PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

A Combined Morphological and Molecular Phylogeny for Sea Urchins (Echinoidea: Echinodermata)

D. T. J. Littlewood and A. B. Smith

Phil. Trans. R. Soc. Lond. B 1995 **347**, 213-234
doi: 10.1098/rstb.1995.0023

A combined morphological and molecular phylogeny for sea urchins (Echinoidea: Echinodermata)*

D. T. J. LITTLEWOOD AND A. B. SMITH

Department of Palaeontology, The Natural History Museum, Cromwell Road, London SW7 5BD, U.K.

SUMMARY

Phylogenetic relationships of higher taxa of echinoids have been investigated using a 163 character morphological data base and molecular sequences from large and small subunit (LSU and SSU) ribosomal RNA (rRNA) genes. The complete SSU rRNA gene has been sequenced for 21 taxa, with representatives from nine of the 14 extant orders of Echinoidea. Partial LSU sequences, representing the first 400 base pairs (b.p.) from the 5′ end were also sequenced for three taxa to complement an existing data base of ten taxa. The two molecular sequences provided a total of 371 variable sites, of which 143 were phylogenetically informative (compared to 145 phylogenetically informative sites from morphological data). Morphological, LSU and SSU data have been analysed separately and together.

Morphological and SSU sequence data generate topologies that are not significantly in conflict (under Templeton's test), but the strong signal pairing arbaciids with clypeasteroids in the LSU derived tree marks the LSU sequence data as anomalous for this taxon. A 'total evidence' approach derived a tree very similar in topology to that derived from morphological data. Rooted on the stem group echinoid *Archaeocidaris*, our total evidence tree suggested relationships of higher taxa as follows: Cidaroida (Phormosomatidae (Echinothuriidae (Diadematidae ((Spatangoida (Clypeasteroida, Cassiduloida)) ((Calycina, Arbacioida) (Stomopneustidae (Glyphocidaridae (Temnopleuridae (Echinometridae (Echinidae, Strongylocentridae))))))))))). Phylogenetic analyses run both with and without key fossil taxa yielded slightly different topologies. It is important to include fossil taxa in a phylogenetic analysis where there are long stem-group branches or where the crown group is highly derived.

1. INTRODUCTION

Echinoids are the best known class of Echinodermata, with just over 900 extant species and around 10000 fossil species described. Their complex endoskeleton provides a wealth of phylogenetically informative characters for systematists, and the taxonomy of the group has been based almost entirely upon hard-part anatomy. Consequently, both fossil and extant taxa can and have been classified on the same suite of skeletal characters. Furthermore, since the more robust parts of the echinoid skeleton also fossilize readily, the geological record of the group is reasonably well known (e.g. Durham *et al.* 1966; Kier 1974; Smith 1984).

Although the broad outlines of the interrelationships of major echinoid groups seems reasonably secure, there remain a number of areas of uncertainty. For example, for a long time irregular echinoids were thought to be polyphyletic (Durham & Melville 1957), but more recent analysis has suggested that they are in fact a monophyletic group (Jensen 1981; Smith 1981). The relationships of camarodonts (a monophyletic group) to various stirodont taxa (possibly a paraphyletic assemblage) also remains uncertain. Two early cladistic analyses of class relationships (Jensen 1981; Smith 1981) have been carried out but neither was rigorous by today's standards. Smith based his analysis on a combination of recent and fossil taxa and gave synapomorphies for branch points, but never tested whether the preferred tree was in fact the most parsimonious. Jensen's analysis identified autapomorphies for major extant groups, based largely on tooth and lantern structure, but almost entirely neglected to document synapomorphies linking these major clades. In recent years, more rigorous cladistic analyses of adult morphology have been carried out within specific echinoid clades (clypeasteroids: Mooi 1987, 1990; cassiduloids: Suter 1994*a*, *b*; holasteroids: David 1987; camarodonts: Smith 1989; cidaroids: Smith & Wright 1990; diadematoids, echinothurioids and Calycina: Smith & Wright 1993). Echinoid larvae also have a complex morphology which has recently been placed into a phylogenetic context (Wray 1992). However, the larval data have simply been optimized onto a tree derived from an analysis of adult morphology rather than being treated as an independent set of phylogenetically informative characters.

In addition to adult and larval morphology, molecular data have also been gathered and applied to the question of echinoid phylogenetic relationships.

* Sequence data presented herein have been deposited with the EMBL database under Accession numbers: Z37116–Z37135, Z37140–Z37149, Z37514, Z37507.

Matsuoka (1980, 1985, 1986, 1987, 1989, 1990) and Matsuoka & Suzuki (1987) have published a series of immunological and allozyme studies of camarodont taxa from which he has calculated genetic distances and inferred phylogenetic relationships. Other molecular phylogenies have also recently started to appear, notably the DNA/DNA hybridization analysis of clypeasteroid relationships by Marshall & Swift (1992), and the analysis of partial nucleic acid sequences of large subunit ribosomal RNA (LSU rRNA) by Féral & Demelle (1991), Smith *et al.* (1992) and Féral *et al.* (1994). Suzuki & Yoshino (1992) and Suzuki *et al.* (1988) have identified the amino acid sequence for the sperm-binding protein in a large number of echinoids and have pointed out the phylogenetic significance of their data. However, none of these molecular studies has tackled the question of how the major lines of echinoids are related. It is this question that we try to answer here, using a combination of morphological and molecular data.

2. METHODOLOGY AND DATA

We have compiled data from three sources; adult morphology, 18*S*-like small subunit ribosomal RNA (SSU rRNA) and partial 28*S*-like large subunit ribosomal RNA (LSU rRNA). We initially planned to build up fully complementary molecular data sets, but technical difficulties meant that we were not always successful in obtaining both SSU and LSU rRNA sequences from the same species. Although taxa were selected to give a broad coverage of extant echinoid orders, not all extant orders have been sampled. Some, such as the Micropygoida, Pedinoida and Holectypoida are represented by just one or two extant species which we did not manage to obtain for our molecular studies. Others, such as the Calycina, are deep-water and, although we obtained frozen material of representative species, we were unable to amplify and sequence gene products from this material. However, we were eventually able to compile both morphological and molecular data for 27 echinoid taxa, representing nine of the 14 extant orders of echinoids recognized in Smith (1984). This study presents the most complete and thorough molecular and morphological data sets available to date.

(*a*) *Morphological data*

A data matrix was compiled for all extant echinoids for which we had molecular sequence data, twenty seven in total. A further five extant taxa were also included in the morphological analysis as representatives of high-rank clades for which we had no molecular sequence data. Finally 13 fossil taxa were also added to the analysis as basal representatives of potential long branches in our tree. Taxa were selected because of their presumed phylogenetic position, close to the base of major extant clades, and their relative anatomical completeness. For example, the pygasteroid *Plesiechinus* was included as one of the earliest and least specialized of the irregular echinoids (Irregularia), and one of the few of this group whose tooth and lantern structure are known. The lantern of *Plesiechinus* is considerably less derived in structure than that of any extant irrregular echinoid (indeed, many have completely lost their lantern as adults) and the inclusion of *Plesiechinus* allowed us to establish the plesiomorphic state for many lantern characteristics within this clade. Similarly the fossil pedinoid *Diademopsis* was included since it is anatomically well known and has been identified as the fossil taxon lying closest to the divergence between aulodonts, stirodonts and irregular echinoids (Smith 1981). Taxa are listed in table 1 together with their current classification.

In total 163 characters were gathered from specimens and from published sources, the principal of which are Mortensen (1928–1951), Hyman (1955), Jensen (1979, 1981), Smith (1981, 1984, 1988), Smith & Wright (1989, 1990, 1993). The great majority of these relate to skeletal attributes. However, some characters, notably those pertaining to pedicellarial structure and lantern structure, are rarely preserved in fossil species and often had to be scored as unknown.

In choosing our characters we attempted to include all morphological traits that had been identified previously as having taxonomic significance (see references above). Thus any character that occurred in two or more states among the taxa listed was included. The list is certainly not exhaustive, but we hope it provides a reasonably unbiased subset of the morphological differences that exist among these echinoids.

For our cladistic analysis we used the advanced stem group echinoid *Archaeocidaris whatleyensis* Lewis & Ensom, from the Carboniferous of southern England (Lewis & Ensom 1982) as our outgroup. It is undoubtedly the best known of any Palaeozoic echinoid and clearly represents an outgroup to all extant echinoids, because it has multiple columns of interambulacral plates instead of the two columns found in all crown group echinoids.

The characters used in our analysis are listed in Appendix 1, together with the data matrix. We carried out parsimony analyses on the data matrix using the computer program PAUP (Macintosh version 3.1.1, Swofford 1993). All characters were treated as unordered and of equal weight. For small data sets we used the 'exhaustive search' or 'branch and bound' options. However, for large data sets this proved too time-consuming and we employed the 'heuristic' search option.

(*b*) *Molecular data*

Genes isolated and sequenced by us were determined directly from polymerase chain reaction (PCR)-amplified and cloned DNA. Sequences incorporated from previous studies were generally determined from the gene products, i.e. the rRNA, and were accessed from published material or electronic databases as indicated below.

(i) *DNA isolation and purification*

Few fresh specimens were available for analysis and most tissues or whole echinoids were either alcohol preserved (fixed in absolute ethanol, shipped and

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Table 1. *Echinoids used in this analysis*

(! = extinct; [SSU] = SSU rRNA sequence available; [LSU] = LSU rRNA sequence available. [1] = this study; [2] = published sequence from Smith *et al.* (1992), data derived from the laboratory of R. Christen; [3] = published sequence from Wada & Satoh (1994); [4] = published sequences from Raff *et al.* (1987).

!Family Archaeocidaridae (advanced stem group member)
!*Archaeocidaris whatleyensis* Lewis & Ensom, 1982
Subclass Cidaroidea Claus, 1880
!Family Miocidaridae
!*Miocidaris keyserlingi* (Geinitz, 1848)
Order Cidaroidea Clause, 1880
Family Cidaridae Gray, 1825
Cidaris cidaris Linnaeus, 1758 [LSU^2]
Eucidaris tribuloides (Lamarck, 1816) [SSU^1]
Subclass Euechinoidea Bronn, 1860
Order Echinothurioida Claus, 1880
Family Phormosomatidae Mortensen, 1934
Phormosoma placenta Thomson, 1872
Family Echinothuriidae Thomson, 1872
Araeosoma ijimai Yoshiwara, 1897 [LSU^1]
Asthenosoma owstoni Mortensen, 1904 [SSU^1, LSU^1]
Order Diadematoida Duncan, 1889
Family Diadematidae Peters, 1855
Diadema setosum (Leske, 1778) [SSU^1]
Centrostephanus coronatus (Verrill, 1867) [SSU^1]
Order Pedinoida Mortensen, 1939
!*Diademopsis* ex. gr. *serialis* (Agassiz, 1840)
Cohort Echinacea Claus 1876
!Family Pseudodiadematidae Pomel, 1883
!*Trochotiara* spp.
Superorder Stirodonta Jackson, 1912
Order Calycina Gregory, 1900
!Family Acrosaleniidae Gregory, 1900
!*Acrosalenia lycetti* Wright 1855
!*Acrosalenia spinosa* Agassiz, 1840
Family Saleniidae Agassiz, 1838
Salenia goessiana Loven, 1874
Order Arbacioida Gregory, 1900
!*Gymnocidaris pustulata* (Forbes, 1849)
!*Glypticus hieroglyphicus* (Goldfuss, 1826)
Family Arbaciidae Gray, 1855
Arbacia lixula (Linnaeus, 1758) [SSU^1, LSU^2]
Order Phymosomatoida Mortensen, 1904
Family Stomopneustidae Mortensen, 1903
Stomopneustes variolaris (Lamarck, 1816) [SSU^1]
Family Glyptocidaridae Jensen, 1981
Glyptocidaris crenularis Agassiz, 1853
Superorder Camarodonta Jackson, 1912
!*Glyptocyphus difficilis* (Agassiz, 1846)
Order Temnopleuroida Mortensen, 1941
Family Temnopleuridae Agassiz, 1872
Temnopleurus hardwickii (Gray, 1855) [SSU^1]
Mespilia globulus (Linnaeus, 1758) [SSU^1]
Salmacis sphaeroides (Linnaeus, 1758) [SSU^1]
Order Echinoida Claus, 1876
Family Echinidae Gray, 1825
Echinus esculentus Linnaeus, 1758 [SSU^1, LSU^2]
Paracentrotus lividus (Lamarck, 1816) [LSU^2]
Psammechinus miliaris (Müller, 1771) [SSU^1, LSU^2]
Family Echinometridae Gray
Subfamily Strongylocentrotidae Gregory, 1900
Strongylocentrotus intermedius Agassiz, 1863 [SSU^3]
Subfamily Echinometridae Gray, 1825
Colobocentrotus atratus (Linnaeus, 1758) [SSU^1]
Heliocidaris erythrogramma (Valenciennes, 1846) [SSU^4]
Heliocidaris tuberculata (Lamarck, 1816) [SSU^4]
Subfamily Toxopneustidae Troschel, 1872
Sphaerechinus granularis (Lamarck, 1816) [SSU^1, LSU^2]
Tripneustes gratilla (Linnaeus, 1758) [SSU^1]
Lytechinus variegatus (Lamarck, 1816) [LSU^2]
Cohort Irregularia
!*Eodiadema minuta* (Buckman, 1845)
!Family Pygasteridae Lambert, 1900
!*Plesiechinus ornatus* (Buckman, 1845)
Order Cassiduloida Claus, 1880
Family Cassidulidae Agassiz & Desor, 1847
Cassidulus mitis Krau, 1954 [SSU^1]
Family Echinolampadidae Gray, 1851
Echinolampas crassa Bell, 1880
Order Clypeasteroida Agassiz, 1872
Suborder Clypeasterina Agassiz, 1872
Family Arachnoididae Duncan, 1889
Fellaster zelandiae (Gray, 1855) [SSU^1]
Suborder Scutellina Haekel, 1896
Family Fibulariidae Gray, 1825
Echinocyamus pusillus (Müller, 1776) [LSU^2]
Family Astriclypeidae Stefanini, 1911
Echinodiscus bisperforatus Leske, 1778 [SSU^1]
Family Mellitidae Stefanini, 1911
Encope aberrans Martens, 1867 [SSU^1, LSU^1]
Order Spatangoida Claus, 1876
Family Spatangidae Gray, 1825
Spatangus purpureus Müller, 1776 [LSU^2]
Family Loveniidae Lambert, 1905
Echinocardium cordatum (Pennant, 1777) [SSU^1, LSU^2]
Family Brissidae Gray, 1855
Meoma ventricosa (Lamarck, 1816) [SSU^1]
Brissopsis lyrifera (Forbes, 1841) [SSU^1]

stored in more than 80% ethanol) or frozen. Table 2 lists the method of preservation, type of tissue and method of tissue extraction used for each echinoid species sequenced for this study. For our purposes the tissues of choice were either gonad material or muscle tissue (usually from the lantern apparatus). Alcohol preserved tissues were always soaked in 2–3 washes of 10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA and 1–2 washes of dH_20 for 2–3 h prior to use.

Approximately 0.01–0.1 g tissue were used for DNA extractions depending on the source of material. DNA was extracted by a variety of methods:

1. CTAB extraction method: adapted from Doyle and Doyle (1987) by M. Black, Rutgers University, New Jersey with modifications described in Littlewood (1994). This was the most successful method of extracting high molecular mass DNA.
2. Tissues lysed by incubation at 37 °C in 50 mM Tris-HCl (pH 8.0), 50 mM EDTA, 100 mM NaCl, 0.5% SDS, 200–500 μg ml^{-1} Proteinase K (Boehringer) for 2–5 h and DNA purified by standard phenol extraction, ethanol precipitation techniques (Sambrook *et al.* 1989).
3. Guanidium thiocyanate method (adapted from Boom *et al.* 1990): 5–10 volumes GE reagent (5 M Guanidium thiocyanate, 0.1 M EDTA) added to tissue. Tissue homogenized; equal volume ice cold 7.5 M ammonium acetate added; left on ice 15 min; extract three times with chloroform/isoamyl alcohol (standard

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Table 2. Tissues selected, method of preservation/storage and method of DNA extraction for echinoids sequenced in this study

(CTAB = CTAB, prot.K = proteinase K, GE = Guanidium thiocyanate, bead = bead beaten, gDNA = genomic DNA supplied by S. Palumbi. Full details given in the text.)

species	tissue	storage	extraction
Araeosoma ijimai	sperm	frozen	CTAB
Asthenosoma owstoni	sperm	frozen	CTAB
Arbacia lixula	gonad	alcohol	CTAB
Brissopsis lyrifera	gonad	fresh	prot.K
Cassidulus mitis	gonad	alcohol	GE
Centrostephanus coronatus	muscle	alcohol	bead/GE
Colobocentrotus atratus	gonad	fresh	gDNA
Diadema setosum	gonad	alcohol	CTAB
Echinocardium cordatum	gonad	alcohol	CTAB
Echinodiscus bisperforatus	gonad	alcohol	CTAB
Echinus esculentus	gonad	alcohol	CTAB
Encope aberrans	gonad	alcohol	prot.K
Eucidaris tribuloides	gonad	alcohol	CTAB
Fellaster zelandiae	gonad	alcohol	prot.K
Meoma ventricosa	gonad	alcohol	CTAB
Mespilia globulus	gonad	alcohol	CTAB
Psammechinus miliaris	gonad	alcohol	CTAB
Salmacis sphaeroides	gonad	alcohol	CTAB
Sphaerechinus granularis	gonad	alcohol	CTAB
Stomopneustes variolaris	gonad	alcohol	CTAB
Temnopleurus hardwickii	gonad	alcohol	CTAB
Tripneustes gratilla	gonad	fresh	gDNA

phenol extraction also used in addition to this method on some occasions); DNA precipitated with equal volume of isopropanol; washed in 70% ethanol; pellet dried and resuspended in dH_20 or TE.

4. Bead beaten: approximately 250 µl tissue sample in TE; equal volume of GE reagent added and 2 volumes of glass beads (0.17–0.18 mm diameter, B. Braun Melsungen AG); beaten twice at 2000 r.p.m. for 1 min with Mikro-Dismembrator U (B. Braun Melsungen AG); spun at 10000 r.p.m. for 5 min; DNA extracted and precipitated from supernatant as in (3).

(ii) *Gene amplification and isolation*

SSU *rRNA gene*. Approximately 1800 b.p. of the nuclear 18*S*-like SSU rRNA gene were amplified by 'hot start' PCR (Hosta & Flick 1991) using universal primers from Medlin *et al.* (1988) or those from Embley *et al.* (1992).

LSU *rRNA gene*. Approximately 1100 b.p. from the 5′ end of the nuclear 28*S*-like rRNA gene were amplified by 'hot start' PCR (Hosta & Flick 1991) using either primers E5, 5′-CGACCTCRGATCGGRCGAGAC-CAC and LSUD4, 5′-tagaagctTCCTGAGGGAAAC-TTCGG, or LSU5, 5′-taggtcgACCCGCTGAAYTT-AAGCA and LSUD4 where lower case letters indicate polylinker sites for directional cloning. E5 was designed to be echinoid specific based on aligned data in Smith *et al.* (1992). LSU5 and LSU3 were designed to be eukaryote specific on the basis of a previously published alignment (Lenaers *et al.* 1989) with the polylinkers based on Medlin *et al.* (1988); further details in Littlewood (1994). LSUD4 and LSU3 are identical except LSUD4 has no polylinker sites.

DNA amplifications were performed in a programmable thermal cycler (Hybaid) with reaction volumes of 25–100 µl (final concentrations: 67 mM Tris-HCl (pH 8.8), 2 mM $MgCl_2$, 1–2 µM each primer, 125–200 µM each dNTP and 1–2 U *Taq* polymerase (Promega). PCR buffer supplied with the *Taq* was also variously used. We found that some templates yielded more amplified product through the addition of 2–5% formamide (derived from Reysenbach *et al.* 1992). Less than 100 ng of template DNA was required for successful amplification. For each gene the PCR program involved 10 cycles of 1 min denaturation at 94 °C, 30 s annealing at 55 °C, and 2 min extension at 72 °C followed by 20 cycles of the same except with the denaturation temperature reduced to 92 °C. The final cycle incorporated a 5 min extension step at 72 °C before cooling to room temperature. Some templates were amplified at lower annealing temperatures (50 °C).

(iii) *Cloning*

PCR products were electrophoresed on 1% agarose/TAE gels, excised and purified with GeneClean (Bio101 Inc.) or QIAEX (QIAGEN Inc.). Some products were digested with SalI and HindIII restriction enzymes prior to sticky-ended directional cloning into the vector pGEM3Zf(−) (Promega). Other products were cloned directly into pGEM-T (Promega). Recombinants were identified by restriction analysis of miniprep DNA. Each recombinant vector was grown up in the JM109 strain of *Escherichia coli*, and purified using column purification kits (Magic MiniPreps, Promega; midi- and maxi-prep columns, QIAGEN Inc.).

(iv) *Sequencing*

Double stranded plasmid DNA was alkaline denatured and sequenced using the Sanger dideoxy-sequencing method (Sanger *et al.* 1977) (Sequenase v. 2.0, USB) with 10% dimethyl sulphoxide added to all stages of the reaction (Winship 1989). Plasmid primers (T7 and SP6) and internal sequencing primers, listed in table 3, were used to sequence the partial LSU rRNA and complete SSU rRNA gene inserts in both directions.

Consensus sequences for each species were assembled using AssemblyLIGN (IBI, Inc.). The entire length of the SSU rRNA gene was sequenced for a total of 21 taxa (deposited with EMBL under accession numbers Z37118–Z37135, Z37140–Z37149, Z37514). Five additional echinoid sequences were available: four (*Heliocidaris erythrogramma*, *H. tuberculata*, *Strongylocentrotus purpuratus and Lytechinus pictus*) had previously been compiled by Raff *et al.* (1988) and are deposited with GenBank (IntelliGenetics, Mountain View, California) under accession numbers M20071–M20073, M20074–M20076, M20117–M20119 and

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Table 3. *Internal sequencing primers for* SSU *and* LSU *rDNA amplification products*

(listed in the order they appear when aligned to a SSU or LSU sequence; F = forward, R = reverse complement of F primer; all primers 5–3 listed for forward primers. Not all primers were needed to sequence the complete plasmid insert. IUPAC codes used throughout.)

SSU		LSU	
325F/R	CCGGAGARGGAGCCTGA	LSU50F	AGCGGAGGAAAGAAAC
557F/R	GCCAGCMGCCGCGGT	D1F/R	TTGYTTGGGAATGCAGC
S-700F	GAGTGTCAAAGCAG	D2F/R	CTTTGAAGAGAGAGTTC
892F/R	CAGAGGAGAAATTCT	D4aF/R	CCCGTCTTGAAACACGG
1125F/R	GAAACTYAAAGGAAT	D3F/R	CCGAAAGATGGTGAACT
1262F/R	GGTGGTGCATGGCCG		
1510F	CAGGTCTGTGATGCCC		
1705F/R	GYAACACACCGCCCGT		

M20089–M20091 respectively. However, these are only partial sequences and omit the two most phylogenetically informative regions at either end of the gene. Consequently they were of little use to our present study and were omitted from all analyses presented here. The other, more complete sequence is of *Strongylocentrotus intermedius*, published by Wada & Satoh (1994), GenBank accession number D14366, included in our analyses. In total there were 279 variable sites among the 22 taxa.

Only three partial LSU gene sequences were successfully determined from the many taxa tried (deposited with EMBL under accession numbers Z37116–Z37117, Z37507). These sequences complemented data of the first 400 b.p. obtained by direct RNA sequencing from R. Christen's laboratory (published in Smith *et al.* 1992). In total 13 taxa were included in the analysis. Aligned sequences have been deposited with EMBL under accession number DS19161. In total there were 92 variable sites among the 13 taxa (shown in Appendix 3).

Molecular sequences were aligned by eye, using the computer program VSM (Christen 1993), working from the highly conserved regions and progressively adding more divergent regions. In fact, alignment posed virtually no problem; both the SSU and partial LSU rRNA sequences could be aligned with little ambiguity. Some minor ambiguity arose in the most highly divergent domain of the SSU rRNA gene (especially positions 685–695) because a few sequences had significant deletions. However, alternate alignment in this region had minimal effect on the phylogenetic results and the full sequence was used in all analyses.

Phylogenetic trees were constructed using both parsimony (PAUP version 3.1.1, Swofford 1993) and maximum-likelihood (DNAML from the PHYLIP, Felsenstein 1993). For the parsimony analyses all characters were treated as equal weight and missing bases were scored as a 'fifth base'. Searches were made with the 'branch and bound' option for small data sets and with the 'heuristic' option for larger data sets. Bootstrap replicates were carried out to establish the robustness of topologies (1000 for small data sets and 250 for larger sets). Maximum-likelihood analyses were performed without global rearrangement, with *Eucidaris tribuloides* treated as the outgroup taxon and with the transition/transversion ratio set at the default (2).

3. RESULTS

(*a*) *Morphological data*

Using the data matrix presented in Appendix 1 we carried out several analyses, differing in the number and kinds of taxa included.

(i) *Full data*

This includes key Recent and fossil taxa for which no molecular data have yet been gathered. Twelve equally parsimonious trees; length = 299 steps, consistency index after removal of autapomorphies (*CI*) of 0.64, and a retention index (*RI*) of 0.89 were found. The semi-strict consensus tree derived from these 12 trees contains three trichotomies, affecting the relative position of *Pseudodiadema* with respect to other stirodont taxa, and the relative positions of the three Echinidae (*Echinus*, *Psammechinus*, and *Paracentrotus*) (figure 1*a*). Relationships among the three temnopleurids were unresolved, but as a clade temnopleurids were identified as sister group to other camarodonts. Rooted on *Archaeocidaris*, cidaroids were placed as sister group to other echinoids (Euechinoidea), while echinothurioids appeared as a paraphyletic grouping. Irregular echinoids are placed as sister group to the combined stirodont+camarodont clade.

Bootstrap replication ($n = 250$) lent strong support (greater than 60%) to a diadematoid clade, a camarodont plus stirodont clade, a clypeasteroid clade, a cassiduloid clade, a spatangoid clade, a camarodont clade and an echinometrid clade. On the whole, high bootstrap values distinguished all of the major clades of irregular echinoids. Irregular echinoids were paired with the stirodont plus camarodont clade by a relatively low bootstrap value (54%), with diadematoids as outgroup. Other relationships were only poorly supported.

(ii) *Recent taxa only*

Thirty three extant were used; a parsimony analysis found 24 trees of minimal length (277 steps); $CI = 0.68$ and $RI = 0.89$. This basically had the same topology as above, but placed echinids and temnopleurids as sister groups at the base of the camarodonts, and placed *Salenia* and *Arbacia* as successive branches leading towards the camarodont clade rather than grouping

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

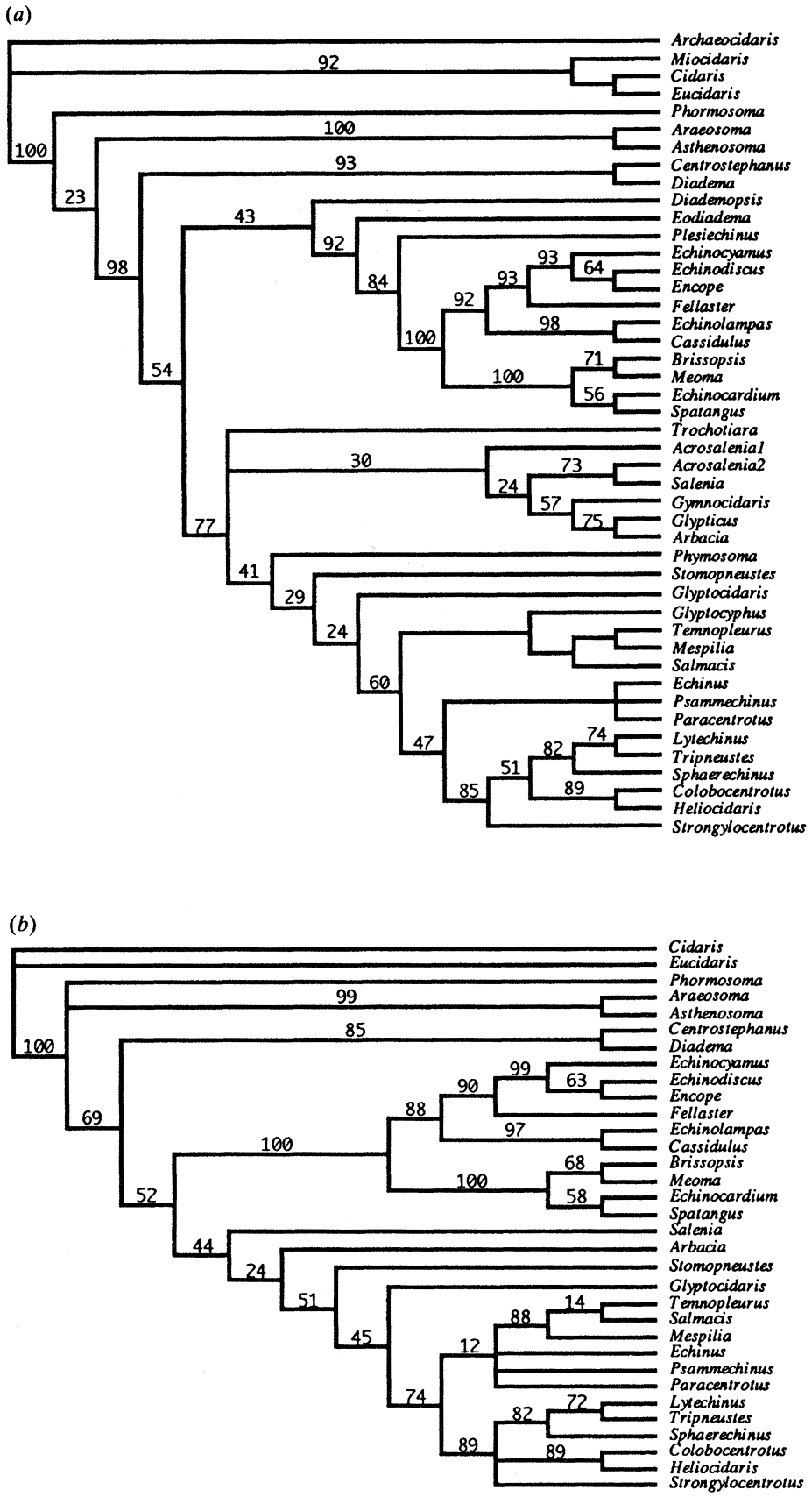

Figure 1. Semistrict consensus trees of morphological data for (*a*) fossil and Recent taxa, and (*b*) Recent taxa only; numbers represent bootstrap percentages (1000 bootstrap replicates).

them as sister taxa (figure 1*b*). Bootstrapping gave low support for the specific branching order of stirodont taxa in relation to the camarodonts, but did support the sister group pairing of a stirodont plus camarodont clade to the irregular echinoids. Irregular echinoids as a clade was supported in 100% of the bootstrap replicates.

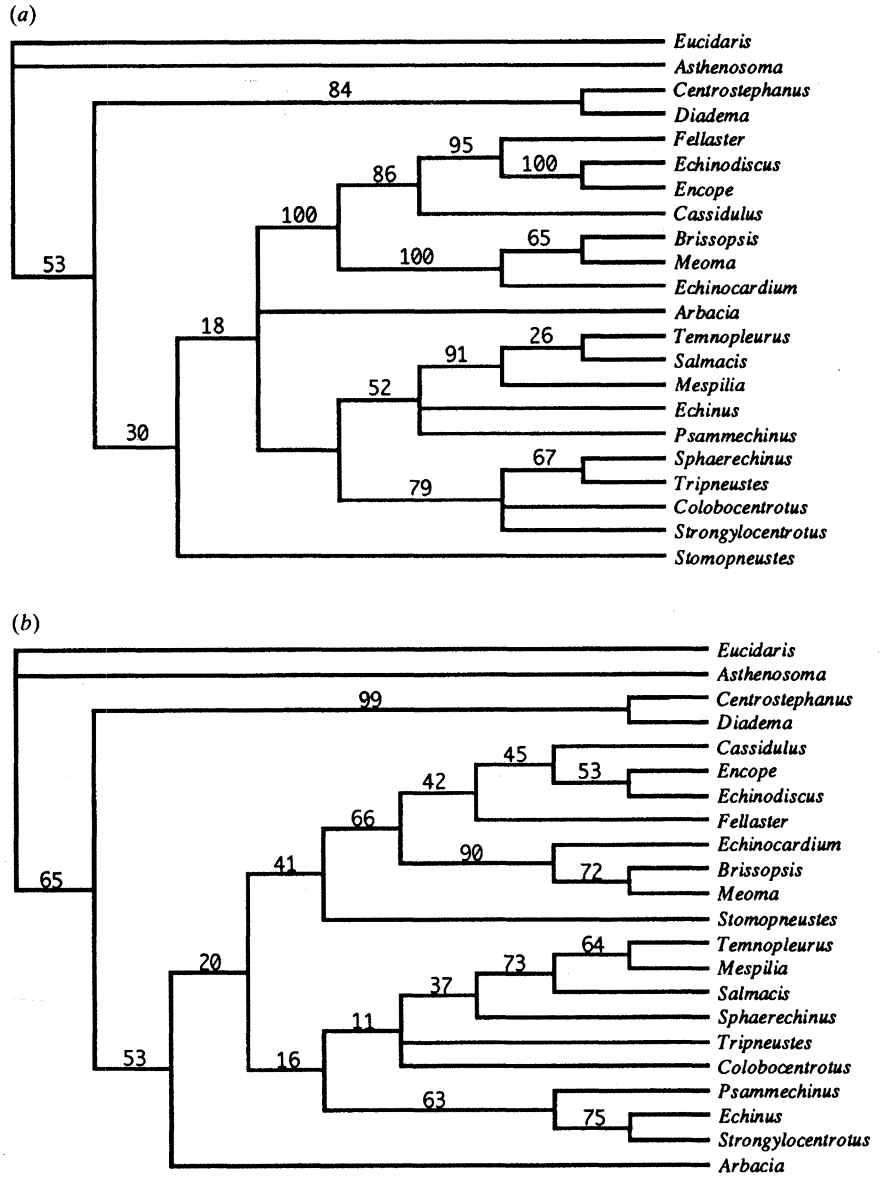

Figure 2. Semistrict consensus trees (*a*) of morphology where we have SSU rRNA sequence data; (*b*) of SSU rRNA sequence data; numbers represent bootstrap percentages (1000 bootstrap replicates).

(iii) *Taxa for which there is ssu rRNA sequence data*

This included 22 taxa and found 12 equally parsimonious trees; length = 249, $CI = 0.72$, $RI = 0.86$. These could be summarized in a single tree by treating the three species of temnopleurid as a trichotomy, and by treating *Echinus*, *Paracentrotus*, *Psammechinus* and the temnopleurids as a trichotomy (figure 2*a*). There are two differences of arrangement compared to the analysis of recent and fossil taxa. Firstly the position of *Arbacia* was unstable, falling either as sister group to irregular echinoids or as sister group to the camarodonts. *Stomopneustes* was consistently identified as outgroup to *Arbacia* plus irregular echinoids and camarodonts. Secondly, the temnopleurids were not identified as sister group to the other camarodonts, but were placed in a trichotomy with *Echinus* and *Psammechinus*. Bootstrapping (1000 replicates) gave strong support (greater than 80%) for a monophyletic irregular clade, a neognathostomate clade (clypeasteroids plus cassiduloids), a spatangoid clade, a diadematoid clade, a camarodont clade and a temnopleuroid clade. Four of the branches were supported in 40% or less of the replicates, including the branches relating to the positioning of *Arbacia*, *Stomopneustes*, irregular echinoids and camarodonts.

(iv) *Taxa for which there is lsu rRNA sequence data*

This included 13 taxa and found a single most parsimonious tree; length = 182 steps, $CI = 0.87$, $RI = 0.91$ (figure 3*a*). This is consistent with the topology found for taxa where there is SSU rRNA sequence data as well as with that found for all Recent taxa. All branches were supported at greater than 60% of 1000 bootstrap replicates.

(v) *Taxa for which there is both SSU and LSU rRNA sequence data*

This included nine taxa, including both *Cidaris* (LSU) and *Eucidaris* (SSU) data for which there are LSU and SSU rRNA sequences respectively. Two equally parsimonious trees were found; length = 178, $CI = 0.82$, $RI = 0.84$ (figure 4*a*). Each branch was supported by greater

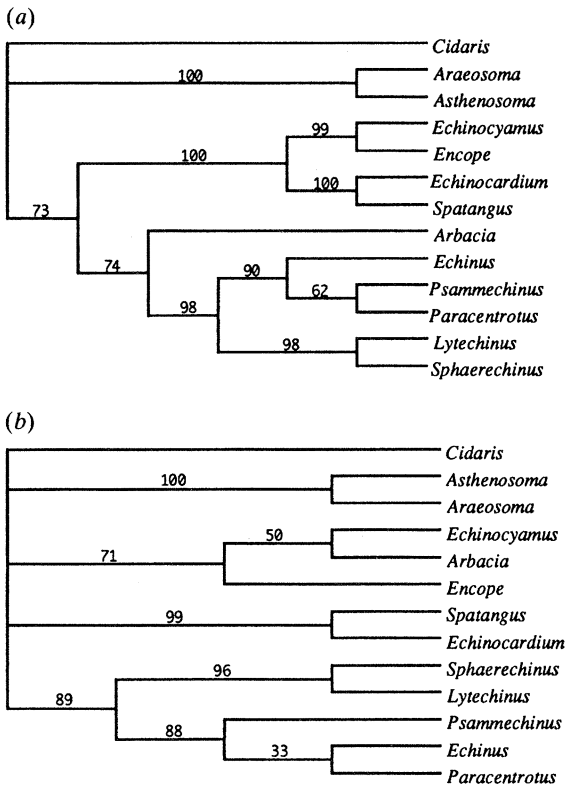

Figure 3. Semistrict consensus trees of (*a*) morphology where we have LSU rRNA sequence data; (*b*) LSU rRNA sequence data; numbers represent bootstrap percentages (1000 bootstrap replicates).

than 50% of 1000 bootstrap replicates. One topology places irregular echinoids as sister group to *Asthenosoma* and the stirodonts and camarodonts, the other places *Asthenosoma* as sister group to irregular echinoids plus stirodonts and camarodonts. The second topology is consistent with trees generated from the full morphological data set described above.

(*b*) *Molecular data*

(i) *LSU rRNA*

Total data. The first 400 b.p. of the LSU rRNA gene were used, position 1 corresponding to position 8 in the mouse (Hassouna *et al.* 1984; GenBank Accession X00525). There were a total of 89 variable characters. Parsimony analysis found eight minimal length trees of 142 steps, $CI = 0.76$, $RI = 0.81$. These recognized a number of the expected monophyletic groups (camarodonts, spatangoids, echinothuriids) which were supported by high bootstrap values, but consistently placed *Arbacia* (a stirodont) within the clypeasteroids (with a bootstrap support of 71%), and failed to identify irregular echinoids as monophyletic. A semistrict consensus of the sixteen trees is shown in figure 3*b*.

Maximum-likelihood yielded a similar topology with camarodonts and spatangoids recognised as monophyletic groups. Again *Arbacia* fell within the clypeasteroids. Unexpectedly the clypeasteroids were placed as a sister group to the camarodonts. Few branches were, however, significantly positive in length.

Data where there are ssu rRNA sequences. Where the cidarid *Cidaris* was retained as the outgroup, maximum

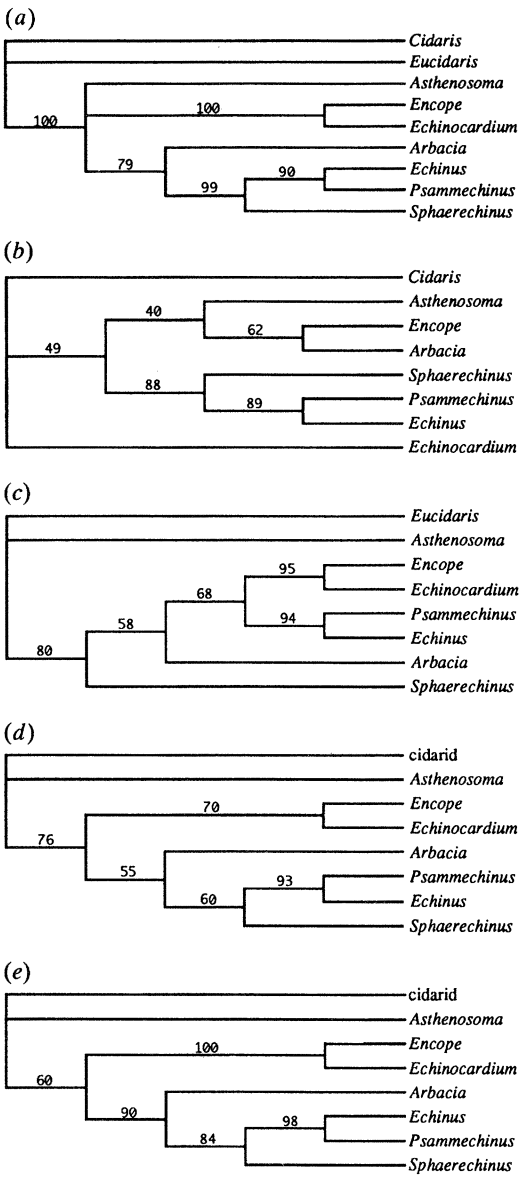

Figure 4. Single most parsimonious trees for taxa where we have morphology, SSU and LSU rRNA sequence data. Data sets treated independently include (*a*) morphology, (*b*) LSU rRNA, (*c*) SSU rRNA, and treated as combined evidence include (*d*) all molecular data (LSU + SSU rRNA), and (*e*) total evidence (morphology + molecular). In the combined data the taxon cidarid was a combination of *Cidaris* and *Eucidaris* characters; see text for details.

parsimony found a single minimal length tree with data from eight taxa (figure 4*b*); 123 steps, $CI = 0.70$, $RI = 0.63$. This conforms to the topology found when all 13 taxa are included. The stirodont *Arbacia* was maintained as the sister taxon to the clypeasteroid.

(ii) *SSU rRNA*

Total data. The entire sequence consisting of 279 variable characters (shown in Appendix 2), was analysed using parsimony and maximum-likelihood. Parsimony analysis found two trees; length = 440 steps, $CI = 0.50$, $RI = 0.58$. These differed only in their placement of *Tripneustes*, and the semistrict consensus tree is shown in figure 2*b*. However, 1000 bootstrap replications found strong support (more than 70%) for

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

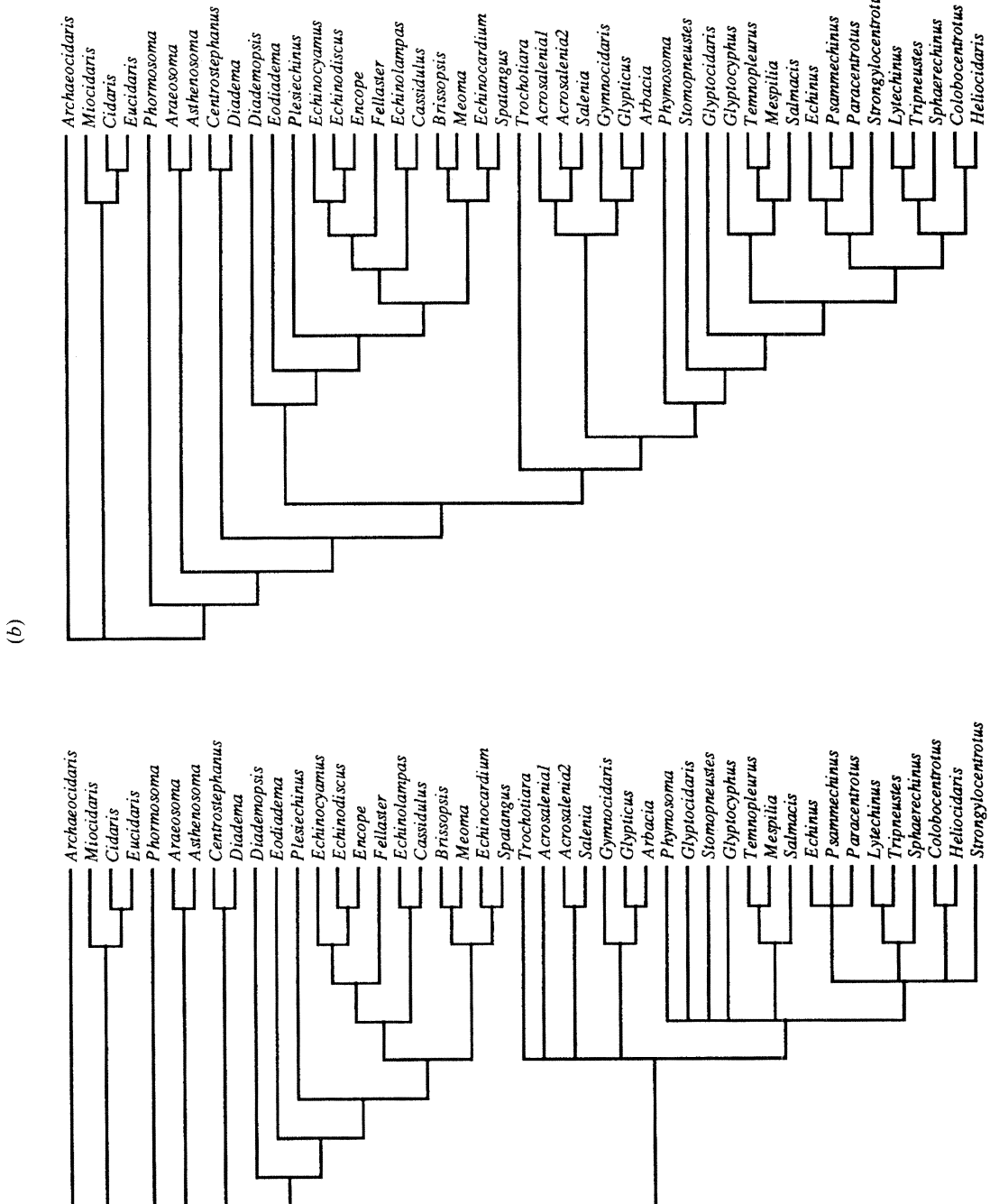

Figure 5. Results of total evidence analysis on all taxa and all data showing (*a*) semistrict consensus tree; (*b*) majority rule consensus tree.

only a small portion of the branches identified. The topology differs from that identified by morphology in the placement of *Stomopneustes* and *Arbacia* relative to the irregular echinoids and camarodonts, though none of these branches are strongly supported. The position of *Cassidulus* also differs, since it falls within the clypeasteroid clade with *Fellaster* as sister group to *Cassidulus* plus *Encope* plus *Echinodiscus*.

Although few branches were significantly positive, maximum-likelihood again generated a tree with a topology conflicting with that determined from morphology. The temnopleurids and spatangoids each appear as monophyletic groupings but with the temnopleurids as a derived clade within the camarodonts, as they were in the parsimony analysis. *Cassidulus*, a cassiduloid, is placed within the clypeasteroids and *Sphaerechinus*, an echinometrid, is sister group to the diadematoids, making the camarodonts paraphyletic.

Data where there are LSU *rRNA sequences*. Maximum parsimony performed on those taxa where we have complementary LSU rRNA data only (eight taxa, retaining *Cidaris* as the outgroup) found a single most parsimonious tree of 233 steps, $CI = 0.65$, $RI = 0.53$. Although the irregulars are held to be monophyletic they were placed as derived clade within the camarodonts which are themselves paraphyletic due to the position of the stirodont *Arbacia* (figure 4*c*).

(iii) *SSU and LSU ribosomal RNA combined*

With a total of eight taxa the combined SSU and LSU rRNA data gave a total of 371 variable sites. LSU data from *Cidaris* and SSU data from *Eucidaris* were combined to give a general cidaroid sequence. All other combined SSU and LSU rRNA data sets were from conspecifics. Parsimony analysis generated a single most parsimonious tree of 143 steps, $CI = 0.63$, $RI = 0.49$ (figure 4*d*). The topology was identical to that generated by morphological data alone with more than 50% of 1000 bootstrap replicates supporting each branch.

Maximum-likelihood analysis placed *Arbacia*, a stirodont, among the irregular echinoids in spite of most branches collapsing into polytomies. Only the camarodonts were monophyletic.

(*c*) *Combined morphological and molecular data: total evidence*

Data were combined in two different ways. First, only taxa were chosen where complementary molecular and morphological data were available. This gave eight taxa with a general cidarid derived from combining morphological and LSU rRNA data from *Cidaris* and SSU rRNA data from *Eucidaris*. Parsimony analysis found a single most parsimonious tree (length = 434, $CI = 0.72$, $RI = 0.66$) supported at each branch by 60% or more of the 1000 bootstrap replicates performed. The topology (figure 4*e*) was identical with that derived from morphology alone. The least well supported branch was that between the cidarid plus echinothurioid clade, though even this was supported by more than 60% of bootstrap replicates.

The second combination of morphological and molecular data included all taxa and where LSU or SSU rRNA sequence data was lacking from specific taxa it was simply scored as missing ('?'). Combining the 163 adult morphological characters with the 371 molecular characters gave a data matrix of 534 characters. Parsimony analysis found 75 minimum length trees; length = 894, $CI = 0.58$, $RI = 0.80$. A semistrict consensus tree (figure 5*a*) was less resolved than using morphological data alone, particularly among the basal (extinct) stirodont taxa which have the greatest amount of missing data. However, the majority-rule consensus tree of the combined data (figure 5*b*) was identical with that produced with morphological data alone (semistrict), except that 50 of these trees (65%) placed *Strongylocentrotus* with the echinids as a sister taxon to the other echinometrids.

4. DISCUSSION

(*a*) *Morphological topology and the effect of including fossil taxa*

It has been demonstrated empirically that fossils play a pivotal role in establishing phylogenetic relationships (Doyle & Donoghue 1987; Gauthier *et al.* 1988; Donoghue *et al.* 1989) and if ignored less resolved or positively misleading topologies result (see review in Smith 1994). In our study, when only those extant taxa for which molecular data had been gathered were selected for morphological analysis, a topology resulted that differed from that found when all major extant groups were included. Furthermore, adding a number of key fossil taxa, even though these were less completely known, altered the topology further. There were two significant changes: (i) When recent taxa alone were sampled, temnopleurids were derived within the camarodonts rather than being sister group to the rest. Camarodonts have a very long stem group that separates crown group members (that diverged over the last 60 million years) from their closest stirodont sister group (divergence *ca.* 160 million years). The addition of one member of this stem group, *Glyptocyphus*, changes the rooting position within the crown group, making temnopleurids sister group to the rest. Unfortunately nothing is as yet known about the pedicellariae or lantern structure of *Glyptocyphus*, so it cannot be scored for its full suite of characters. However, the apical disc structure of *Glyptocyphus* is very similar to that of *Phymosoma* and differs from that seen in either *Glyphocidaris* or *Stomopneustes*, or temnopleurids. The lack of knowledge of pedicellarial and lantern features of many fossil taxa precludes firm placement of camarodont phylogenetic relationships, but *Stomechinus* and *Glyphocidaris* are identified as closest extant outgroups to the Camarodonta in a number of our analyses. (ii) The only other significant change to result from the addition of fossil taxa concerns the relative positioning of the stirodonts *Arbacia* and *Salenia*. When just extant taxa were considered, *Salenia* was identified as sister group to all other stirodonts plus camarodonts, with *Arbacia* as the next branch. However, the addition of early (fossil) members of both the Calycina and Arbacioida clades changes the topology

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

such that *Arbacia* and *Salenia* are recognized as sister taxa.

In both cases the addition of fossil taxa belonging to the stem group of clades effected a change in cladogram topology. This confirms the views put forward by Doyle & Donoghue (1987), Gauthier *et al.* (1988), Donoghue *et al.* (1989), Huelsenbeck (1991), Novacek (1992*a*, *b*) and Smith (1994) that the addition of fossil data, despite its sometimes incomplete nature, can sometimes have a significant effect. Of course, the addition or removal of any taxon, Recent or fossil, may affect the result of a phylogenetic analysis, and the addition of fossil taxa is in effect only adding to the density of our sampling of the Echinoidea. However, some clades such as the Camarodonta, have a diverse crown group and a very long stem lineage that cannot be sampled other than through the fossil record. Rooting the crown group correctly in this case requires the addition of an appropriate stem group member.

The accuracy of the phylogenetic reconstruction decreases further when only a subset of Recent taxa belonging to the ingroup (in this case Echinoidea) are sampled, as with the data sets constructed for taxa for which we have SSU or LSU rRNA sequence data. Kluge & Wolf (1993) and Eernisse & Kluge (1993) have argued that parsimony analysis should be conducted on the total evidence that can be amassed otherwise it may lead to inaccurate topologies. Although they argue primarily for total evidence in terms of combining all available data matrices, it is significant to note that the density at which taxa are sampled is also critically important. Clearly the most accurate reconstruction comes from including a comprehensive spread of taxa from among the ingroup.

A phylogenetic analysis of extant taxa that does not have good systematic coverage of ingroup taxa can produce misleading topologies, because incomplete sampling leads to homoplasies being mistaken for homologies. In some instances, even dense sampling of extant taxa may give misleading topologies if the clade has a long stem lineage in which significant change has occurred. Thus the addition of key fossil taxa help to provide a denser sampling of character distributions than would be possible from extant taxa alone.

(*b*) *Comparison of SSU and LSU rRNA molecular phylogenies*

Hillis & Dixon (1991) reviewed the relationship between phylogenetic resolution afforded by SSU and LSU nuclear gene sequences and the divergence of taxa through geological time. They concluded that in general SSU rRNA nuclear gene sequences appear to be the best choice of molecule for elucidating Precambrian divergences and LSU rRNA nuclear gene sequences for Palaeozoic and Mesozoic divergences. These authors were considering the complete gene sequences, which are approximately 1800 b.p. and 5000 b.p. for the SSU and LSU rRNA genes respectively. Each mature gene product is functional as a folded rRNA molecule with characteristic divergent domains and conserved regions where base changes are more or less likely to occur (see Hill *et al.* 1990 and papers therein). Although the region of the LSU rRNA gene used in our study spans only one divergent domain (D1 according to the nomenclature of Larsen *et al.* 1993) it has been considered sufficiently informative to resolve phylogenetic divergence events occurring in the Mesozoic (Qu *et al.* 1988).

The two echinoid phylogenies derived from SSU and LSU rRNA sequence data sets have much in common, but also differ in a number of key aspects. The most obvious difference is in their ability to resolve relationships: whereas the SSU rRNA data give a highly resolved cladogram, the LSU rRNA data leaves deep-branch relationships unresolved. One possibility is that this difference may arise because of differences in the number of phylogenetically informative positions derived from each gene, since the SSU rRNA sequence is *ca.* 1800 b.p. long whereas the LSU rRNA partial sequence is only some 360 b.p. long. However, in the combined data matrix for the eight taxa for which we have both SSU and LSU rRNA sequence data, 38 phylogenetically informative sites are derived from the SSU rRNA sequence and 30 from the LSU rRNA sequence. Therefore there is no significant difference in the number of phylogenetically informative sites contained in the two sequences. Since the SSU rRNA sequence is almost five times as long as the LSU rRNA sequence we are using, the partial LSU rRNA sequence must be accumulating fixed mutations at a higher rate than the SSU rRNA sequence. Consequently variable sites in the LSU rRNA partial sequence are presumably more likely to become overprinted over time than those in the SSU rRNA sequence. This faster rate of fixed point mutational accumulation in the LSU rRNA sequence explains why deeper branches are less well resolved in the phylogeny derived from LSU rRNA sequence data compared to SSU rRNA sequence data. Phylogenetic information about deeper branches has been more heavily overprinted in the LSU rRNA sequence and any signal about closely spaced branching order has been lost through time. This accords with the relative rates of mutation previously observed and documented (see review in Hillis & Dixon 1991).

The most unexpected aspect of the LSU rRNA sequence data is that they consistently place the stirodont *Arbacia* within the clypeasteroids with high bootstrap support, making both clypeasteroids and irregular echinoids non-monophyletic. This is clearly wrong since it contradicts both morphological and SSU rRNA phylogenies, yet the reason for this placement of *Arbacia* among clypeasteroids is not at all obvious. Although there are some sites in our LSU rRNA partial sequence which suggest the placement of *Arbacia* as a sister group to camarodonts, there are a greater number of sites linking *Arbacia* and clypeasteroids (especially *Echinocyamus*). However, we have not sequenced the LSU rRNA gene of *Arbacia* ourselves and the first step must be to confirm the published sequence derived from direct RNA sequencing.

Similarly the placement of *Cassidulus* among the clypeasteroids is an unexpected result from the SSU rRNA sequence data. This topology is only weakly supported from the molecular data, with only one uncontradicted base change in the SSU rRNA sequence

to support the inclusion of *Cassidulus* within Clypeasteroidea. The monophyly of Clypeasteroidea is firmly established on morphological grounds (e.g. Mooi 1990) and a total evidence approach places *Cassidulus* as sister group to the Clypeasteroidea. The position of *Cassidulus* indicated from SSU rRNA sequence data is therefore weakly supported and may change with additional data.

SSU rRNA data originally compiled by Raff *et al.* (1988) and used for phylogenetic analysis of echinoderm class relationships, omitted the two most variable regions of the sequence (positions 35–59 and 1709–1761 in our alignment). Consequently the resolving power of these partial sequences (which align to regions 184–556, 722–1127, and 1405–1644 respectively in our alignment; EMBL accession number DS19161) was very limited (Smith 1989). We initially included Raff *et al.*'s partial sequences, scoring the missing variant sites as unknown. However, this greatly increased the number of equally parsimonious solutions that were found using parsimony analysis, and were consequently omitted from further consideration. Any phylogenetic analysis of relatively recent divergences (e.g. within 250 million years) should include these two variable domains for best results.

The results from maximum-likelihood were disappointing given the good performance of this technique on a reduced LSU rRNA data set (Smith 1993). Much of the resolution power of our parsimony analysis must come from the treatment of missing bases as a fifth character. When missing bases were treated as unknown a highly unexpected result was achieved: *Asthenosoma* paired with *Temnopleurus* among the temnopleurids and *Sphaerechinus* was placed as outgroup to stirodonts, irregular echinoids and other camarodonts, exactly as was found using maximum-likelihood. Maximum-likelihood calculations ignore sites at which there is missing data and, in this case, would appear to omit important phylogenetic information.

(*c*) *Total evidence as the best estimate of the echinoid phylogenetic tree*

In comparing morphological and molecular phylogenies care must be taken in rejecting one or other topology. As Larson (1994) notes: 'If the most parsimonious topology for the molecular data constitutes a suboptimal tree for the morphological data and vice versa, it is incorrect to conclude that the data sets conflict. We cannot reject the hypothesis that both data sets are estimating with error the same phylogenetic topology.' To test whether the most parsimonious topologies for each data set constitute suboptimal topologies for the other one Larson (1994) recommends the non-parametric test of Templeton (1983). Using the cases where we have taxa in common between morphology and molecular data Templeton's test indicates that the LSU rRNA tree (figure 3*b*) is significantly rejected by the morphological data (figure 3*a*) and vice versa ($p < 0.001$, Wilcoxon's matched-pairs signed-ranks test; see Larson 1994) whereas the differences between the SSU rRNA tree (figure 2*b*) and the morphological tree is not significant at the $p > 0.01$ level (figure 2*a*). Differences between the trees (figures 2*a*, *b*) can be attributed to random error. Thus the topology derived from the morphology is judged to be suboptimal for the SSU rRNA molecular data and vice versa, whereas the topology derived from the LSU rRNA conflicts with its morphological counterpart significantly. This evidence might argue for a consensus approach (de Queiroz 1993) when combining the LSU rRNA data with morphology as they each have strong but conflicting phylogenetic signal. However, it is really only the position of *Arbacia* in the LSU rRNA tree which is problematic. Due to the unstable position of taxa such as *Arbacia*, a consensus tree from the three data sets would be largely uninformative, and would cause all higher taxa to collapse into a major polytomy. Therefore we have opted for a combined evidence approach, as advocated by Eernisse & Kluge (1993) and Kluge & Wolf (1993).

Our combined data set has been compiled in one of two ways. Firstly we removed all taxa that were absent from one or more of the three data sets, leaving just eight taxa in total; and secondly we constructed a data matrix for all 46 taxa, with missing data scored as unknown ('?'). Both approaches gave congruent topologies, and consequently we have opted for the full combined data tree (figure 5*b*) as our best estimate of the phylogenetic relationships of higher taxa of echinoids. Our preferred topology identifies a number of hidden homoplasies within the molecular data. Whereas treelength for the most parsimonious topology derived from SSU rRNA data alone was 440 steps, the same data optimized over the combined evidence tree has a tree length of 455 (an increase of 3.3%). Similarly tree length calculated from LSU rRNA data on its own is 142 steps, but the same data optimized over the combined evidence topology has a tree length of 147 (an increase of 3.4%). The increase in tree length estimates for morphological data is much less, where tree length increased by only 2 steps (0.6%).

Our results agree to a large extent with previously proposed echinoid phylogenies. Cidaroids are identified as sister group to all other echinoids while echinothurioids are placed as the least derived of the euechinoid clade. However, our results indicate that echinothurioids are not a monophyletic group, despite their apparent similar derived style of ambulacral compounding. *Phormosoma* (Phormosomatidae) is the first branch and is sister group to Echinothuriidae (*Araeosoma* and *Asthenosoma*) plus other euechinoids. Irregular echinoids form a monophyletic group, as suggested by Smith (1981) and Jensen (1981) and within this clade spatangoids are the sister group to cassiduloids plus clypeasteroids, as first suggested by Kier (1974). Insufficient cassiduloids and clypeasteroids are sampled to say anything about the relationships within neognathostomates.

The relationship of irregular echinoids to other groups has always been uncertain. Our analysis suggests that irregular echinoids are sister group to stirodont plus camarodont echinoids (regular echinoids with keeled teeth). The diadematoids are sister group to the larger clade composed of irregular echinoids plus camarodonts and stirodonts.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Finally, among stirodonts it would appear that Calycina and Arbacioida are paired taxa and sister group to other stirodont plus camarodont taxa. The immediate extant outgroup to camarodonts may be *Glyphocidaris*, though we do not have molecular data to confirm this. Our analysis suggests that stirodonts plus camarodonts divide into Calycina plus Arbacioida, with Acrosaleniidae as part of their stem group, and *Stomopneustes*, *Glyphocidaris* and the camarodonts, with phymosomatoids as part of their stem group. The position of *Pseudodiadema* as a trichotomy at the base of the stirodont plus camarodont clade fits well with the fossil record that suggests this is one of the earliest representatives to appear (Kier 1977). We can now be slightly more confident about relationships among camarodont families, although we have no molecular data from LSU rRNA yet for the Temnopleuridae. The Temnopleuridae are probably sister group to other camarodont families. Toxopneustidae and Echinometridae are sister groups with Strongylocentrotidae as their combined sister group.

Despite our wealth of data, there still remain a few outstanding phylogenetic problems. Our work has failed to derive a statistically robust solution to the order of branching of stirodont taxa leading up to the camarodonts, and relationships within stirodonts still remain relatively weak.

We thank the following individuals for help in locating, providing and transporting echinoid material: S. Amemiya, R. Aronson, A. Baker, A. Bentley, H. Dixon, S. Donovan, R. Emson, C. Freire, P. Gayle, J. Gage, M. Gibbons, M. Hart, H. Hayashi, G. Hendler, W. Hide, T. Kikuchi, J. Korrubel, B. Lafay, D. McKenzie, T. Matsuoka, P. Mikkelsen, B. Morton, T. Motokawa, D. Nichols, S. Palumbi, N. Suzuki, J. Taylor, P. Tyler, T. Uehara, J. Woodley and C. Young. This study benefitted from the SEQNET facilities and was supported by a NERC grant (GR37960).

REFERENCES

Boom, R., Sol, C.J.A., Salimabs, M.M.M., Lansen, C.L., Wertheim-van Dillen, P.M.E. & van der Noordaa, J. 1990 Rapid and simple method for purification of nucleic acids. *J. clin. Microbiol.* **28**, 495–503.

Christen, R. 1993 VSM (Computer data base and alignment package). Distributed by the author, Station Zoologique, Centre Océanologique, Villefranche-sur-Mer, 06230, France.

David, B. 1987 Origins of the deep-sea holasteroid fauna. In *Echinoderm phylogeny and evolutionary biology* (ed. C.R.C. Paul & A.B. Smith), pp. 331–346. Oxford: Clarendon Press.

De Queiroz, A. 1993 For consensus (sometimes). *Syst. Biol.* **42**, 368–372.

Donoghue, M.J., Doyle, J., Gauthier, J., Kluge, A. & Rowe, T. 1989 The importance of fossils in phylogeny reconstruction. *A. Rev. Ecol. Syst.* **20**, 431–460.

Doyle, J. & Donoghue, M.J. 1987 The importance of fossils in elucidating seed plant phylogeny and macroevolution. *Rev. Paleobot. Palynol.* **50**, 63–95.

Doyle, J.J. & Doyle, J.L. 1987 A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **19**, 11–15.

Durham, J.W. & Melville, R.V. 1957 A classification of echinoids. *J. Paleont.* **31**, 242–272.

Eernisse, D.J. & Kluge, A.G. 1993 Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules, and morphology. *Molec. biol. Evol.* **10**, 1170–1195.

Embley, T.M., Finlay, B.J., Thomas, R.H. & Dyall, P.L. 1992 The use of rRNA sequences and fluorescent probes to investigate the phylogenetic positions of the anaerobic ciliate *Metopus palaeoformis* and its archaeobacterial endosymbiont. *J. gen. Microbiol.* **138**, 1479–1487.

Felsenstein, J. 1993 PHYLIP: Phylogeny Inference package version 3.5c. Distributed by the author, Department of Genetics, University of Washington, Seattle.

Féral, J.-P. & Derrelle, E. 1991 Partial sequence of the 28S ribosomal RNA and the echinoid taxonomy and phylogeny. Application to the Antarctic brooding schizasterids. In *Echinoderm biology* (ed. T. Yanagisawa, I. Yasumasu, C. Oguro, N. Suzuki & T. Motokawa), pp. 331–337. Rotterdam: A.A. Balkema.

Féral, J.-P., Derrelle, E. & Philippe, H. 1994 Inferred phylogenetic trees of antarctic brood-protecting schizasterid echinoids from partial 28S ribosomal RNA sequences. In *Genetics and evolution of aquatic organisms* (ed. A.R. Beaumont), pp. 199–206. London: Chapman & Hall.

Gauthier, J., Kluge, A.G. & Rowe, T. 1988 Amniote phylogeny and the importance of fossils. *Cladistics* **4**, 105–209.

Hassouna, N., Michot, B. & Bachellerie, J.P. 1984 The complete nucleotide sequence of mouse 28S rRNA gene. Implications for the process of size increase of the large subunit rRNA in higher eukaryotes. *Nucl. Acids Res.* **12**, 3563–3583.

Hill, W.E., Dahlberg, A., Garrett, R.A., Moore, P.B., Schlessinger, D. & Warner, J.R. 1990 *The ribosome: structure, function and evolution*. Washington, D.C.: American Society for Microbiology.

Hillis, D.M. & Dixon, M.T. 1991 Ribosomal DNA: molecular evolution and phylogenetic inference. *Q. Rev. Biol.* **66**, 411–452.

Hosta, L. & Flick, P. 1991 Enhancement of specificity and yield in PCR. *U.S. Biochem. Comments* **18**(3), 1–5.

Huelsenbeck, J.P. 1991 When are fossils better than extant taxa in phylogenetic analysis? *Syst. Zool.* **40**, 458–469.

Hyman, L.H. 1955 *The invertebrates*, vol. 4 (*Echinodermata*). New York: McGraw–Hill.

Jensen, M. 1979 Primary plates of sea urchin teeth (Echinoidea). *Vidensk. Meddr dansk naturh. Foren.* **141**, 7–27.

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

Kier, P.M. 1974 Evolutionary trends and their functional significance in the post-Paleozoic echinoids. *J. Paleont.* **48** (Suppl.) *Paleont. Soc. mem.* **5**, 1–95.

Kluge, A.G. & Wolf, A.J. 1993 Cladistics: what's in a word. *Cladistics* **9**, 183–200.

Larsen, N., Olsen, J.G., Maidak, B.L., McCaughley, M.J., Overbeek, R., Macke, T.J., Marsh, T.L. & Woese, C.R. 1993 The ribosomal database project. *Nucl. Acids Res.* **21**(Suppl.), 3021–3023.

Larson, A. 1994 The comparison of morphological and molecular data in phylogenetic systematics. In *Molecular approaches to ecology and evolution* (ed. B. Schierwater, B. Streit, G.P. Wagner & R. DeSalle), pp. 371–390. Basel: Birkhäuser Verlag.

Lenaers, G., Maroteaux, L., Michot, B. & Herzog, M. 1989 Dinoflagellates in evolution. A molecular phylogenetic analysis of large subunit ribosomal RNA. *J. molec. Evol.* **29**, 40–51.

Lewis, D.N. & Ensom, P.C. 1982 *Archaeocidaris whatleyensis* sp. nov. (Echinoidea) from the Carboniferous limestone of Somerset and notes on echinoid phylogeny. *Bull. Br. Mus. Nat. Hist.* (*Geol.*) **36**, 77–104.

Littlewood, D.T.J. 1994 Molecular phylogenetics of cupped oysters based on partial 28S rRNA gene sequences. *Molec. phylog. Evol.* **3**, 221–229.

Maddison, W.P. & Maddison, D.R. 1992 MacClade: Analysis of phylogeny and character evolution, version 3.0. Sunderland, Massachusetts: Sinauer Associates. (Computer program and manual).

Marshall, C.R. & Swift, H. 1992 DNA-DNA hybridization phylogeny of sand dollars and highly reproducible extent of hybridization values. *J. molec. Evol.* **34**, 31–44.

Matsuoka, N. 1980 Immunological relatedness of sea-urchin glucose-6-phosphate dehydrogenases: phylogenetic implication. *Comp. Biochem. Physiol.* **66**B, 605–607.

Matsuoka, N. 1985 Biochemical phylogeny of the sea-urchins of the family Toxopneustidae. *Comp. Biochem. Physiol.* **80**B, 767–771.

Matsuoka, N. 1986 Further immunological study on the phylogenetic relationships among sea-urchins of the order Echinoida. *Comp. Biochem. Physiol.* **84**B, 465–468.

Matsuoka, N. 1987 Biochemical studies on the taxonomic situation of the sea-urchin *Pseudocentrotus depressus*. *Zool. Sci.* **4**, 339–347.

Matsuoka, N. 1989 Biochemical systematics of four sea-urchin species of the family Diadematidae from Japanese waters. *Biochem. Syst. Ecol.* **17**, 423–429.

Matsuoka, N. 1990 Evolutionary relationships of sea-urchins at the molecular level. *Comp. Biochem. Physiol.* **97**B, 31–36.

Matsuoka, N. & Suzuki, H. 1987 Electrophoretic study on the phylogenetic relationships among six species of sea-urchins of the family Echinometridae found in the Japanese waters. *Zool. Sci.* **6**, 589–598

Medlin, L., Elwood, H.J., Stickel, S. & Sogin, M.L. 1988 The characterization of enzymatically amplified 16S-like rRNA-coding regions. *Gene* **71**, 491–499.

Mooi, R.C. 1987 A cladistic analysis of the sand dollars (Clypeasteroida: Scutellina) and the interpretation of heterochronic phenomena. Unpublished Ph.D. thesis, Department of Zoology, University of Toronto, Canada.

Mooi, R. 1990 Paedomorphosis, Aristotle's lantern, and the origin of the sand dollars (Echinodermata: Clypeasteroida). *Paleobiology* **16**, 25–48.

Mortensen, 1928–1951 *A monograph of the Echinoidea*, volumes 1–5. Copenhagen: C.A. Reitzel.

Novacek, M.J. 1992 *a* Fossils, topologies, missing data, and the higher level phylogeny of eutherian mammals. *Syst. Biol.* **31**, 366–375.

Novacek, M.J. 1992 *b* Fossils as critical data for phylogeny. In *Extinction and phylogeny* (ed. M.J. Novacek & Q.D. Wheeler), pp. 46–88. New York: Columbia University Press.

Qu, L.-H., Nicoloso, M. and Bachellerie, J.-P. 1988 Phylogenetic calibration of the 5′ terminal domain of large rRNA achieved by determining twenty eucaryotic sequences. *J. molec. Evol.* **28**, 113–124.

Raff, R.A., Field, K.G., Ghiselin, M.T., Lane, D.J., Olsen, G.J., Parks, A.L., Parr, B.A., Pace, N.R. & 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.

Reysenbach, A.-L., Giver, L.J., Wickham, G.S. & Pace, N.R. 1992 Differential amplification of rRNA genes by polymerase chain reaction. *Appl. env. Microbiol.* **58**, 3417–3418.

Sambrook, J., Fritsch, E.F. & Maniatis, T. 1989 *Molecular cloning: a laboratory manual*, 2nd edn. Cold Spring Harbor Laboratory Press.

Sanger, F., Nicklen, S. & Coulson, A.R. 1977 DNA sequencing with chain terminating inhibitors. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467.

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: George Allen & Unwin.

Smith, A.B. 1988 Patterns of diversification and extinction in early Palaeozoic echinoderms. *Palaeontology* **27**, 431–459.

Smith, A.B. 1989 Echinoid evolution from the Triassic to lower Jurassic. *Cahiers de l'Université Catholique de Lyon, série science* **3**, 79–117.

Smith, A.B. 1994 *Systematics and the fossil record: documenting evolutionary patterns*. Oxford: Blackwell Scientific.

Smith, A.B. & Wright, C.W. 1989 British Cretaceous echinoids. Part 1, General introduction and Cidaroidea. *Monogr. Palaeontogr. Soc.* 1–101, pls 1–32 (publn no. 578, part of volume 141 for 1987).

Smith, A.B. & Wright, C.W. 1990 British Cretaceous echinoids. Part 2, Echinothurioida, Diadematoida and Stirodonta (1, Calycina). *Monogr. Palaeontogr. Soc.* 101–198, pls 33–72 (publn no. 583, part of volume 143 for 1989).

Smith, A.B. & Wright, C.W. 1993 British Cretaceous echinoids. Part 3, Stirodonta 2 (Hemicidaroida, Arbacioida and Phymosomatoida, part 1). *Monogr. Palaeontogr. Soc.* 199–267, pls 73–92 (publn no. 593, part of volume 147 for 1993).

Smith, A.B., Lafay, B. & Christen, R. 1992 Comparative variation of morphological and molecular evolution through geologic time: 28*S* ribosomal RNA versus morphology in echinoids. *Phil. Trans. R. Soc. Lond.* B **338**, 365–382.

Suter, S.J. 1994*a* Cladistic analysis of cassiduloid echinoids: trying to see the phylogeny for the trees. *Biol. J. Linn. Soc.* **53**, 31–72.

Suter, S.J. 1994*b* Cladistic analysis of the living cassiduloids (Echinoidea), and the effects of character ordering and successive approximations weighting. *Zool. J. Linn. Soc.* **112**, 363–387.

Suzuki, N. & Yoshino, K.-I. 1992 The relationship between amino acid sequences of sperm activating peptides and the taxonomy of echinoids. *Comp. Biochem. Physiol.* **102**B, 679–690.

Suzuki, N., Yoshino, K., Kurita, M., Yamaguchi, M. & Amemiya, S. 1988 Taxonomical significance of respiratory stimultion of sea urchin spermatazoa by egg associated substances. In *Echinoderm biology* (ed. R.D. Burk, P.V. Mladenov, P. Lambert & R.L. Parsley), pp. 213–218. Rotterdam: A.A. Balkema.

Swofford, D.L. 1993 PAUP: Phylogenetic Analysis Using Parsimony, version 3.1.1. Washington D.C.: Smithsonian Institution.

Templeton, A. 1983 Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution* **37**, 221–244.

Wada, H. & Satoh, N. 1994 Details of the evolutionary history from invertebrates to vertebrates, as deduced from the sequences of 18S rDNA. *Proc. natn. Acad. Sci. U.S.A.* **91** 1801–1804.

Winship, P.R. 1989 An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide. *Nucl. Acids Res.* **17**, 1266.

Wray, G.A. 1992 The evolution of larval morphology during the post-Paleozoic radiation of echinoids. *Paleobiology* **18**, 258–287.

Received 28 July 1994; accepted 31 August 1994

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

APPENDIX 1.

Morphological characters used in the phylogenetic analysis

Character states are given in parentheses. All characters were treated as unordered.

(*a*) *Gross shape and general test features*

1. Test pentaradially symmetrical and circular in outline (0); pentaradial but bilaterally symmetric along the III-5 axis. (1); pentaradial but bilaterally symmetric along the II-4 axis. (2).
2. Test basically globular or ovoid (0); test flattened and discoidal in profile (1).
3. Test coronal plates firmly sutured together forming a solid test (0); test plates imbricate, not sutured together (1).
4. Plate suture faces smooth (0); embellished with a system of pegs and pits for interlocking (1).
5. Internal butressing absent (0); present with pillars and partitions uniting apical and adoral surfaces (1).
6. Coronal plates unornamented without pits (0); with a system of sutural pits (1).
7. Plates smooth and unornamented (0); plate surface ornamented by glassy epistroma (1).
8. Test not cordiform in outline (0); cordiform in outline (1).

(*b*) *Apical disc*

9. Periproct surrounded by plates of the apical disc system (0); periproct opening lying partially or completely outside the apical disc plates (1).
10. Periproct in contact with plates of the apical disc system (0); periproct opening entirely separate from the apical disc plates (1).
11. Five genital pores in adults (0); four gential pores in adults (1).
12. Apical disc includes genital plate V (0); genital plate V absent from apical disc (1).
13. Apical disc plates arrangment: monocyclic (0); dicyclic (1); hemicyclic (asymmetric, with posterior oculars exsert) (2); monobasal (3); tetrabasal (4).
14. Apical disc without madreporite plate extending posteriorly (0); with madreporite plate extending well to the posterior (ethmolytic pattern) (0).
15. Suranal plate absent (0); present during at least early ontogeny (1).
16. Enlarged periproctal plate ('suranal plate') undifferentiated in adults (0); one such plate significantly enlarged and forming an integral part of the apical disc (1).
17. Apical disc plates firmly sutured to coronal plates (0); only loosely attached and caducous upon death (1).
18. Periproctal plates not valve-like (0); arranged as a four-plate valve (1).
19. Periproct not enlarged (0); enlarged as an anal cone (1).
20. Periproct situated on aboral surface (0); on posterior (1); on oral surface (2).
21. Periproct not associated with interambulacral plates 2/3 (0); periproct firmly bound with interambulacral plates 2/3 (1).
22. Genital plates single (0); formed of a mosaic of small platelets (1).
23. Madrepores confined to a single genital plate (0); extending onto periproctal plates as well as the genital plate (1).
24. Apical disc flush with the corona (0); forming an elevated cap (1).
25. Interior of plates without projections (apophyses); with such projections (1).
26. Periproctal opening basically circular or teardrop-shaped (0); longitudinal (1); keyhole (3).
27. Genital plates pentagonal but not penetrating deep into the interambulacra (0); genital plates penetrate deeply along the interradius giving the disc a pentagonal outline (1).
28. Periproct rim – angular (0); smooth and oval, not indented by plates of the periproctal system (1).

(*c*) *Ambulacra*

29. Ambulacra simple (0); compound (1).
30. Polygeminate compound plates: not developed (0); present (1).
31. Bigeminate compound plates: absent (0); present (1).
32. Style of compounding: simple (0); acrosaleniid style, with two united elements alternating with a simple element (1); echinid style, with the lower element the largest (2); diadematoid-arbaciid style, with upper and lower elements reduced and typically occluded from the perradius (3); stomopneustid style, with demiplates inserted both above and below the middle element (4); pseudocompound, without overgrowth of a primary tubercle (5); phymosomatoid style with demiplates inserted between the middle and upper element (6).
33. Occluded plates never developed (0); always present in adults (1).
34. Respiratory tube-feet absent (0); arbaciid-type respiratory tube-feet (1); petaloid tube-feet (2).
35. Pore-pairs unmodified for respiration (0); modified into obvious petals (1).
36. Ambulacra I, II, IV, V, flush with the corona adapically (0); sunken to form pouches (1).
37. Ambulacrum III flush or hardly sunken (0); strongly sunken and modified for mucous string feeding (1).
38. Petals tapering towards apex and more or less closed (0); widening adapically and open (1).
39. Penicillate tube feet absent from around the mouth (0); present (1).
40. One tube-foot to each ambulacral plate (0); multiple accessory tube-feet to each ambulacral plate (1).
41. Ambulacrum III tube-feet undifferentiated from other ambulacra adapically (0); specialized and differentiated (1).
42. Funnel-building tube-feet absent adapically from ambulacrum III (0); present (1).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

43. Ambulacral pores confined to ambulacral plates (0); also extending into adjacent interambulacral plates (1).
44. Pores and tube-feet not organised into discrete rows (0); arranged into close-packed rows to form characteristic comb-like areas (1).
45. Food grooves absent (0); present as simple perradial channel (1); present as bifurcating complex network (2).
46. Ambulacra notches or lunules absent (0); present (1).
47. Anal lunule absent (0); present (1).
48. Pore-pairs double (0); single below the petals (1).
49. Subanal penicillate tube-feet absent (0); present (1).
50. Bourrelets and phyllodes absent (0); developed around the peristome (1).
51. Aboral tube-feet with thin walls and lacking suckered disc (0); with suckered disc (1).

(*d*) *Interambulacra*

52. Primary tubercles without clearly defined areole or surrounding ring of scrobicular tubercles (0); cidarid style, with sunken areole and a dense circle of scrobicular spines and tubercles (1); calycinid style, with scrobicular tubercles at corners of plates (2).
53. Both ambulacral columns border the peristome (0); only the primordial basicoronal plate borders the peristome (1).
54. All five interambulacral zones basically similar (0); the posterior interambulacrum strongly differentiated to form a labrum and sternum (1);
55. Just one large primary tubercle dominates each interambulacral plate (0); two or three large tubercles to each primary plate, often with the primary slightly larger (1); four to six, equal-sized tubercles to each plate (2); large numbers of very small, irregularly arranged tubercles (3).
56. Plate construction: outer layer composed of rectilinear stereom (0); entire plate composed of laminar stereom (1); composed of an outer layer of galleried stereom and an inner layer of labyrinthic stereom (2).
57. Plates abut without membranous gaps (0); membranous gaps developed between plates (1).
58. Naked, tubercle free zones absent (0); present down each interradius (1); present on oral surface along interambulacrum 5 (2).
59. More than two columns of interambulacral plates to each interambulacral zone (0); two columns only developed (1).
60. Subanal fasciole, when developed: diamond-shaped (1); bilobed (2).
61. Labral plate short and broad (0); long and narrow (1).
62. Aboral interradial suture imbricate (0); rigid (1).

(*e*) *Peristome*

63. Circular in outline (0); pentagonal, elongate laterally (1); D-shaped and opening towards the anterior (2).
64. Pharyngeal expansion sacs absent (0); present (1).
65. Peristome border smooth (0); notched for pharyngeal expansion sacs (1).
66. No tag associated with buccal notches (0); tags present (1).
67. Buccal notches, when present, relatively shallow, not penetrating more than to the second ambulacral plate (0); buccal notches deep, typically penetrating to the level of the fourth ambulacral plate (1).
68. Ambulacral plates migrate onto peristome throughout growth (0); no migration of ambulacral plates onto the peristome (1).
69. Multiple tube-feet per column on the peristomial membrane (0); single circlet of 10 buccal tube feet on the peristomial membrane (1); no tube feet on the peristomial membrane (2).
70. Peristome basically flush or only gently invaginated (0); with a pronounced well that bears modified spines that form a protective grill (1).
71. Mouth relatively large compared to the test diameter ($> 30\%$) (0); mouth relatively small in comparison to the test ($< 30\%$) (1).
72. Radial water vessel extends onto the peristome (0); not extending over peristome, but giving rise to lateral branches, each with an internal ampullae lying some distance from the buccal tube-foot (1).
73. Peristomial plates wedge-shaped and stacked imbricately (0); peristomial plates basically flat and ovoid (1).

(*f*) *Spines and tubercles*

74. Spines moderate in length, 0.5 or greater than test diameter (0); spines relatively short, 0.5–0.2 test diameter (1); spines very short (2).
75. Sphaeridia absent (0); present (1).
76. Sphaeridia in shallow pits (0); enclosed within the plate (1).
77. Sphaeridia situated along the perradial midline (o); close to the tube-feet along the adradial margin (1)
78. Sphaeridia: single (0); a pair (1); multiple (2).
79. Sphaeridia extending aborally to the apex (0); confined to the adoral region (1); extending to the ambitus (2).
80. Spines, hollow in cross-section (0); solid (1).
81. Spines ribbed, lacking a cortex layer (0); with a cortex layer (1).
82. Spines without cortical hairs (0); with cortical hairs (1).
83. Spines not verticillate (0); with verticillate arrangement of thorns (1).
84. Ambulacral and interambulacral spines broadly similar in size (0); ambulacral spines greatly reduced, typically forming a protective canopy of spatulate spines protecting the tube-feet (1).
85. General miliary spines without any glandular sac (1); terminating in a large glandular sac (1).
86. Scutellid boot-shaped spines absent (0); present (1).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

87. Aboral spines basically a heterogeneous armament (0); a uniform felt of dense spines (1);
88. Spines pointed or with a blunt end (0); spines spatulate-tipped (1).
89. Aboral tubercles with deeply sunken areole (0); with areole more or less flush to the surface (1).
90. Large defensive spines and sunken tubercles undifferentiated (0); scattered over aboral surface (1).
91. No specially differentiated oral spines for locomotion (0); plastronal spines differentiated to provide the main propulsive force (1).
92. Tubercles imperforate (0); perforate (1).
93. Tubercles non-crenulate (0); crenulate (1).
94. Primary ambulacral tubercle on every compound plate (0); on alternate compound plates only (1).
95. Modified mucous-producing miliaries (clavulae) absent (0); present(1).
96. Subanal fasciole absent (0); present (1).
97. Inner fasciole absent (0); present (1).
98. Peripetalous fasciole absent (0); present (1).
99. Anal branch to subanal fasciole: absent (0); present (1).
100. Spines without poison sac (0); ending in poison sac (1).
101. Spines without terminal hyaline hoof (0); ending in a hyaline hoof (1).
102. Secondary spines present (0); absent (1).

(*g*) *Pedicellariae*

103. Ophicephalous pedicellariae absent (0); present (1).
104. Triphyllous pedicellariae absent (0); present (1).
105. Globiferous pedicellariae absent (0); present (1).
106. Tridactylous pedicellariae absent (0); present (1).
107. Globiferous pedicellariae: poison gland external (0); internal, opening through a distal pore in the blade (1).
108. Globiferous pedicellariae: poison sacs undifferentiated, with valves surrounded by a thick skin (0); with clear glandular sacs on each valve (1).
109. Globiferous pedicellariae: poison glands associated with the valves only (0); carried on the stalk and valves (1).
110. Globiferous pedicellariae: without terminal fang (0); with terminal fang (1).
111. Globiferous pedicellariae: without ring of small teeth surrounding distal opening (0); with such a ring of teeth (1).
112. Globiferous pedicellariae: lacking lateral teeth on blade (0); with multiple paired lateral teeth (1); with single asymmetric tooth (2).
113. Globiferous pedicellariae: blade an open meshwork (0); a narrow fused cylindrical structure (1).
114. Globiferous pedicellariae: poison gland lacking muscles (0); with muscles (1).
115. Globiferous pedicellariae: without muscular neck (0); with long flexible muscular neck (1).
116. Globiferous pedicellariae: poison sacs single (0); double (1) on each valve.
117. Stalk of globiferous pedicellariae: solid rod (0); fused tubular cylinder (1); unfused mass of calcite fibres (2).
118. Tridentate pedicellariae: only with long narrow blades (0); with rostrate forms with bulbous blade (1).
119. Triphyllous pedicellariae without widened apophysis on blades (0); with widened apophyses forming a plate (1).
120. Bidentate pedicellariae absent (0); present (1).

(*h*) *Perignathic girdle*

121. Perignathic girdle absent (0); composed of apophyses (1); composed of auricles (2); composed of single interambulacral projection (3).
122. Protractor muscles attach to interambulacral plates (0); attach to ambulacral plates (1).
123. Retractor muscles attach to interambulacral plates (0); attach to ambulacral plates (1).
124. Perignathic girdle more or less pentaradially symmetric (0); asymmetric (1).

(*i*) *Lantern*

125. Present throughout life (0); present in juveniles, resorbed in adults (1); completely absent (2).
126. Hemipyramids: with very shallow (< 25% length) foramen magnum (0); with deep foramen magnum (*ca.* 30–40% length) (1); with no foramen magnum (2).
127. Hemipyramid without ala exterior (0); with expanded ala exterior (1).
128. Hemipyramid without ala interior (0); with expanded ala interior (1).
129. Hemipyramid not extremely flattened (0); extremely flattened and modified (1).
130. Hemipyramids without processi super alveolaris (0); with processi super alveolaris (1).
131. Compasses present (0); absent (1).
132. Rotulae: hinge-type articulation (0); ball and socket-type articulation (1).
133. Epiphyses small, scale-like (0); with prominent demi-arcs (1).
134. Epiphyses separate (0); meet and fuse above the hemipyramids and support the tooth (1).
135. Teeth shape in cross-section: U-shaped (cidaroid) (0); crescentic (diadematoid) (1); keeled (2); wedge-shaped (3).
136. Lantern symmetric (0); strongly asymmetric (1).
137. Central Lamellae-Needle-Prisms system (CLNP) and Lateral Lamellae-Needle-Prisms system (LLNP) undifferentiated (0); forming discrete systems (1).
138. Aboral portion of CLNP composed of cylindrical rods (0); flattened flabelliform leaves (1).
139. Oral portion of CLNP composed of obliquely orientated lamellae (0); composed of forks and/or tines perpendicular to edge(1); lamellae with forks oblique to edge (2).
140. Highly differentiated series of marginal tines absent (0); present forming a comb-like edge (1).
141. CLNP without upper layer of longitudinal lamellae (0); with such a layer (1).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

142. CLNP a single integrated system (0); composite, composed of an upper and lower layer (1).
143. Prisms relatively small and unspecialized (0); large and leaf-like (1).
144. LLNP system present (0); absent (1).
145. CLNP system with lamellae basically arranged in parallel (0); composed of a longitudinal band flanked by two transverse zones of lamellae (1).
146. Teeth with secondary plate (0); secondary plate absent, only primary plate developed (1).
147. Secondary tooth plate at stage III form a triangular lappet that extends along the median-lateral section of the primary tooth plate: median portion of central section narrow (0); extends along the distal portion of the central section of the primary tooth plate as well as along the median-lateral section: median lateral sections relatively large (1).
148. Secondary tooth plate with a carinal appendage (0); with a wing-like appendage (1); with pencil-like projections (2).
149. Carinal appendage on secondary tooth plate: long and thin (0); spatulate (1).
150. Primary tooth plates: median-lateral sections not separated by a sharp angle (0); separated by a sharp angle (1).
151. On primary tooth plates umbo-lateral margin distance less than umbo-oral distance (0); vica versa (1).
152. Aboral surface with a median groove (0); with a crest, not a groove (1).
153. Median section of tooth well developed (0); strongly reduced, with central and median sections forming a single unit (1).

(*j*) *Soft tissue*

154. Siphon absent from gut (0); present (1).
155. Siphon attached to the gut along its entire length (0); separated and free along its proximal third (1).
156. Secondary siphon absent whether or not primary siphon is developed (0); secondary siphon present in addition to primary siphon (1).
157. Mucous cells found in stomach wall (0); absent from stomach (1).
158. Mucous cells found in intestine wall (0); absent form intestine (1).
159. Caecum absent (0); present at start of stomach (1).
160. Stewart's organs absent (0); present (1).
161. Gut makes two complete circuits (0); makes only a single circuit (1).
162. Longitudinal body wall muscles absent (0); present (1).
163. Periproctal, perianal and genital sinuses remain discrete (0); fused into a single sinus (1).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

(*k*) *Character matrix* (*a–j*; 1–163)

node	1111111111111111111111111111111111111111111111111111111111111111 11111111112222222222333333333344444444445555555555666666666677777777778888888888999999999900000000001111111111222222222233333333334444444444555555555566666 1234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890123
Archaeocidaris	0010000000000000000000000000000000000000000000?10001000??00000?000000000???00001?0000001000000?000000??????????0000000000000000010??????????????????????????
Miocidaris	0010000000???0??0??00????00000000000000000000000000010000001??00000???00????????00001?000000110000000?00???????????????????000000000????0??????????011?0?0??????????
Cidaris	00000000000010000000000000000000000000000000010000001??10000?000000000???1110100000001000000000000101001100000000010000000000100000000000?0011?000000000001000
Eucidaris	00000000000010000000000000000000010000000000000000010000001??10000?000000000???1110100000001000000000000101000100000000010000000000100000000000?0011?000000000001000
Phormosoma	00100000000000000000010000101005100000000000000000000021001??00100000000001012000000000000010000000100010000??????????01020100100000000100000000?0012?0000000100101 0
Araeosoma	0010000000000000000000000101005100000000000000000000021101??00100000000001012000000000000010000000100101???????????1102010010001001010100010000012?0000000100101 0
Asthenosoma	0010000000000000000000010001010051000000000000000000000021101??00100000000001012000000000000010000000100101???????????1102010010001001010100010000012?0000000100101 0
Centrostephanus	00000000000020000000000000101001000000000000000000000012001??001110110011010120000100000000110000000001110000010110001102010010001001010110010000011?0000000100?000
Diadema	00000000000020000010000000101001000000000000000000000022001??001110110011010120000100000000110000000001100???????????01020100100010010101100100000011?00000001001000
Diademopsis	0000000000002000000000000001001000000000000000000000022001??1011001100?1010???00000?0000001000000?00????????????????????201001000000000101??????0?0?1?0000??????????
Eodiadema	000000000???0??0??00????00?1001000000000000000000000000002001??1011001?00??010???00000?0000001100000?00????????????????????20100100000???030?????????1???1?11??????????
Plesiechinus	100?00001011200000000000002001001000000000000000000000022001??1011001?00??1101??00000?0100001100000?00????????????????????2010010000000030?????????1???1?01??????????
Echinocyamus	10001000111130001002100000000?00000210000100000000100001032001??10000?12110?21100100000000101001100000000001100???????????000300102100010003111100?01101???1011100110 0101
Fellaster	11001000111130001000000000?101502100001000110010000103200 1??10000?12110?21101100000000101001100000000000100???????????000211102100010003 1?100?01101???10111001100101
Dendraster	1100100011113000100210000000?000002100001001020010000103200 1??10000?12110?21100100000111010011000000000001100???????????000300102110010003 1??00??1??1???10111001100101
Echinodiscus	11001000111130001002100000000?00000210000100102101000010320 01??10000?12110?21100100000111010011000000000001100???????????000300102111010003 1??00??1??1???10111001100101
Encope	11001000110130001002100000000?000002100001001021110000103200 1??10000?12110?21100100000111010011000000000000100???????????00130010211001000 31??00??1??1???10111001100101
Echinolampas	100000001111300010020000010?100502100000000000001010010320 21??11000?12110?21102100000001010011000000000001100???????????000201111000010003111100?01101???10111001100001
Cassidulus	100000001111300010010000010?100502100000000000001010010320 21??11000?12110?21102100000001010011000000000001100???????????000201111000010003111100?01101???10111001100001
Brissopsis	10000000111114100100100001 00?0000021110101100000110001132001 2112000?12010?21012100000001100111011010000111000001011000100 0???2???????????????????????????????1111110001
Meoma	1000000011111410010010000100?00000211001010000001100011320 01101 2000?12010?21012100000001100111011010000111000001011000100 0???2???????????????????????????????1111110001
Echinocardium	1000000011111410010010000100?0000021011101100000110001132001 1012000?12010?21012100000001100111011110100011100000101100010 00???2???????????????????????????????1101110001
Spatangus	1000000011111410010010000100?0000021011101000001100011320 01201 2000?12010?21012100000001101111011000000111000001011000100 0???2???????????????????????????????1101110001
Pseudodiadema	00000000000000?00000000000000010010000000000000000000000002011??1011001100??01????10000?00000011000000000????????????????????20100100010010 20?????????0???0?00??????????
Acrosalenia	00000000000000?000000000000010010000000000000000000000002000 2001??1011001100?101012?11001?0000001100000 0?00????????????????????2010010001001020?????????0???0?00??????????
Acrosalenia	0000000000002011100000000000010010000000000000000000020002001??1011001100?101012?11001?000000110000000?00????????????????????20100100010010 20?????????0???0?00??????????
Salenia	0000000000002011100000001000010110000000000000000000020002001??10110011001101012211001000000001000000000 1100???????????00020100100010010201100100000100000001001100000
Gymnocidaris	000000000000010?01?0000000001100101000000000000000000000020002001??1011001?00??0????11001?0000001100000?00????????????????????20100???
Glypticus	0000001000001000 1?00000000001100101000000000000000000000020002001??1011001?00??0????11000?00000000000000?00????????????????????20100???
Arbacia	000100100000100011000000000011003010000000000000000000010220 01??1011101100110100011100000000000000000000011100???????????0002010010001001020112001000000000001001100000
Phymosoma	0000000000000?000000000000001104000000000000000000000100012001??1011001100?101012?11000?000000010000000?00????????????????????20100100010010 20?????????0???0?00??????????
Glyptocidaris	00000000000020?0100000000000111060000000000000000000000100012001??1011001100111101 22110000000000000100000000111001010011110002010010001001020110010000010000000100110?000
Stomopneustes	0000000000002000100000000000111060000000000000000000000100012001??10110011001111012010000000000000000000000011100001011000000 2010010001001020110010000010000000100 1101000
Glyptocyphus	000?010000000201000000000000100200000000000000000000000100012001??1011001100????????1?2000?00000001000000000??????????????????201001000?????????????????????????????????????
Temnopleurus	0001010000001010000000000000110020000000000000000000000100012001??101100110011110022100000000000010000000001110010101110120002010010000001120110000000001010000100110 0000
Mespilia	00010100000010000000000000001100200000000000000000000000100012011??101100110011110022100000000000010000000001110010101010120002010010000001120110000000001010000100110 0000
Salmacis	000101000000101000000000000011002000000000000000000000000100022001??1011001100111100221000000000000010000000001110010102110120002010010000001120110000000001010000100110 0000
Echinus	00000000000010100000000000001100200000000000000000000000100012001??101100110011110021100000000000000100000000111001010111022000201001000000112011100000000001000010011 00000
Psammechinus	0000000000001010000000000000110020000000000000000000000100012001??10110011001111002110000000000000000000000111001010101022000201001000000112011100000000001000010011 00000
Paracentrotus	000000000000101000000000000011102000000000000000000000000100012001??10110011001111002110000000000000000000000111001010101022000201001000000112011100000000001000010011 00000
Lytechinus	00000000000020000000000000001100200000000000000000000000100022001??10110111001111002110000000000000000000000111001110010210002010010000001120111100001000101001001100000
Sphaerechinus	000000000000201000000000000011102000000000000000000000000100022001??10110111001111002110000000000000000000000111001110010210002010010000001120111100001000101001001100000
Tripneustes	0000000000002000000000000000110020000000000000000000000100022001??101101110011110021100000000000001000000000111001110010210002010010000001120111100000000101001001100000
Colobocentrotus	20000000000010100000000000001130200000000000000000000000100012001??1011001100112100211000000000000000000000000111001110211010000201001000000112011110000100010100100110 0000
Heliocidaris	00000000000020100000000000001130200000000000000000000000100012001??1011001100112100211000000000000000000000000111001110211010000201001000000112011110000100010100100110 0000
Strongylocentrot	00000000000020100000000000001110200000000000000000000000100012001??1011001100111100211000000000000000000000000111001010011120000201001000000112011110000100010100100110 0000

APPENDIX 2

ssu rRNA molecular data matrix. Full alignment of complete small subunit rRNA gene sequences has been deposited with EMBL under accession number DS19161 and is available via FTP. Full species names appear in Table 1; dashes (–) indicate alignment gaps and are treated as fifth characters; IUPAC codes are used throughout; ? indicates unknown. Individual ssu rRNA sequences determined in this study have been deposited with EMBL under accession numbers Z37118–Z37135, Z37140–Z37149, Z37514.

```
Eucidaris  UCCAAA-UGU-CCA-UGUCUU-GCCAC-ACA-CCUCGUCUACGGCGUUCCAAAGA-UACCGGGAUUUUCGCGGGA-AUA-A-UGGAGCAGAUAG--GUUAG
Asthenosoma UGUCGGUGU-CCA-UGCAUUCCCCGA-GCAACCCC----------GUUACC-A-UCUACGUGCUUCGCGCGG-AAA-A-UAAGGCGGGUAA-GACUGC
CentrostephUUCAAAGUGU-CCA-UGCCUU-GCCAC-ACA-CCUU----------GCCUCU-ACUACCCGGAUUUCGCGCGA-AUA-A-UGGAGCAGAUAG--GUUAG
Diadema    UUCAAAGUGU-CCA-UGUCUU-GCCAC-ACA-CCUU----------GCCACU-ACUACCCGGAUUUCGCGCGA-AUA-A-UGGAGCAGAUAG--GUUAG
Cassidulus UUCAAAGUUU-CCA-UGCCUU-GUCA-AACA-CUCU----------GCCAAC-A-UACCCGGAUUUCGCGCGA-AUACA-UGGAGCAGAUAG--GUUAG
Fellaster  UUCAAAGCUU-GCA-UGCCUU-GUCC-AACA-CUCU----------GCCAAC-A-UACUUGGAUUUCGCGCGA-AUA-A-UGGAGCAGAUAG--GUUAG
Encope     UUCAAAGUGU-CCA-UGCCUU-GUCC-AACA--UCU----------GCCAAA-A-UACCUGGAUUUCAUGCGA-AUA-A-UGGAGCAGAUAG--GUUAG
Echinodiscu UUCAAAGUUU-CCA-UGCCUU-GUCC-AACA-CUCU----------GCCAAA-A-UACCCGGAUUUAGCGCGA-AUA----UGGAUUAGAUAG-GGUUAG
Echinocardi UUCAAAGUGU-CCA-UGCCUU-GUCC-AACA-CCCU----------GCCAAC-A-UACUUGGAUUUCGCGCGA-AUG-A-UGGAGCAGAUAG--GUUAG
Brissopsis UUCAAAGUGU-CCA-UGCCUU-GUCC-AACA-CCCC----------GCCAAC-A-UACCCGGAUUUCGCGC-A-AUA-AUUGGAGCAGAUAGU-GUUAG
Meoma      UUCAAAGUGU-CCA-UGCCUU-GUCC-AACA-CCCC----------GCCAAC-A-UACCCGGAUUUCGCGCGA-AUA-A-UGGAGCAGAUAG--GUUAG
Stomopneust UUCAAAGUGU-CCA-UGCCUU-GUCCCAACA-CUCU----------GCCAAC-A-UACUCUGAUUUCGCGCGAA-UA-A-UGGAGCAUAUAG-GGUUAG
Arbacia    UUCAAAGU-U-C-A-UGCCUU-GCCCCAACA-CUUU----------GCCACC-A-UACUCGGAUAUCGCGCGA--UA---UGGAGCAGAUAG--GUUAG
Temnopleuru CUCACAGU-U-C-A-UGCAUC-GCCCCAACG-CCCU----------GCACAC-G-CACUCGGGUUUCCGCGCGA-RUA---UGGAGCAGACAG--GUUAG
Salmacis   UUCAAAGU-UACCG-UCCCUC-GCCCCAACA-CUCU----------GCACAC-G-UACCCGGGUUUCGCGCGA-AUA-A-UGGAGCAGAUAG--GUUAG
Mespilia   UUCAAAGU-UACCA-CGCUUC-GCCCCAACA-CCCU----------GCACAC-G-UACACGGGUUUCGCGCGA-AUA-A-CGGAGCAGAUAG--GUUAG
Psammechinu UUCAAAGUGU-UCA-UGCCUC-GUUCCAACA-CUCU----------GCCCAC-A-UACUUGGAUUUCGCGCGA-AUA-A-UGGAGCAGAUAG--GUUAG
Echinus    UUCAAAGUGU-UCA-UGCCCU-GUUCCAACA-CUCU----------GCCCAC-A-UACUUGGAUUUCGCGCGA-AUA-A-UGGAGCAGAUAG--GUCAG
Sphaerechin UUCUAAGU--ACCG-UGCCUU-GCUCCAACA-CCUC----------GCUACC-G-UACCCGGAUUUCGCGCGA-AUA-A-UGGAGCAGAUGG--GUUAG
Tripneustes UUCAAAGUGU-CCA-UGCCUU-GUGCCAAAA-CUCU----------GCCCAC-A-UACCCAAAUUUCGCGCGA-AUA-A-UGGAGCAGAUAG--GUUAG
Colobocentr UUCAAAGUGU-UCA-UGCCUU-GCUCCAACA-CUCU----------GCACAC-A-UACCCGGAUUUCGCACGA-AU?-A-UGGAGCAGAUAG--GUUAG
Strongyloce???????????-???GUGCCUU-GUUCCAACA--UCU----------GCCCAC-A-UACUUGGAUUUCGCGCGA-AUA-A-UGGAGCAGAUAG-GGUUAG

Eucidaris  CCGCU-CAAUGUUUCC-UAAUGA-AGG-AAU--UCGU-UUCG-GAUAG-GCACCU?2UAACCGCCGGCGUCGGUUGACUUGAAG-GAAACGCUGCCGCGCG
Asthenosoma CUGCG-C-CUGGCU-C-CAAUGG-AUA-AAUA-UCGC-ACCG-GAUAA-GAACAU?2UAACUAGCGGCGUCUGUCC-GUU-AAG-GAAGC-CGUCCCGCG
CentrostephCUUUU-C-AGGUUU-C-UAACGA-AGG-AAU--UCGC-UUCG-GACAG-GCCCCU?2UAACCGCCGGCGUCGGUUAACUU-AAG-AGAACGCUUCCCGCG
Diadema    CUGUU-C-AGGUUU-C-UAACGA-AGG-AAU-?UCGC-UUCG-GAUAG-GCACCU?2UAACCGCCGGCGUCGGUUAACUU-AAG-AAAACGCUUCCCGCG
Cassidulus CCG-U-C-AUGUUU-C-UAACGA-UAG-AAU--UCGC-UUCG-GAUAG-GCACCU?2UAACCGCCGGCGUCG-UUAACUU-AAG-AAAACGCUUCCCGCG
Fellaster  UUGGU-U-AUAUUU-U-UAACGA-UAG-AAU--U-GC-UUCG-GAUCG-GCACCU?2UAGCCGCCGGCGUCGGUUAACUU-AAG-AAAACGCUUCCCGCG
Encope     CCG-U-U-AUGUUU-U-UAACGA-UAG-AAU--UCGCGUUCG-AAUAG-GCACCU?2UAACCGCCGGCGUCGGUUAACUU-AAG-AAAACGCUUCCCGCG
Echinodiscu CCG-U-C-AUGUUU-U-UAGCGA-UAG-AAU--CCGC-UUCA-GAUAG-GCACCU?2UGACCGCCGGCGUCGGUUAACUU-AAG-AAAACGCUUCCCGCG
Echinocardi CCG-U-C-AUGUUU-C-CAACAA-AAG-AAU-AUCGC-UUCG-GAUAG-GCACCU?2UAACCGCCGGCGUCGGUUAACUU-AAG-AAAACGCUUCCCGCG
Brissopsis CCGCUUC-AUGUUU-C-UAAUGA-AAG-AAU--UCUC-UUCG-GAUAG-GCACCU?2UAACCGCCGGCGUCGGUUAACUU-AAG-AAAACGCUUUCCGCA
Meoma      CCG-U-C-AUGUUU-C-UAACGA-AAG-AAU--UCGC-UUCG-GAUAG--CACCU?2UAACCGC------CGGUUAACUU-AAG-AAAACGCUUCCCGCG
StomopneustCCG-U-C-AUGUUA-C-CAGUGA-AAG-AAU--UCGC-UUCG-GAUAG-GCACCU?2UAACCGCCGGCGUCGGUUGACUU-GAG-AAAACGCUUCUCGCG
Arbacia    CCG-U-C-AUGUUU-C-UAACGA-AAG-AAU--UCGC-UUCGAGAUAG-GCACCA?2UAAUCGCCGGCGUCGGUUAACUU-AAG-AA-ACGCUUCCCGCG
Temnopleuru CCGGG-C-CUGUCU-C-CGAUGGUAAGC--U--UCGC-UUCG-GGUAG-GCACCA?2UAAUCGCCGGCGUCGGCCCCCUU-AAG-GAAACGUUUCCCGCG
Salmacis   CCG-G-C-CUGUUU-C-CAGUGA-AAG-AAU--UCGC-UUCG-GAUAG-GCACCU?2UAACCGCCGGCGUCGGUUGACUU-AAG-AAAAUGCUUCCCGCG
Mespilia   CCG-G-C-CUGUUU-C-CGAUGGUAAG-AAU--UCGC-UUUG-GAUAGAGCACCU?2UAACCGCCGGCGU-GGUUGACAA-AAGGAAAACGCUUCCCGCG
Psammechinu UCGCU-C-UUGUUU-C-UAAUGA-AAG-AAC-AUCGC-UUCG-GAUAG-GCACCU?2UAACCGCCGGCGUCGGUUAACUU-AAG-AAAACGCUUCCCGCG
Echinus    CCG-U-C-AUGUUU-CUCAAUGA-AAG-AAU--UCGC-UUCG-GAUAG-GCACCU?2UAACCGCCGGCGUCGGUUAACUU-AA--AAAACGCUUCCCGCG
Sphaerechin CCG-U-C-AUGUUU-C-UGAUGA-AAG-AAU--UCGC-UUUG-GAUAG-GCACCU?2UAACCGCCGGCGUCGGUUAACUU-AAG-AAAACGCUUCCCGCG
Tripneustes CCG-U-C-AUGUUU-C-CGAUGA-AAG-AAU--?CGC-UUUG-GAUAG-GCACCU?AUAACCGCCGGCGUCGGUUGACUU-AAG-AAAACGCUUCCCGCG
Colobocentr CCG-U-C-AUGUUU-C-CAAUGA-AAG-AAU--UCGC-UUCG-GAUAG-GCACCUCGCAACCGCCGGCGUCGGUUAACUU-AGG-AAAACGCUUCCUGCG
Strongyloce CCG-U-C-AUGUUU-C-CAAUGA-AAG-AAU--UCGC-UUCG-GAUAG-GCAACUGCUAACCGCCGGCGUCGGUUAACUU-AAGGAAAACGCUUCCC--GG
```

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES


```
Eucidaris  -UCAGCUCCCACUACACUGCCGGUAUCU-AAUAUUUG-CCGUCCUCGGUCUCCGCGCACGUC-GCGA-G-ACGUUACGU 279
Asthenosoma-UAAGAUCCUAUCACACUGCUGGUGCUG-ACUAUUUGUAC-UUCCCUGA-----C-CA--UU-G------ACAUAACGU
Centrosteph-UAAGAUCCCACUACACUGCCAGUAUCU-AAUGUUUG-CCGUCGUCGGACUCUG-CCUCGCC-GCAU-G-ACGAUAC??
Diadema    -UAAGAUCCCACUACACCGCCAGUAUCU-AAUAUUCG-CCGUCGUCGGACUCUG-CCUCGCC-GCAU-G-ACGUUAC??
Cassidulus -UAAGAUCCCACUACGCUUCCGGUAUCU-AAUAUUUG-CCGGCGUCGAACUCUGCUCUCGUCUGCAU-GCCCGUUGCGU
Fellaster  -UAAGAUCCCACUAAACUGCCGGUAUCU-AAUAUUUG-CCGUCGUCGGACUCUGCUCUCGUCUGCAUGGUACGUUAU??
Encope     -UAAGAUCCCACUAUCUCUCCGGUAUCU-AAUAUCUG-CCGGCGAAGGG-UCU-CUCUC-UCUGCAU-GCUCGUUACGU
Echinodiscu-UAAGAUCCCACUAUUUUUCCGGUAUCU-GAUAUUUG-CCGGCGGCGGACCCUGCUCUCGCCUGCAC-GCCCGUUACGU
Echinocardi-UAAGAUCCCACUCUUGUU-CGGUAUUG-AAUAUUUG-CCGUCGUCGGACUCUGCUCUCGUCUGCAU-G-ACGUUACGU
Brissopsis CUAAGAUCCCACUCUUGUU-CGGUAUCUCAAUAUUUG-CCGUCGUCGGACUCUGCUCUCGUCUGCAU-G-ACGUUAC??
Meoma      -UAAGAUCCCACUCUUGUU-CGGUAUCU-AAUAUUUG-CCGUCGUCGGACUCUGCUCUCGUCUGCAU-G-ACGUUACGU
Stomopneust-UAAAAUCCCACUAUCUCUCCGGUAUCU-AAUAUUUG-CCGUCGUCGGAUUCU-CCCUCGCC-GCAU-GCACGUUACGU
Arbacia    -UAAGAUCCCACUACGCUGCCGGUAUCU-AAUAUUUG-CCGUCGUCGGACUCCGCUCACGAC-GCAUUGUACGUUACGU
Temnopleuru-CAAGACCCCACUACGCUGCCGGUAUCU-AAUAUUUG-CUGUCGUCGGACUCCGCGCCCGCC-GCAU-GCACGUUACCG
Salmacis   -CAAGAUCCCACUACACUGCCGGUAUCU-AAUAUUUG-CUGUCGUCGGACUUCG-CCCCGCC-GCGUUG-ACGUUACGU
Mespilia   -CAAGAUUCCACUACACUGCCGGUAUCU-AAUAUUUA-CUGUCGUCGGACUCCG-CCCCGCC-GCAU-GCACGUUACGU
Psammechinu-UAAGAU?CCACUACACUGCCGGUAUCU-AAUAUUUG-CCGUCGUCGGACUUCGCUGCCGCC-CCAU-GUACGUUACGU
Echinus    -UAAGAUCCCACUACACUGCCGGCAUCU-GACAUUUG-CCGUCGUCGGACUUCG-CCUCGCC-GCAU-GUACGUUAC??
Sphaerechin-CAGGAU?CCGCUACACC??CG-UAUCU-AAUAUUUG-CCGUCGUCGGACUCCG-CCUCGCC-GCAU-GCACGUUACGU
Tripneustes-UAAGAUCCCACUACACCGCCGGUAUCU-AAUAAUUG-CCUUCGUCGGACUUCG-CCUCGCC-GCAU-GCAGGUUACGU
Colobocentr-UAAGAUCCCACUACACUGCCGGUAUCU-AAUAUUUG-CCGUCGUCGGACUCCG-CCUCGCC-GCAU-GUACGUUAU??
Strongyloce-UAAGUUG-CACUACACUGCCGGUAUCU-GACAUUUG-CCGUCGUCG-ACUUUGCCCUCGCC-GCAU-GUACG??????
```

APPENDIX 3

LSU rRNA molecular data matrix. Full alignment of partial large subunit rRNA gene sequences has been deposited with EMBL under accession number DS19161 and is availabe via FTP. Full species names appear in table 1; dashes (–) indicate alignment gaps and are treated as fifth characters; IUPAC codes are used throughout; ? indicates unknown. Individual LSU rRNA sequences determined in this study have been deposited with EMBL under accession numbers Z37116–Z37117, Z37507.

```
Cidaris        AGCAGUUCCGAGUGGA--GACCGGCUCGGGGGUGGCGGUCC--AACCGCUCACAG-CC?GGGGUGU-CGC-GAAAGUUAC---G----U-UG 92
Asthenosoma    ??CGCGUCAGUAUAGCC-G-CCGGC------GUUCCG-UUCUUGACUCAUUUUGUGGCCCAGGAAU-GGAGGAGUGAUCCAA-GUGGAC-GA
Araeosoma      ??CAAGUCAGUGUAGCC-G-GCGGC------GUUCCG-UUCUUGACUAACUUUGUGGCCCAG?CGU-CGAGUGGUGAGUCAAAGUUUACAGA
Echinocyamu    AGCACUCUCGAGCAAAGUGACCG?CCUCGGGGCGGCUGUUC--GGUC-ACUUCAGGGCCCGAGCAA-CGCGGAAAGAUUC---G----U-GA
Encope         ??CACUUCC-AUUGGGGCGACCGACCCC-GGGUGGCUGUUC--AGUC-ACUUCAGGGCCCGAGCAA-CGCGGAAAGAUUC---G----U-GA
Spatangus      AGCACUUCCGAGUACAG-GACCGGUUGCGGGGUGGCGGUCC--AACC-ACUUCAG-G?CCGGGUGU-CGCGGAAAGAUUC---C----U-GA
Echinocard     AGCACUUCCGAGUACAG-GACCGGUUGCGGGGUGGCGGUCC--AACC-ACUUCAG---CCGGGUGU-CGCGGAAAGAUUC---C----U-GA
Arbacia        AACACUCUCGU?C?GA--GACCGGCUCGGGGCCGAUGGGUU--AGCC-ACUUCAGGGCCCGACC-ACCGCAGAAAAAUUU---G----U-GA
Sphaerechin    GACACUCCCGAGCAGCG-CUC-AGCUCUGGGGCGGCGGUCC--AACC-ACUUCAG-GCGCGGGCGA-CAAGGAAAGAUUC---G----U-GA
Lytechinus     GACACUCCCGAGCAGCG-CUC-AGCUCUGGGGCGGCGGUCC--AACC-ACUUCAG-GCGCGGGCGA-CAAGGAAAGAUUC---G----U-GA
Psammechinu    AACACUUCCGAGCAGCG-UUC--GUCCUGGGCCGGCGGUCC--AACC-ACUUCAG--CCCGGGUGU-CAAGGAAAGAUUC---A----U-GA
Echinus        AAUAUUUCCGAGCAGCG-UUC--GUCCUGGGCCGGCGGUCC--AACC-ACUUCAG-GCGCGGGUGU-CAAGGAAAGAUUC---A----U-GA
Paracentrot    AACACUUCCGAGCAGCG-UUC--GUCCUGGGCCGGCGGUCC--AACC-ACUUCAG-GCGCGGGUGU-CAAGGAAAGAUUC---A----U-GA
```


PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES