary interambulacral tubercle: a, on every plate (0); b, **on alternate plates in a column** (1).
47. Modified miliaries (clavulae): a, absent (0); b, **present** (1).
48. Inner (aboral) fasciole: a, absent (0); b, **present** (1).
49. Subanal fasciole: a, absent (0); b, **present** (1).
50. Peripetalous fasciole: a, absent (0); b, **present** (1).

Pedicellariae

51. Ophicephalous pedicellariae: a, absent (0); b, **present** (1).
52. Triphylous pedicellariae: a, absent (0); b, **present** (1).
53. Globiferous pedicellariae: a, absent (0); b, present (1) [U].
54. Globiferous pedicellariae: poison gland a, external, opening not through pore (0); b, **internal, opening through pore in blade** (1).
55. Globiferous pedicellariae: a, without terminal fang (0); b, with terminal fang (1) [U].
56. Globiferous pedicellariae: a, without ring of small teeth around distal opening (0); b, **with ring of small teeth** (1) [U].
57. Globiferous pedicellariae: a, lacking paired lateral teeth (0); b, with paired lateral teeth (1) [U].
58. Globiferous pedicellariae: poison glands carried a, on the valves only (0); b, on the valves and stalk (1) [U].
59. Blade of globiferous pedicellariae: a, open meshwork (0); b, **narrow fused cylinder** (1).
60. Stalk of globiferous pedicellariae: a, unfused mass of calcite fibres (0); b, fused and tubular (1); c, a solid meshwork (2) [U].
61. Tridentate pedicellariae: a, with long, narrow blades (0); b, **include rostrate form with bulbous blades** (1).

Lantern

62. Lantern: a, present (0); b, **absent** (1).
63. Hemipyramids: a, with small foramen magnum (cidaroid-type) (0); b, **with deep v-shaped foramen magnum (euechinoid-type)** (1).
64. Epiphysis: a, without large projection (0); b, **with prominent lateral projection supporting tooth** (1).
65. Epiphysis: a, not meeting and fused (0); b, **fused above foramen magnum** (1).
66. Rotula: a, hinge-socket type joint (0); b, **ball and socket type joint** (1).
67. Hemipyramids: a, not enlarged (0); b, **enlarged into a processus superalveolaris** (1).

Teeth (see Jensen 1981)

68. a, grooved (0); b, **keeled** (1).
69. Central lammellae-needles-prisms (CLNP) complex: a, lacking flabelliform lamellae (0); b, with flabelliform lamellae (1) [U].
70. CNLP complex: a, without tines (0); b, with tines (1) [U].
71. CLNP/LLNP (lateral lamellae-needles-prisms) complexes: a, undifferentiated (0); b, the two complexes distinct (1) [U].
72. Tooth plate: a, umbo-lateral distance less than umbo-oral distance (0); b umbo-lateral distance greater than umbo oral distance (1) [U].
73. Primary tooth plate: a, undifferentiated (cidaroid) (0); b, arbacioid (1); c, echinoid (2) [U].
74. Secondary tooth plate (lappet): a, a narrow simple plate (0); b, well developed with a carinal appendage (1) [U].

Soft tissue anatomy

75. a, three sinuses (periproctal, perianal and genital) all separated (0); b, these three sinuses represented by one small sinus (1) [U].
76. Syphon on intestine: a, absent (0); b, present (1) [U]
77. Circum-oesophageal ring with: a, 5 discrete spongy bodies (?pollian vesicles) (0); b, a continuous spongy ring (1) [U].
78. Mucous glands found: a, throughout the gut (0); b, **in the oesophagus and pharynx only** (1); c, **in a small ring at the top of the pharynx only** (2).
79. Outer marginal duct of the digestive haemal system: a, without a collateral duct (0); b, with a collateral duct (1) [U].
80. Caecum at start of stomach: a, absent (0); b, prominent (1).
81. Frontal ambulacrum: a, unmodified for food collecting (0); b, **with modified glandular spines developed for mucous string feeding** (1).

Character matrix

outgroup	00000	00000	00?00	?00?0	00000	00000	?000?	0?000	00000	00000	00???	?????	00000	000??	?????	?????	0
Cidaris	00000	00100	00000	?00?0	00000	00000	?000?	10011	00000	00000	00111	10002	00000	10000	00000	00000	0
Echinocardium	11001	12011	00?10	000?1	11101	11011	12110	01200	11010	01110	11100	10012	11???	?????	????1	10201	1
Spatangus	11001	12011	00?10	000?1	11101	01001	11110	01200	11011	01010	11100	10012	11???	?????	????1	10201	0
Brissus	11001	12011	00?10	100?1	11101	01101	20111	01200	11010	01011	11100	10012	11???	?????	????1	10201	0
Arbacia	00100	11200	01100	?1101	00020	00000	?000?	11111	00100	00000	110??	?????	00100	01111	11110	11100	0
Echinus	00100	11200	10100	?1200	00010	00000	?000?	11110	00100	10000	11101	01010	00101	00111	10210	11110	0
Psammechinus	00100	11200	10100	?1200	00010	00000	?000?	11110	00100	00000	11101	01000	00101	00111	10210	11110	0
Paracentrotus	00100	11200	10100	?1210	00010	00000	?000?	11110	00100	00000	11101	01000	00101	00111	10210	11100	0
Sphaerechinus	00110	11200	00101	?1210	00010	00000	?000?	11110	00100	00000	11101	00111	00111	00111	11210	11110	0
Lytechinus	00110	11200	00101	?1200	00010	00000	?000?	11110	00100	00000	11101	00111	00111	00111	11210	111?0	0

APPENDIX 2

Molecular data: full data set. Regions underlined could not be aligned unambiguously between echinoderm classes and were treated as unknown for outgroup purposes in some of our analyses. A ‘v’ on the top line mark sites that are variant among the echinoids. Bases identical to those on the top line are replaced by a dash. Deletions are marked as a ‘*’.

```
                     v        v                                 v                                  v          vv
M. musculu  CGCGACCUCA  GAUCAGACGU  GGCGACCCGC  UGAAUUUAAG  CAUAUUAGUC  AGCGGAGGAA  AAGAAACUAA  CCAGGAUUCC  CUCAGUAACG  GCGAGUGAAC
C. cidaris  NNN-------  ----G-G--A  -A-C------  ----------  -----C-C-A  ----------  ----------  -G--------  -------G--  ---------G
E. cordatu  NNNNNNN---  ----G-G--A  -A-C------  ----------  -----C-C-A  ----------  ----------  ----------  -------G--  ---------G
S. purpure  NNN-------  ----G-G--A  -A-C------  ----------  -----C-C-A  ----------  ----------  ----------  -------G--  ---------G
B. unicolo  NNNNN-----  ----G-G--A  -A-C------  ----------  -----C-C-A  ----------  ----------  ----------  -------G--  ---------G
E. pusillu  NU--------  ----G-G--A  -A-C------  ----------  -----C-C-A  ----------  ----------  ----------  -CU----G--  ---------G
A. lixula   NU--------  ----G----A  -A-C------  ----------  -----C-C-A  ----------  ----------  ----------  -CU----G--  ---------G
E. esculen  NNN-------  ----G----A  -A-C------  ----------  -------C-A  ----------  ---------   -U--------  -------G--  ---------G
P. lividus  NU--------  ----G----A  -A-C------  ----------  -----C-C-A  ----------  ---------   ----------  -------G--  ---------G
P. miliari  NNN-------  ----G----A  -A-C------  ----------  -----C-C-A  ----------  ----------  ----------  -------G--  ---------G
S. granula  NU-------G  ----G----A  -A-C------  ----------  -----C-C-A  ----------  ----------  ----------  -C-----G--  ---------G
L. variega  -U-------G  ----G----A  -A-C------  ----------  -----C-C-A  ----------  ----------  ----------  -C-----G--  ---------G
A. bifida   NN--------  ----G-G--A  -AAU------  ----------  -------C-A  ----------  ----------  -A--------  ------G--   ---------G
S. papposa  NNNN------  -G--G-G--A  -AG-------  --G-------  -----C-C-A  ----------  ----------  ----------  -C-----G--  -----C---G
A. gibbosa  NU--------  -G--G-G--A  -AG-------  --G-------  -----C-C-A  ----------  ----------  ----------  -C-----G--  -----C---G
E. seposit  NNU-------  -G--G-G--C  -AG-------  C-G-------  -N---C-C-A  ----------  ----------  -A--------  -------G--  -----C---G
H. sanguin  NNN-------  -G--G-G--C  -A-U------  --G-------  -----C-C-A  ----------  ----------  -A--------  -------G--  -----C---G
C. tenuisp  NNNN------  -G--G-G--A  -AG-------  --G-------  -----C-C-A  ----------  ----------  -A--------  -CU----G--  -----C---G
M. glacial  NNNNN-----  -G--G-G--A  -AG-------  --G-------  -----C-C-A  ----------  ----------  -A--------  -CU----G--  -----C---G
A. rubens   NNN-------  -G--G-G--A  -AG-------  --G-------  -----C-C-A  ----------  ----------  -A--------  -CU----G--  -----C---G
A. squamat  NNN-------  -C--GAG---  -A--------  --G-------  -------C-A  ----------  ----------  ----------  -------G--  -N-------G
O. fragili  NNN-------  ----GAG--A  -A-U------  ----------  -----C-C-A  ----------  ----------  ----------  -------G--  -N-------G
H. forskal  NNN-------  ----G-N--A  -AUA------  ----------  -----C-C-A  ----------  ----------  -A--------  -CU----G--  -----C---G
L. inhaere  NNN-------  -C--G-G---  -AUA------  ----------  -----C-C-A  ----------  ----------  -A--------  -------G--  ---------G
```

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

APPENDIX 2 (*Continued*)

```
                 v            v        vvvvv    vvvvvv     vvvv     vv                          vvv     vv      vv  vv
M. musculu  AGGGAAGAGC CCAGCGCCGA AUCCCCGCCG CGCGUCG*** *CGGCGUGGG AAAUGUGGCG UACGGAAGAC CC*ACUCCCC GGCGCCGCUC GUGGGGGGCC
C. cidaris  C--------- ---U---U-- -----*-*GA *C-*GACCG- -G-CUCG--- -GU------- -GA--CG-CU U--GGG-G-G C-A-U-**** *CAC-AC---
E. cordatu  C--------- ---U---U-- -----ACAG* *C-*GACCG- -G-UU-C--- -GU------- -GA--CG-CU U--GGG-G-G C-A-U-**** *CAC-AC---
S. purpure  C--------- ---U---U-- -----ACAG* *C-*GACCG- -G-UU-C--- -GU------  -GA--CG-CU U--GGG-G-G C-A-U-**** *CAC-AC---
B. unicolo  C--------- ---U---U-- -----ACAG* *C-*GACCG- -G-UU-C--- -GU------- -GA--CG-CU U--GGG-G-G C-A-U-**** *CAC-AC---
E. pusillu  C--------- ---U------ ----NAAAGU *C-*GACCG- -GNC-UC--- -GC------- -GA--CG-CU U--GGG-G-U C-A-UU**** *C-C--U---
A. lixula   C----UN--- ---U------ -NNNNNN*GA *C-*GACCG- -G-CUGG--- -CC------- -GA--CG-CU U--GGG-AUG C-A-GU**** *-AC--C---
E. esculen  C--------- ---U------ -----A--G* *C-U--*--- -G-U-C---- -CC------- -GA--CG-CU U--GGG-G-G C-A-U-**** *CAC-AC---
P. lividus  C--------- ---U------ -----A--G* *C-U--*--- -G-U-C---- -CC------- -GA--CG-CU U--GGG-G-G C-A-U-**** *CAC-AC---
P. miliari  C--------- ---U------ -----A--G* *C-U--*--- -G-U-C---- -CC------- -GA--CG-CU U--GGG-G-G C-A-U-**** *CAC-AC---
S. granula  C--------- ---U------ -----A--G* *C-C--A--- -G-CUC---- -GC------- -GA--CG-CU U--GGG-G-G C-A-U-**** *CAC-AC---
L. variega  C--------- ---U------ -----A--G- *C-C--A--- -G-C-C---- -GC------- -GA--CG-CU U--GGG-G-G C-A-U-**** *CAC--C---
A. bifida   C----UC--- ---------- ---U---UG* *C-UGAACG- -GUA--C--- ---------- --A-A---C* ---UACU--G C---UGUG** *CUC--A*-U
S. papposa  C--------- ---------- -------GGC -C--G--AGG -GUC--A--- --C------- --A-----*U ---GA-*--G C-*-U-**** *G-C--C---
A. gibbosa  C--------- ---------- -----N-GNN -C-CC-NCGG -GCCG-A--- --C------- --A-----*U ---GAG*--G C-GC-G**** *G-C--C---
E. seposit  C--------- ---U*----- -------*GC -C-*G-AAGU CU-C--A--- ---------- --A-----*U -G-GUCU--G C-*-U--G-U CGUC--C---
H. sanguin  C--------- ---U------ -------GG- GC**G--AGU --CC--AN-- ---------- --A-----*U ---GU*U-** -C*-U-CG** *G-C--C---
C. tenuisp  C--------- ---------- -------**C -C-*G--AGG -GCC--A*-- ---------- -NA-----*U NN-N**--N- --GC-G**** *GCC--C---
M. glacial  C--------- ---------- -------GGC -CN**--AGG -GCU--A--- --C------- --A-----*U -*-G-C--GA ---C****** *GCA--C---
A. rubens   C--------- ---U------ -------GGC -C-AC--AGG -GCU--A--- --C------- --A-----*U -----C---G ----U***** *GCC--C---
A. squamat  U--------- ---U---UU- ----U---GU -C-***AGUG -GC-U----- -G---C---- -GA--U--** --GUGG-UUA UCG--GA*** *C-C------
O. fragili  C--------- ---U------ ---U----GU **-U**AGCG -GC-U-A-A- ---------- -GA-----** --GUGG-UAA CCG-AGA*** *C-U--A---
H. forskal  G-----A--- -------U-- -------UG* -C-U**-UCG -GU---A--- --C---A--- --A----A** -UG-A-U--G U---UN-*** *C-NNC-U-G
L. inhaere  U--------- ---U---U-- --------AC GC-*G-AAGG UGU---C--- --U------- -GA----A*A -AG-A---UG -C--UUC-** *AA--CU---
```

APPENDIX 2 (*Continued*)

```
                                  v        v v                  vv vvv    v     vv       vv   vvv v
M. musculu  CAAGUCCUUC U*GAUCGAGG CCC******* AGCCCGUGGA CGGUGUGAGG CCGGUAGCGG CCCCC****G GCGCGCCGGG *CUCGGGUCU UCCCG***GA
C. cidaris  G-G-----C- -------G-- -G-CU----C -A---AGA-C G-----C-*C -N---UCG-- NGU-G----- -UCUCG---* --CU--AG-C A--AA-----
E. cordatu  G-G-----C- -------G-- -A-CU----- UA---AGA-C G-----C-** *-C--UCG-- -GU-G----- -UCUCG---- --CU--AG-C A--AA-----
S. purpure  G-G-----C- -------G-- -A-CU----- UA---AGA-C G-----C--* N-C--UCG-- -GU-G----- -UCUCG---- --CU--AG-C A--AA-----
B. unicolo  G-G-----C- -------G-- -A-CU----- UA---AGA-C G-----C--* -GC--UCG-- -GU-G----- -UCUCG---- --CU--AG-C A--AA-----
E. pusillu  G-G-----C- -------G-- -A-CU----- UA---AGA-C G-----C--- --C--UCA-- -G--G----- AACUGG---- --CU--AG-C A--AA-----
A. lixula   G-G-----C- -------G-- -A-CU----- UA---AGA-C G-----C--- --C--UCA-- ----G----- A-CUCG-A-- --CU--AG-C A--AA-----
E. esculen  G-G-----C- -------G-- -A-CU----- UA---AGA-C G-----C--* -GC--UCG-- -GU-G----- -UCUCAA--- --CU--AG-C A--AA-----
P. lividus  G-G-----C- -------G-- -A-CU----- UA---AGA-C G-----C--* -GC--UCG-- -GU-G----- -UCUCAA--- --CU--AG-C A--AA-----
P. miliari  G-G-----C- -------G-- -A-CU----- UA---AGA-C G-----C-** -CC--UCG-- -GU-G----- -UCUCAA--- --CU--AG-C A--AA-----
S. granula  G-G-----C- -------G-- -A-CU----- UA---AGA-C G-----C--* -GC--UCG-- -G--G----- -ACUCAA--- --CU--AG-C A--AA-----
L. variega  G-G-----C- -------G-- -A-CU----- UA---AGA-C G-----C--- --C--UCG-- -G--G----- -ACUCAA--- --CU--AG-C A--AA-----
A. bifida   --C------- ---------- --*------- UU---A-A-- G-----C--* --CC**CA-- --UGG---GC ---U--G-A* -**-CAC*GG -UUUCAG---
S. papposa  G-G------- ---------- ---CC-AUCG U*---AA--C G-----C--- --**-GCG-- --GG*---UC CG--NG-NNN -**-CA*GGG CNU-CGG---
A. gibbosa  G-G------- ---------- ---CC-UCCG U*---A---C G-----C--- --**-GCG-- -G-*G---GC CG--CG---A -**-CA*GGG CUU-CGG---
E. seposit  G-G------- ---------- --UCCGCGCG U*---AG--C G-----C--- --**-UCG-- --GG*---UC CG--CG---A AU----A-GG CUU-CGG---
H. sanguin  G-G------- ---------- --UCU-U-CG U*---AC--C G-----C--- --**-GCG-- -G--G---UC CG-UCG---A A**-CAA-GG CUU-CAG---
C. tenuisp  G-G------- ---------- NNUGC-UCCG U*---AG--N G-----C-NN G-CAGGCGNN *G--G---GN NGN*CG-NNN -**-CAAGGG CUU-CGG---
M. glacial  G-G------- ---------- -GUGC-UCCG U*---AG--C G-----C--- --CAGGCG-N NGNN*---GC CG-NCU-NNA -**-CA*GGG CUU-CGG---
A. rubens   G-G------- -----U---- -GUGC-UCCG U*---AG--C G-----C--- --*AGGCG-- -G--A---GC CG-UCG---A -**-*AAGGG C*U-CGG---
A. squamat  G-C------G ----AG---- -U*-U----A U*---ACA-- GU----U--- --CCC-CA-- -U---GUU-C CG-U-G-AU* -**-*A**GG CUA-UAA---
O. fragili  G-C------G ----AG---- -U*-U----A U*---A-A-- GU----U--- --CCCGCA-- -*G--GUUAC CG-U-G-AU* -**-*A**GG CUU-UAG---
H. forskal  GCU-----C- -U--GU-G-- -U*------- ***----A-C G--N--A--- --C---CG-N GGUGG---CU CG-NNGGAA* -**G-AU**G CUUUCCAA--
L. inhaere  G-G----CCA ----C--G-- -U-CU----- U**--NNNNC G-----A-N- --C--UCA-- --AGG---G* CG-UCGG-N* -**GUAU**G CUUUCCAA--
```

APPENDIX 2 (*Continued*)

```
               v   v                v   v       v                                                                  v   v
M. musculu  GUCGGGUUGC  UUGGGAAUGC  AGCCCAAAG*  CGGGUGGUAA  ACUCCAUCUA  AGGCUAAAUA  CC*GGCACGA  GACCGAUAGU  CA*ACAAGUA  CCGUAAGGGA  AAGU      :
C. cidaris  ---------U  --U-------  ---------A  G-------G-  --------C-  ----------  -U--------  --------UC  GGACA*----  ----G-----  ----      :
E. cordatu  -------A-U  --U-------  U--------A  --------G-  --------C-  ----------  -U--------  ---------C  AGACA*----  ----G-----  ----      :
S. purpure  -------A-U  --U-------  U--------A  --------G-  --------C-  ----------  -U--------  ---------C  AGACA*----  ----N-----            :
B. unicolo  -------A-U  --U-------  U--------A  --------G-  --------C-  ----------  -U--------  ---------C  AGACA*-N--  ----N-                :
E. pusillu  -------A-U  --U-------  U--------A  G-------G-  --------C-  ----------  -U--------  ---------C  AGACA*-:
A. lixula   ---A---A-U  --U-------  U---U----A  G-------G-  --------C-  ----------  -U--------  ---------C  AGACA*----  ----G-----  --        :
E. esculen  -------A-U  --U-------  U--------A  A-------G-  --------C-  ----------  -U--------  ---------C  AGACA*----  ----G-----  ----      :
P. lividus  -------A-U  --U-------  U--------A  A-------G-  --------C-  ----------  -U--------  ---------C  AGACA*----  ----G-----  ----      :
P. miliari  -------A-U  --U-------  U--------A  A-------G-  --------C-  ----------  -U--------  ---------C  AGACA*----  ----G-----  ----      :
S. granula  -------A-U  --U-------  U--------A  G-------G-  --------C-  ----------  -U-------   ---------C  AGACA*----  ----G-----  ----      :
L. variega  -------A-U  --U-------  U--------A  G-------G-  --------C-  ----------  -U--------  ---------C  AGACA*---                          :
A. bifida   ---------U  --U-------  ---------A  ----------  --------C-  ----------  -U--------  -U-------C  G-ACA*----  ----G-----  ----      :
S. papposa  ---------U  --U-U-GC--  ---------G  ----NN----  --------C-  ----------  ---U------  ---------A  GCACA*----  ----G-----  ----      :
A. gibbosa  ---------U  --U-U-GC--  ---------G  ----------  -------C-   ----------  ---U------  ---------A  GCACA*----  ----G-----  ----      :
E. seposit  ---------U  -CU-U-U---  -----C---G  -----N----  --------C-  ----------  ----------  ---------C  G-ACA*----  ---CG-----  ----      :
H. sanguin  ---------U  -CU-U-GC--  --------G   -----N----  --------C-  ----NN----  ----------  ---------A  GCACA*----  ----G-----  ----      :
C. tenuisp  ---------U  --U-U-GC--  ----N----G  ----NN----  --------C-  ----NN----  --UU------  ----NN---A  GCACA*----  ----G:
M. glacial  ---------U  --U-U-GC--  ---------G  ----NN----  --------C-  ----------  --UU------  ---------A  GCACA*----  ----G-NN--  ----      :
A. rubens   ---------U  --U-U-GC--  ---------A  A---------  --------C-  ----------  --UU------  ---------A  GCACA*---   ----G-----  ----      :
A. squamat  ---------U  --U-------  ---------A  ----------  --------C-  ----------  -UAA------  ---------A  -GACA*----  --A-G-----  ----      :
O. fragili  ---------U  --UA--U---  -----U---A  ----------  --------C-  ----------  --CA------  ----------  UCACA*----  --A-G-----  ----      :
H. forskal  ---------U  --U-------  ----NN---A  ----NN--G-  -U------C-  ----------  UUAA------  ---------A  -GACA*----  ----G-----  -         :
L. inhaere  ---------U  --UA------  ----N----A  G--NNN--G-  -----N--C-  ----------  UUAA------  ----------  AGACA*----  ----G-----  -         :
```

APPENDIX 3

Divergence dates based on Harland *et al.* (1989), rounded to nearest 5 Ma.

I. Cidaroid–Euechinoid divergence. Taken as earliest occurrence of definitive cidaroid: *Miocidaris keyserlingi* (Döderlein), Ford Reef Formation, Kazanian, Upper Permian, of Sunderland, northern England (Smith & Hollingworth 1990). 255 Ma.

II. Irregularia–Echinacea divergence. Taken as earliest occurrence of definitive irregular '*Plesiechinus*' *hawkinsi* Jesionek–Szymanska, Sunrise Formation, Sinemurian, Lower Jurassic, of New York Canyon, Nevada, U.S.A. (Jesionek-Szymanska 1970). 205 Ma.

III. Spatangoid divergence. The three spatangoids *Echinocardium*, *Spatangus* and *Brissus* are all closely related genera that appear to have diverged during the Eocene. *Echinocardium* is a loveniid and the earliest confirmed occurrence of a loveniid *Lovenia sulcata* (Haime), Upper Eocene, Biarritz, France (Mortensen, 1951, in 1928–1951; BMNH collections). The earliest confirmed record of *Brissus* is from the Middle–Late Eocene of Cuba (Kier 1984). The earliest confirmed spatangid is Oligocene, but *Maretia*, from which *Spatangus* may be derived through loss of peripetalous fasciole, occurs from the Middle Eocene onwards. All three taxa appear to have diverged between 40 and 45 Ma.

IV. Stirodont–Camarodont divergence. The oldest definitive Arbaciidae is *Glypticus* from the late Callovian, Middle Jurassic (155 Ma). The origin of camarodonts remains obscure but no camarodont is known prior to the Lower Cretaceous. Arbaciids are probably derived from hemicidarids and thus divergence from echinaceans could be as far back as 175 Ma. Here we take the first appearance of true Arbaciids as a divergence date.

V. Echinidae–Echinometridae divergence. 50 Ma is the latest divergence date possible because both families were differentiated by the middle Eocene (Smith 1989). However, *Scoliechinus* may be late Cretaceous (S. K. Donovan, personal communication) and would take the divergence back to around 70 Ma. Here an intermediate divergence date of 60 Ma is adopted.

The *Echinus–Paracentrotus–Psammechinus* divergence. The fossil record of *Paracentrotus* and *Echinus* extends back only to the late Miocene and thus the latest divergence is placed at 15 Ma. *Psammechinus* may be older but recognition of this genus has proved impossible from test material alone (Smith 1989). Earliest divergence within this tricotomy is very poorly constrained at between 15 and 50 Ma and is arbitrarily placed at 30 Ma.

The *Lytechinus–Sphaerechinus* divergence. The earliest *Lytechinus* is *L. pictus*? from the Upper Oligocene of Oregon, U.S.A. Divergence is thus placed at 25 Ma.