Phylogenomic analyses of echinoid diversification prompt a re-evaluation of their fossil record

Nicolás Mongiardino Koch[1,2]*, Jeffrey R Thompson[3,4], Avery S Hatch[2], Marina F McCowin[2], A Frances Armstrong[5], Simon E Coppard[6], Felipe Aguilera[7], Omri Bronstein[8,9], Andreas Kroh[10], Rich Mooi[5], Greg W Rouse[2]

[1] Department of Earth & Planetary Sciences, Yale University, New Haven CT, USA

[2] Scripps Institution of Oceanography, University of California San Diego, La Jolla CA, USA

[3] Department of Earth Sciences, Natural History Museum, Cromwell Road, SW7 5BD London, UK

[4] University College London Center for Life's Origins and Evolution, London, UK

[5] Department of Invertebrate Zoology and Geology, California Academy of Sciences, San Francisco CA, USA

[6] Bader International Study Centre, Queen's University, Herstmonceux Castle, East Sussex, UK

[7] Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias Biológicas, Universidad de Concepción, Concepción, Chile

[8] School of Zoology, Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel

[9] Steinhardt Museum of Natural History, Tel-Aviv, Israel

[10] Department of Geology and Palaeontology, Natural History Museum Vienna, Vienna, Austria

** Corresponding author. Email: nicolas.mongiardinokoch@yale.edu*

Abstract

Echinoids are key components of modern marine ecosystems. Despite a remarkable fossil record, the emergence of their crown group is documented by few specimens of unclear affinities, rendering their early history uncertain. The origin of sand dollars, one of its most distinctive clades, is also unclear due to an unstable phylogenetic context. We employ seventeen novel genomes and transcriptomes to build a dataset with a near-complete sampling of major lineages and use it to revise the phylogeny and divergence times of echinoids. We introduce the concept of a chronospace—a multidimensional representation of node ages—and use it to explore the effects of using alternative gene samples, models of molecular evolution, and clock priors. We find the choice of clock model to have the strongest impact on divergence times, while gene sampling and the use of site-heterogeneous models show little effects. Crown group echinoids originated in the Permian and diversified rapidly in the Triassic. We clarify the relationships among sand dollars and their close relatives, showing that *Apatopygus* represents a relict lineage with a deep Jurassic origin. Sand dollars are confidently dated to the Cretaceous, implying ghost ranges spanning approximately 50 million years, a remarkable discrepancy with their fossil record.

Keywords: Echinoidea, Sea urchins, Sand dollars, Phylogenomics, Time calibration, Site heterogenous models

Background

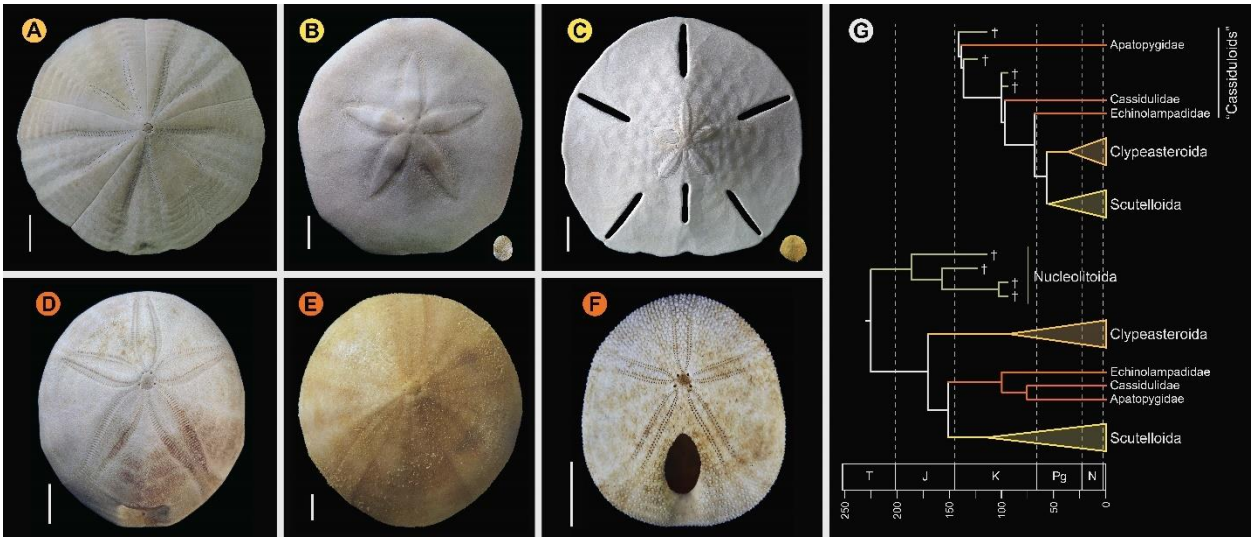
The fossil record represents the best source of primary data for constraining the origins of major lineages across the tree of life. However, the fossil record is not perfect, and even for groups with an excellent fossilization potential, constraining their age of origin can be difficult [1, 2]. Furthermore, as many traditional hypotheses of relationships have been revised in light of large-scale molecular datasets, the affinities of fossil lineages and their bearings on inferred times of divergence have also required a reassessment. An exemplary case of this is Echinoidea, a clade comprised of sea urchins, heart urchins, sand dollars, and allies, for which phylogenomic trees have questioned the timing of previously well-constrained nodes [3, 4].

Echinoids are easily recognized by their spine-covered skeletons or tests, composed of numerous tightly interlocking plates. Slightly over 1,000 living species have been described to date [5], a diversity that populates every marine benthic environment from intertidal to abyssal depths [6]. Echinoids are usually subdivided into two morpho-functional groups with similar species-level diversities: “regular” sea urchins, a paraphyletic assemblage of hemispherical, epibenthic consumers protected by large spines; and irregulars (Irregularia), a clade of predominantly infaunal and bilaterally symmetrical forms covered by small and specialized spines. In today’s oceans, regular echinoids act as ecosystem engineers in biodiverse coastal communities such as coral reefs [7] and kelp forests [8], where they are often the main consumers. They are first well-known in the fossil record on either side of the Permian-Triassic (P-T) mass extinction event when many species occupied reef environments similar to those inhabited today by their descendants [9, 10]. This extinction event was originally thought to have radically impacted the macroevolutionary history of the clade, decimating the echinoid stem group and leading to the radiation of crown group taxa from a single surviving lineage [11, 12]. However, it is now widely accepted that the origin of crown group Echinoidea (i.e., the divergence between its two

61 main lineages, Cidaroidea and Euechinoidea) occurred in the Late Permian, as supported by molecular
62 estimates of divergence [13, 14], as well as the occurrence of Permian fossils with morphologies typical
63 of modern cidaroids [15, 16]. However, a recent total-evidence study recovered many taxa previously
64 classified as crown group members along the echinoid stem, while also suggesting that up to three
65 crown group lineages survived the P-T mass extinction [3]. This result increases the discrepancy between
66 molecular estimates and the fossil record and renders uncertain the early evolutionary history of crown
67 group echinoids. Constraining the timing of origin of this clade relative to the P-T mass extinction
68 (especially in the light of recent topological changes [3, 4]) is further complicated by the poor
69 preservation potential of stem group echinoids, and the difficulty assigning available disarticulated
70 remains from the Late Palaeozoic and Early Triassic to specific clades [11, 12, 17-20].

71 Compared to the morphological conservatism of regular sea urchins, the evolutionary history of
72 the relatively younger Irregularia was characterized by dramatic levels of morphological and ecological
73 innovation [21-24]. Within the diversity of irregulars, sand dollars are the most easily recognized (figure
74 1). The clade includes greatly flattened forms that live in high-energy sandy environments where they
75 feed using a unique mechanism for selecting and transporting organic particles to the mouth, where
76 these are crushed using well-developed jaws [25, 26]. Sand dollars (Scutelloida) were long thought to be
77 most closely related to sea biscuits (Clypeasteroida) given a wealth of shared morphological characters
78 [19, 25]. The extraordinary fossil record of both sand dollars and sea biscuits suggested their last
79 common ancestor originated in the early Cenozoic from among an assemblage known as “cassiduloids”
80 [23, 25], a once diverse group that is today represented by three depauperate lineages: cassidulids (and
81 close relatives), echinolampadids, and apatopygids [19, 27]. These taxa not only lack the defining
82 features of both scutelloids and clypeasteroids but have experienced little morphological change since
83 their origin deep in the Mesozoic [24, 27-29]. However, early molecular phylogenies supported both
84 cassidulids and echinolampadids as close relatives of sand dollars (e.g., [14, 30]), a topology initially

disregarded for its conflicts with both morphological and paleontological evidence, but later confirmed using phylogenomic approaches [4]. While many of the traits shared by sand dollars and sea biscuits have since been suggested to represent a mix of convergences and ancestral synapomorphies secondarily lost by some “cassiduloids” [3, 4], the strong discrepancy between molecular topologies and the fossil record remains unexplained. Central to this discussion is the position of apatopygids, a clade so far unsampled in molecular studies. Apatopygids have a fossil record stretching more than 100 million years and likely have phylogenetic affinities with even older extinct lineages [3, 19, 28, 29]. Although current molecular topologies already imply ghost ranges for scutelloids and clypeasteroids that necessarily extend beyond the Cretaceous-Paleogene (K-Pg) boundary, the phylogenetic position of apatopygids could impose even earlier ages on these lineages (figure 1). Constraining these divergences is necessary to understand the timing of origin of the sand dollars, one of the most specialized lineages of echinoids [24-27]. Resolving some phylogenetic relationships within scutelloids has also been complicated by their recurrent miniaturization and associated loss of morphological features (figure 1; [25, 31, 32]).



101 **Figure 1:** Neognathostomate diversity and phylogenetic relationships. **A.** *Fellaster zelandiae*, North
102 Island, New Zealand (Clypeasteroida). **B.** Large specimen: *Peronella japonica*, Ryukyu Islands, Japan;
103 Small specimen: *Echinocyamus crispus*, Maricaban Island, Philippines (Laganina: Scutelloida). **C.** Large
104 specimen: *Leodia sexiesperforata*, Long Key, Florida; Small specimen: *Sinaechinocyamus mai*, Taiwan
105 (Scutellina: Scutelloida). **D.** *Rhyncholampas pacificus,* Isla Isabela, Galápagos Islands (Cassidulidae). **E.**
106 *Conolampas sigsbei*, Bimini, Bahamas (Echinolampadidae). **F.** *Apatopygus recens*, Australia
107 (Apatopygidae). **G.** Hypotheses of relationships among neognathostomates. Top: Morphology supports
108 a clade of Clypeasteroida + Scutelloida originating after the K-Pg boundary, subtended by a paraphyletic
109 assemblage of extant (red) and extinct (green) "cassiduloids" [19]. Bottom: A recent total-evidence
110 study split most of cassiduloid diversity into a clade of extant lineages closely related to scutelloids, and
111 an unrelated clade of extinct forms (Nucleolitoida) [3]. Divergence times are much older and conflict
112 with fossil evidence. Cassidulids and apatopygids lacked molecular data in this analysis. Scale bars = 10
113 mm.

114

115 Echinoidea constitutes a model clade in developmental biology and genomics. As these fields
116 embrace a more comparative approach [13, 33, 34], robust and time-calibrated phylogenies are
117 expected to play an increasingly important role. Likewise, the extraordinary fossil record of echinoids
118 and the ease with which echinoid fossils can be incorporated in phylogenetic analyses make them an
119 ideal system to explore macroevolutionary dynamics using phylogenetic comparative methods [3, 31]. In
120 this study, we build upon available molecular resources with 17 novel genome-scale datasets and build
121 the largest molecular dataset for echinoids yet compiled. Our expanded phylogenomic dataset extends
122 sampling to 16 of the 17 currently recognized echinoid orders (plus the unassigned apatopygids) [35], is
123 the first to bracket the extant diversity of both sand dollars and sea biscuits, and include members of all
124 three lineages of living "cassiduloids" (cassidulids, echinolampadids, and apatopygids). We also

125 incorporate a diverse sample of outgroups, providing access to the deepest nodes within the crown
126 groups of all other echinoderm classes (holothuroids, asteroids, ophiuroids, and crinoids). With it, we
127 reconstruct the phylogenetic relationships and divergence times of the major lineages of living echinoids
128 and place their diversification within the broader context of echinoderm evolution.

129

130 ## Methods

131 ### (a) Sampling, bioinformatics, and matrix construction

132 This study builds upon previous phylogenomic matrices [3, 4], the last of which was augmented
133 through the addition of eight published datasets (mostly expanding outgroup sampling), as well as 17
134 novel echinoid datasets (15 transcriptomes and two draft genomes). Final taxon sampling included 12
135 outgroups and 50 echinoids. For all novel datasets, tissue sampling, DNA/RNA extraction, library
136 preparation, and sequencing varied by specimen, and are detailed in the supplementary information.

137 Raw reads for all transcriptomic datasets were trimmed or excluded using quality scores with
138 Trimmomatic v. 0.36 [36] under default parameters. Further sanitation steps were performed using the
139 Agalma 2.0 phylogenomic workflow [37], and datasets were assembled *de novo* with Trinity 2.5.1 [38].
140 For genomic shotgun sequences, adapters were removed with BBDuk
141 (https://sourceforge.net/projects/bbmap/), and UrQt v. 1.0.18 [39] was used to filter short reads (size <
142 50) and trim low-quality ends (score < 28). Datasets were then assembled using MEGAHIT v. 1.1.2 [40].
143 Draft genomes were masked using RepeatMasker v. 4.1.0 [41, 42], before obtaining gene predictions
144 with AUGUSTUS 3.2.3 [43]. A custom set of universal single-copy orthologs (USCOs) obtained from the
145 latest *Strongylocentrotus purpuratus* genome assembly (Spur v. 5.0) was employed as the training

146 dataset. Further details of these analyses, as well as statistics for raw reads and assemblies can be found
147 in tables S1 and S2.

148 Multiplexed transcriptomes were sanitized from cross-contaminants using CroCo v. 1.1 [44], and
149 likely non-metazoan contaminants were removed using alien_index v. 3.0 [45] (removing sequences
150 with AI values > 45). Datasets were imported back into Agalma, which automated orthology inference
151 (as described in [37, 46]), gene alignment with MACSE [47], and trimming with GBLOCKS [48]. The amino
152 acid supermatrix was reduced using a 70% occupancy threshold, producing a final dataset of 1,346 loci
153 (327,695 sites). As a final sanitation step, gene trees were obtained using ParGenes v. 1.0.1 [49], which
154 performed model selection (minimizing the Bayesian Information Criterion) and inference using 100
155 bootstrap (BS) replicates. Trees were then used to remove outlier sequences with TreeShrink v. 1.3.1
156 [50]. We specified a reduced tolerance for false positives and limited removal to at most three terminals
157 which had to increase tree diameter by at least 25% (-q 0.01 -k 3 -b 25).

158

159 **(b) Phylogenetic inference**

160 Inference was performed under multiple probabilistic and coalescent-aware methods, known to
161 differ in their susceptibility to model violations. Coalescent-based inference was performed using the
162 summary method ASTRAL-III [51], estimating support as local posterior probabilities [52]. Among
163 concatenation approaches, we used Bayesian inference under an unpartitioned GTR+G model in
164 ExaBayes 1.5 [53]. Two chains were run for 2.5 million generations, samples were drawn every one
165 hundred, and the initial 25% was discarded as burn-in. We also explored maximum likelihood inference
166 with partitioned and unpartitioned models. For the former, the fast-relaxed clustering algorithm was
167 used to find the best-fitting model among the top 10% using IQ-TREE 1.6.12 [54, 55] (-m MFP+MERGE -
168 rclusterf 10 -rcluster-max 3000), and support was evaluated with 1,000 ultrafast BS replicates [56]. For

169 the latter, we used the LG4X+R model in RAxML-NG v. 0.5.1 [57] and evaluated support with 200
170 replicates of BS. Finally, we also implemented the site-heterogenous LG+C60+F+G mixture model using
171 the posterior mean site frequency (PMSF) approach to provide a fast approximation of the full profile
172 mixture model [58], allowing the use of 100 BS replicates to estimate support. Given some degree of
173 topological conflict between the results of the other methods (see below), multiple guide trees were
174 used to estimate site frequency profiles, but the resulting phylogenies were identical.

175 Given conflicts between methods in the resolution of one particular node (involving the
176 relationships among Arbacioida, Salenioida, and the clade of Stomopneustoida + Camarodonta), all
177 methods were repeated after reducing the matrix to 500 and 100 loci selected for their phylogenetic
178 usefulness using the approach described in [3, 59] and implemented in the *genesortR* script
179 (https://github.com/mongiardino/genesortR). This approach relies on seven gene properties routinely
180 used for phylogenomic subsampling, including multiple proxies for phylogenetic signal—such as the
181 average BS and Robinson-Foulds (RF) similarity to a target topology—as well as several potential sources
182 of systematic bias (e.g., rate and compositional heterogeneities). Outgroups were removed before
183 calculating these metrics. RF similarity was measured to a species tree that had the conflicting
184 relationship collapsed so as not to bias gene selection in favor of any resolution. A principal component
185 analysis (PCA) of this dataset resulted in a dimension (PC 2, 17.6% of variance) along which phylogenetic
186 signal increased while sources of bias decreased (figure S1), and which was used for loci selection. For
187 the smallest dataset, we also performed inference under the site-heterogeneous CAT+GTR+G model
188 using PhyloBayes-MPI [60]. Three runs were continued for > 10,000 generations, sampling every two
189 generations and discarding the initial 25%. Convergence was confirmed given a maximum bipartition
190 discrepancy of 0.067 and effective sample sizes for all parameters > 150.

191 Two other approaches were implemented in order to assist in resolving the contentious node.
192 First, we implemented a likelihood-mapping analysis [61] in IQ-TREE to visualize the phylogenetic signal

193 for alternative resolutions of the quartet involving these three lineages (Arbacioida, Salenioida, and
194 Stomopneustoida + Camarodonta) and their sister clade (Irregularia; other taxa were excluded). Second,
195 we estimated the log-likelihood scores of each site in RAxML (using best-fitting models) for the two most
196 strongly supported resolutions found through likelihood mapping. These were used to calculate gene-
197 wise differences in scores, or δ values [62]. In order to search for discernable trends in the signal for
198 alternative topologies, genes were ordered based on their phylogenetic usefulness (see above) and the
199 mean per-locus δ values of datasets composed of multiples of 20 loci (i.e., the most useful 20, 40, etc.)
200 were calculated.

201

202 **(c) Time calibration**

203 Node dating was performed using relaxed molecular clocks in PhyloBayes v4.1 using a fixed
204 topology and a novel set of 22 fossil calibrations corresponding to nodes from our newly inferred
205 phylogeny (listed in the supplementary information). Depending on the node, we enforced both
206 minimum and maximum bounds, or either one of these. A birth-death prior was used for divergence
207 times, which allowed for the implementation of soft bounds [63], leaving 5% prior probability of
208 divergences falling outside of the specified interval. We explored the sensitivity of divergence time
209 estimates to gene selection, model of molecular evolution, and clock prior. One hundred loci were
210 sampled from the full supermatrix according to four different sampling schemes: usefulness (calculated
211 as explained above, except incorporating all echinoderm terminals), phylogenetic signal (i.e., smallest RF
212 distance to species tree), clock-likeness (i.e., smallest variance of root-to-tip distances), and level of
213 occupancy. For clock-likeness, we only considered loci that lay within one standard deviation of the
214 mean rate (estimated by dividing total tree length by the number of terminals [64]), as this method is
215 prone to select largely uninformative loci otherwise (figure S2; [59]). A fifth sample of randomly selected

loci was also evaluated. These five datasets were run under two models of molecular evolution, the site-homogenous GTR+G and the site-heterogenous CAT+GTR+G, and both uncorrelated gamma (UGAM) and log-normal autocorrelated (LN) clocks were implemented. The combination of these settings (loci sampled, model of evolution, and clock prior) resulted in 20 analyses. For each, two runs were continued for 20,000 generations, after which the initial 25% was discarded and the chains thinned to every two generations (see log-likelihood trace plots in figure S3). To explore the sensitivity of divergence times to these methodological decisions, 500 random chronograms were sampled from each analysis (250 from each run), and their node dates were subjected to multivariate analysis of variance (MANOVA) and between-group PCA (bgPCA) using package *Morpho* [65] in the R statistical environment [66]. bgPCA involves the use of PCA on the covariance matrix of group means (e.g., the mean ages of all nodes obtained under UGAM or LN clocks), followed by the projection of original samples onto the obtained bgPC axes. The result is a multidimensional representation of divergence times—a chronospace—rotated so as to capture the distinctiveness of observations obtained under different settings. Separate bgPCAs were performed for loci sampling strategy, clock prior and model of molecular evolution, and the proportion of total variance explained by bgPC axes was interpreted as an estimate of the relative impact of these choices on inferred times of divergence. Finally, lineage-through-time plots were generated using *ape* [67].

Results

Relationships inferred from the full dataset were remarkably stable, with all nodes but one being identically resolved and fully supported across inference methods (figure 2A). While recovering a topology similar to those of previous molecular studies [3, 4, 13, 14, 30, 68], this analysis is the first to sample and confidently place micropygoids and aspidodiadematoids within Aulodonta, as well as resolve

239 the relationships among all major clades of Neognathostomata (scutelloids, clypeateroids and the three
240 lineages of extant "cassiduloids"). Our results show that *Apatopygus recens* is not related to the
241 remaining "cassiduloids" but is instead the sister clade to all other sampled neognathostomates. The
242 strong support for this placement, as well as for a clade of cassidulids and echinolampadids
243 (Cassiduloida *sensu stricto*) as the sister group to sand dollars, provides a basis for an otherwise elusive
244 phylogenetic classification of neognathostomates. Our topology also confirms that *Sinaechinocyamus*
245 *mai*, a miniaturized species once considered a plesiomorphic member of Scutelloida based on the
246 reduction or loss of diagnostic features (figure 1), is in fact a derived paedomorphic lineage closely
247 related to *Scaphechinus mirabilis* [32].

248

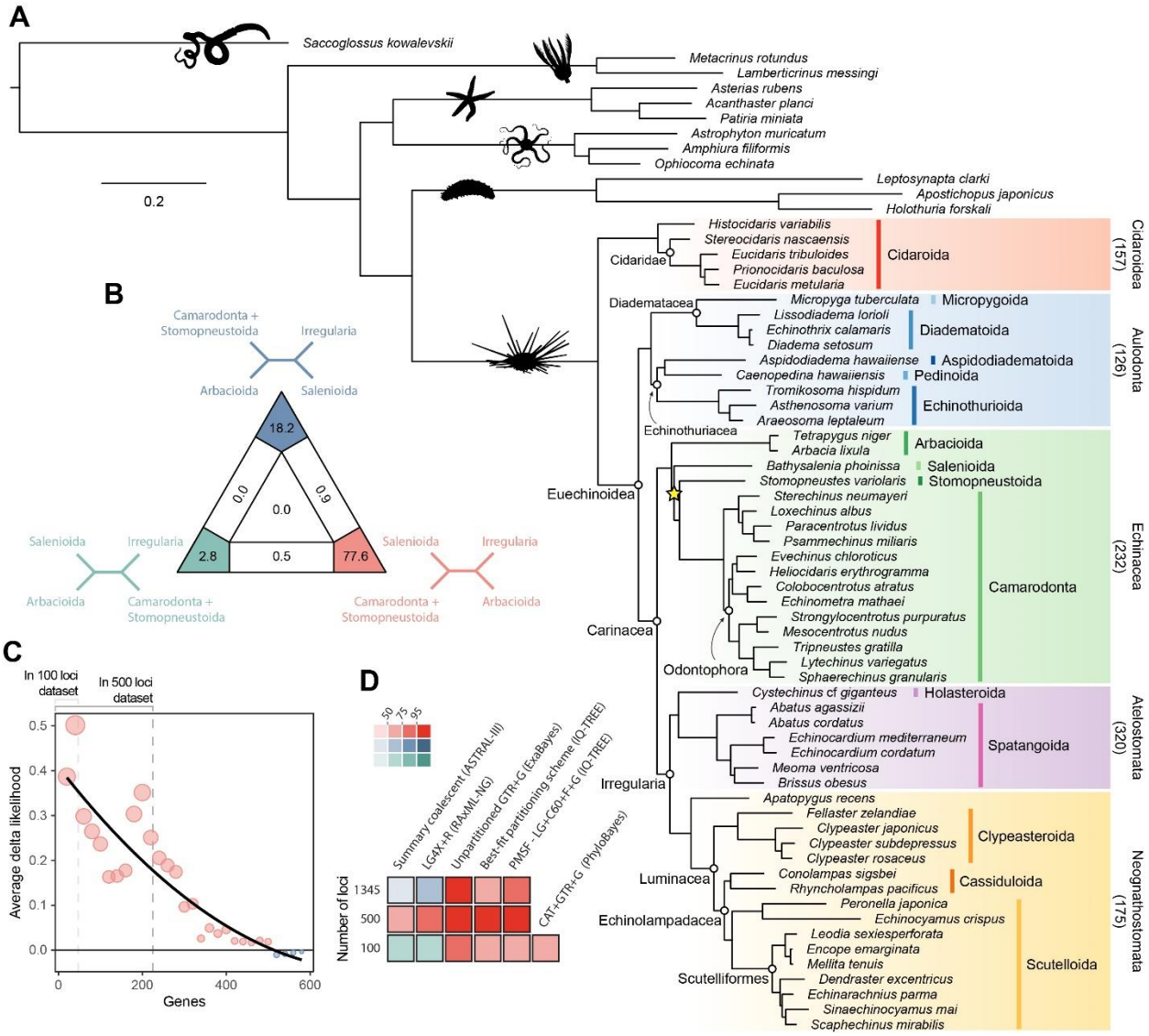


Figure 2: Phylogenetic relationships among major clades of Echinoidea. **A.** Favored topology. With the exception of a single contentious node within Echinacea (marked with a star; see Discussion), all methods supported the same pattern of relationships. Numbers below major clades correspond to the current numbers of described living species (obtained from [5]). **B.** Likelihood-mapping analysis showing the proportion of quartets supporting different resolutions within Echinacea. While the majority of quartets support the topology depicted in **A** (shown in red), a significant number support an alternative resolution that has been recovered in morphological analyses [19] (shown in blue). **C.** Support for a

257 clade of Salenioida + (Camarodonta + Stomopneustoida), as depicted in **A**, is predominantly
258 concentrated among the most phylogenetically useful loci. This signal is attenuated in larger datasets
259 that contain less reliable genes, eventually favoring an alternative resolution. Only the 584 loci
260 containing data for the three main lineages of Echinacea were considered. The line corresponds to a
261 second-degree polynomial regression. **D.** Resolution and bootstrap scores (color coded) of the topology
262 within Echinacea found using datasets of different sizes and alternative methods of inference.

263

264 Salenioida is another major lineage sampled here for the first time, and whose exact position
265 among regular echinoids proved difficult to resolve. While some methods supported salenioids as the
266 sister group to a clade of camarodonts, stomopneustoids, and arbacioids (a topology previously
267 supported by morphology [19]), others recovered a closer relationship of salenioids to Camarodonta +
268 Stomopneustoida, with arbacioids sister to them all. These results do not stem from a lack of
269 phylogenetic signal, but rather from the presence of strong and conflicting evidence in the dataset
270 regarding the position of salenioids (figure 2B). However, a careful dissection of these signals shows that
271 loci with high phylogenetic usefulness favor the topology shown in figure 2A, with the morphological
272 hypothesis becoming dominant only after incorporating less reliable loci (figure 2C). In line with these
273 results, moderate levels of gene subsampling (down to 500 loci) unambiguously support the placement
274 of arbacioids as sister to the remaining taxa regardless of the chosen method of inference (figure 2D).
275 More extreme subsampling (down to 100 loci) again results in disagreement among methods. This
276 possibly stems from the increasing effect of stochastic errors in smaller datasets, as less than half of the
277 sampled loci in these reduced datasets contain data for all branches of this quartet (see figure 2C). This
278 result shows the importance of ensuring that datasets (especially subsampled ones) retain appropriate
279 levels of occupancy for clades bracketing contentious nodes [69]. Despite these disagreements, several
280 lines of evidence favour the topology shown in figure 2A, including the results of likelihood mapping,

281 and the increased support for this resolution among the most phylogenetically useful loci and when
282 using more complex methods of reconstruction, such as partitioned and site-heterogenous models,
283 which favour this topology regardless of dataset size (figure 2D).

284 While alternative methods of inference had minor effects on phylogenetic relationships, they
285 did impact the reconstruction of branch lengths (figure S4). Site-heterogenous models (such as
286 CAT+GTR+G) returned longer branch lengths overall, but also uncovered a larger degree of molecular
287 change among echinoderm classes. Branches connecting these clades were stretched to a much larger
288 extent than those within the ingroup (figure S4), which might affect the inference of node dates. We
289 tested this hypothesis by exploring the sensitivity of divergence times to the use of alternative models of
290 molecular evolution, as well as different gene sampling strategies and clock priors, all of which
291 significantly affected inferred divergence times (MANOVA; all $P < 2.2\text{x}10^{-16}$). While nodes connecting
292 some outgroup taxa were among the nodes most sensitive to these decisions, large effects were also
293 seen among nodes relating to the origin and diversification of the echinoid clades Cidaroidea and
294 Aulodonta (figure 3A). All of these nodes varied in age by more than 60 Ma—and up to 115 Ma—among
295 the consensus topologies of different analyses (figure 3A).

296

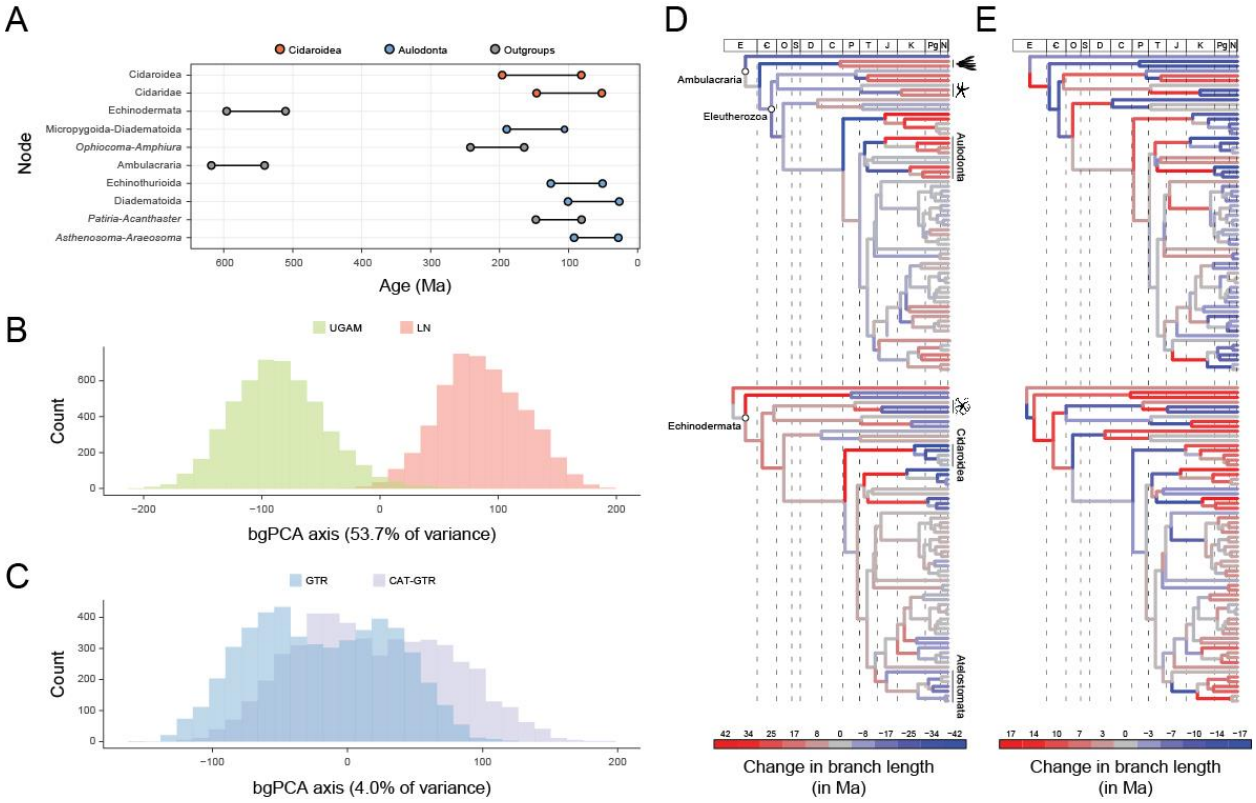


Figure 3: Sensitivity of divergence time estimation to methodological decisions. **A.** The ten most sensitive ages are found within Cidaroidea, Aulodonta, and among outgroup nodes. For each, the range shown corresponds to the ages found among the consensus topologies of the 40 time-calibrated runs. **B-C.** Scores for the between-group PCA (bgPCA) separating chronograms based on the clock model (**B**) or the model of molecular evolution (**C**) employed. While the first factor has a major impact on inferred ages, the second one has only minor effects. **D-E.** Change in branch lengths along the bgPCA axis defined by discriminating based on clock model (**D**) or model of molecular evolution (**E**). Branches are colored according to their contraction/expansion given a score of – 1 (top) or + 1 (bottom) standard deviations along the bgPCA axis relative to the mean. Clades/nodes showing age differences larger than 20 Ma are labelled (none in the case of **E**).

309 In order to isolate and visualize the impact of these factors on reconstructed ages, their major
310 effects were extracted with the use of bgPCAs (figures 3B-E, S5). This revealed that the single dimension
311 of chronospace maximizing the distinctiveness of trees obtained under different clock models explained
312 53.7% of the total variance in node ages across all analyses. In contrast, the choice of different loci or
313 models of molecular evolution showed a much lesser effect, explaining only 10.5% and 4.0% of the total
314 variance, respectively. Even though all of these decisions affected a similar set of sensitive nodes (those
315 mentioned above, as well as some relationships within Atelostomata; figures 3B-E, S5), the choice of
316 clock model modified the ages of 17 of these by more than 20 Ma (figure S6), a magnitude of change
317 that was not induced by either alternative loci or models of evolution on any node (figures S7-S8).

318 Even when the age of crown Echinodermata was constrained to postdate the appearance of
319 stereom (the characteristic skeletal microstructure of echinoderms) in the Early Cambrian [70, 71], only
320 analyses using the most clock-like loci recovered ages concordant with this (i.e., median ages younger
321 than the calibration enforced; figure S9). Instead, most consensus trees favored markedly older ages for
322 the clade, in some cases even predating the origin of the Ediacaran biota [72] (figures 3A, 4, S7). Despite
323 the relative sensitivity of many of the earliest nodes to methodological choices (figure 3), the split
324 between Crinoidea and all other echinoderms (Eleutherozoa) is always inferred to have predated the
325 end of the Cambrian (youngest median age = 492.2 Ma), and the divergence among the other major
326 lineages (classes) of extant echinoderms are constrained to have happened between the Late Cambrian
327 and Middle Ordovician (see figure S9). Our results also recover an early origin of crown group
328 Holothuroidea (sea cucumbers; range of median ages = 351.6–383.2 Ma), well before the crown groups
329 of other extant echinoderm classes. These dates markedly postdate the first records of holothuroid
330 calcareous rings in the fossil record [73, 74], and imply that this trait does not define the holothuroid
331 crown group but instead evolved from an echinoid-like jaw-apparatus along its stem [75]. The other
332 noteworthy disagreement between these results and those of previous studies [76] involves dating

crown group Crinoidea to times that precede the P-T mass extinction (range of median ages = 268.0–327.7 Ma, although highest posterior density intervals are always wide and include Triassic ages).

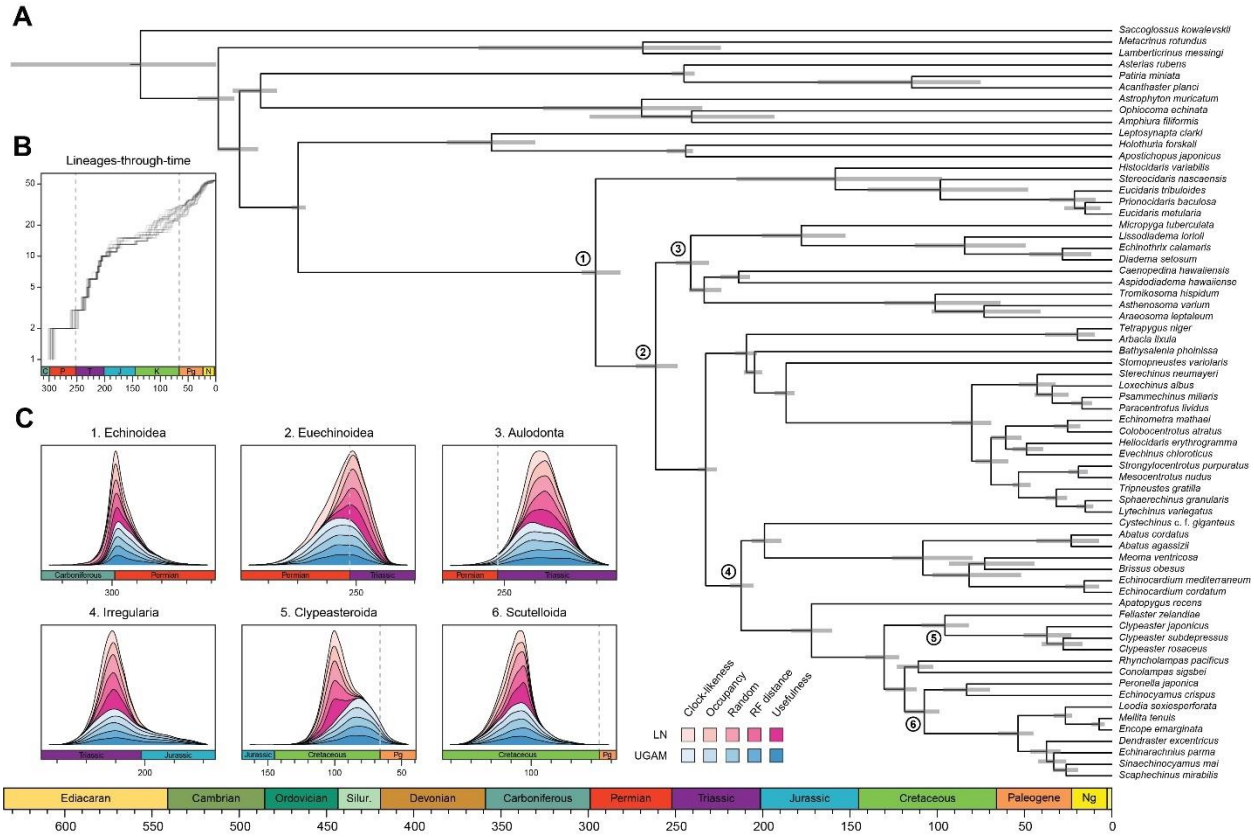


Figure 4: Divergence times among major clades of Echinoidea and other echinoderms. **A.** Consensus chronogram of the two runs using clock-like genes under a CAT+GTR+G model of molecular evolution and a LN clock prior. Node ages correspond to median values, and bars show the 95% highest posterior density intervals. **B.** Lineage-through-time plot, showing the rapid divergence of higher-level clades following the P-T mass extinction (shown with dashed lines, along with the K-Pg boundary). Each line corresponds to a consensus topology from among the 40 time-calibrated runs performed. **C.** Posterior distributions of the ages of selected nodes (marked in the chronogram). The effects of models of molecular evolution are not shown, as they represent the least important factor (see figure 3);

distributions using site-homogenous and heterogenous models are combined for every combination of targeted loci and clock prior.

Across all of the analyses performed, the echinoid crown group is found to have originated somewhere between the Pennsylvanian and Cisuralian, with 42% posterior probability falling within the late Carboniferous and 58% within the early Permian (figure 4C). An origin of the clade postdating the P-T mass extinction is never recovered, even when such ages were allowed under the joint prior (figure S10). While the posterior distribution of ages for Euechinoidea spans both sides of the P-T boundary, the remaining earliest splits within the echinoid tree are constrained to have occurred during the Triassic, including the origins of Aulodonta, Carinacea, Echinacea, and Irregularia (figures 4, S9). Many echinoid orders are also inferred to have diverged from their respective sister clades during this period, including aspidodiadematoids, pedinoids, echinothurioids, arbacioids, and salenioids. LTT plots confirm that lineage diversification proceeded rapidly throughout the Triassic (figure 4B). Despite the topological reorganization of Neognathostomata, the clade is dated to a relatively narrow time interval in the Late to Middle Jurassic (range of median ages = 169.48-179.96 Ma), in agreement with recent estimates [3]. Within this clade, the origins of both scutelloids and clypeasteroids confidently predate the K-Pg mass extinction (posterior probability of origination before the boundary = 0.9995 and 0.9621, respectively), despite younger ages being allowed by the joint prior (figure S10).

Discussion

(a) The echinoid Tree of Life

366 In agreement with previous phylogenomic studies [3, 4], echinoid diversity can be subdivided
367 into five major clades (figure 2A). Cidaroids form the sister group to all other crown group echinoids
368 (Euechinoidea). Some aspects of the relationships among sampled cidaroids are consistent with previous
369 molecular [77] and morphological studies [19], including an initial split between *Histocidaris* and the
370 remaining taxa, representing the two main branches of extant cidaroids [5, 35]. Others, such as the
371 nested position of *Prionocidaris baculosa* within the genus *Eucidaris,* not only implies paraphyly of this
372 genus but also suggests the need for a taxonomic reorganization of the family Cidaridae. Within
373 euechinoids, the monophyly of Aulodonta is supported for the first time with sampling of all of its major
374 groups. The subdivision of these into a clade that includes diadematoids plus micropygoids (which we
375 propose should retain the name Diadematacea), sister to a clade including echinothurioids and
376 pedinoids (Echinothuriacea *sensu* [4]) is strongly reminiscent of some early classifications (e.g., [78]).
377 Our expanded phylogenomic sampling also confirms an aulodont affinity for aspidodiadematoids [3, 35],
378 and places them within Echinothuriacea as the sister group to Pedinoida.

379 The remaining diversity of echinoids, which forms the clade Carinacea (figure 2), is subdivided
380 into Irregularia and their sister clade among regulars, for which we amend the name Echinacea to
381 include Salenioida. Given the striking morphological gap separating regular and irregular echinoids, the
382 origin of Irregularia has been shrouded in mystery [19, 23, 78]. Our complete sampling of major regular
383 lineages determines Echinacea *sensu stricto* to be the sister clade to irregular echinoids. A monophyletic
384 Echinacea was also supported in a recent total-evidence analysis [3], but the incomplete molecular
385 sampling of that study resulted in a slightly different topology that placed salenioids as the sister group
386 to the remaining lineages. However, an overall lack of morphological synapomorphies uniting these
387 clades had previously been acknowledged [19]. While the relationships within Echinacea proved to be
388 difficult to resolve even with thousands of loci, multiple lines of evidence lead us to prefer a topology in

389 which salenioids form a clade with camarodonts + stomopneustoids, with arbacioids sister to all of these
390 (figure 2).

391 As has been already established [3, 4, 14, 19, 30], the lineages of irregular echinoids here
392 sampled are subdivided into Atelostomata (heart urchins and allies) and Neognathostomata (sand
393 dollars, sea biscuits, and "cassiduloids"). Despite the former being the most diverse of the five main
394 clades of echinoids (figure 2), its representation in phylogenomic studies remains low, and its internal
395 phylogeny poorly constrained [35]. On the contrary, recent molecular studies have greatly improved our
396 understanding of the relationships among neognathosomates [3, 4, 68], revealing an evolutionary
397 history that dramatically departs from previous conceptions. Even when scutelloids and clypeasteroids
398 were never recovered as reciprocal sister lineages by molecular phylogenies (e.g., [13, 14, 30]), this
399 result was not fully accepted until phylogenomic data confidently placed echinolampadids as the sister
400 lineage to sand dollars [4]. At the same time, this result also rendered the position of the remaining
401 "cassiduloids", a taxonomic wastebasket with an already complicated history of classification [19, 27, 28,
402 79], entirely uncertain. An attempt to constrain the position of these using a total-evidence approach [3]
403 subdivided the "cassiduloids" into three unrelated clades: Nucleolitoida, composed of extinct lineages
404 and placed outside the node defined by Scutelloida + Clypeasteroida, and two other clades nested
405 within it (see figure 1). Extant "cassiduloids" were recovered as members of one of the latter clades,
406 representing the monophyletic sister group to sand dollars. Here, we show that *Apatopygus recens* does
407 not belong within this clade but is instead the sister group to all other extant neognathostomates. Given
408 this phylogenetic position, as well as the morphological similarities between *Apatopygus* and the
409 entirely extinct nucleolitids [19, 28, 79-81], it is likely that the three extant species of apatopygids
410 represent the last surviving remnants of Nucleolitoida, a clade of otherwise predominantly Mesozoic
411 neognathostomates [3]. Because of the renewed importance in recognizing this topology, we propose
412 the name Luminacea for the clade uniting all extant neognathostomates with the exclusion of

413 *Apatopygus* (figure 2). This nomenclature refers to the dynamic evolutionary history of the Aristotle's
414 lantern (i.e., the echinoid jaw-apparatus) within the clade (present in the adults of both clypeasteroids
415 and scutelloids, but found only in the juveniles of Cassiduloida *sensu stricto*), the inclusion of the so-
416 called lamp urchins (echinolampadids) within the clade, and the illumination provided by this hitherto
417 unexpected topology. The previous misplacement of *Apatopygus* within this clade ([3]; see figure 1G) is
418 likely a consequence of tip-dating preferring more stratigraphically congruent topologies [82], an effect
419 that can incorrectly resolve taxa on long terminal branches [83]. Given the generally useful phylogenetic
420 signal of stratigraphic information [84], this inaccuracy further highlights the unusual evolutionary
421 history of living apatopygids.

422

423 **(b) Echinoid diversification**

424 Calibrating phylogenies to absolute time is crucial to understanding evolutionary history, as the
425 resulting chronograms provide a major avenue for testing hypotheses of diversification. However, the
426 accuracy and precision of the inferred divergence times hinge upon many methodological choices that
427 are often difficult or time-consuming to justify (e.g., calibration strategies, prior distributions on node
428 ages, clock models, etc.) [85-90]. Understanding the impact of each of these sources of uncertainty can
429 be difficult. Here we explore the sensitivity of node ages to different ways of modelling heterogeneities
430 among lineages in the rates and patterns of molecular evolution, as well as alternative criteria to sample
431 molecular loci. To do so, we introduce an approach to visualize and measure the overall effect of these
432 decisions on inferred dates. Our results reveal that divergence times obtained under site-homogenous
433 and site-heterogenous models (such as CAT+GTR+G) are broadly comparable. This happens despite
434 these types of models estimating higher levels of sequence divergence and stretching branches in a non-
435 isometric manner (figure S3). A similar result was recently found when comparing site-homogenous

436 models with different numbers of parameters [91], suggesting that relaxed clocks adjust branch rates in
437 a manner that buffers the effects introduced by using more complex (and computationally-intensive)
438 models of evolution. Similarly, the choice between different loci also has a small effect on inferred ages,
439 with little evidence of a systematic difference between the divergence times supported by randomly-
440 chosen loci and those found using targeted sampling criteria, such as selecting genes for their
441 phylogenetic signal, usefulness, occupancy, or clock-likeness. A meaningful effect was restricted to a few
442 ancient nodes (e.g., Echinodermata), for which clock-like genes suggested younger ages that are more
443 consistent with fossil evidence. While this validates the use of clock-like genes for inferring deep
444 histories of diversification, the choice of loci had no meaningful effect on ingroup ages. Finally, the
445 choice between alternative clock models induced differences in ages that were five to ten times
446 stronger than those of other factors, emphasizing the importance of either validating their choice (e.g.,
447 [92]) or—as done here—focusing on results that are robust to them.

448 The origin and early diversification of crown group Echinoidea have always been considered to
449 have been determined (or at least strongly affected) by the P-T mass extinction [11, 12, 16, 20].
450 However, estimating the number of crown group members surviving the most severe biodiversity crisis
451 in the Phanerozoic [93] has been hampered by both paleontological and phylogenetic uncertainties [3,
452 10, 13-15, 18]. Our results establish with confidence that multiple crown group lineages survived and
453 crossed this boundary, finding for the first time a null posterior probability of the clade originating after
454 the extinction event. While the survival of three crown group lineages is slightly favored (figure S11),
455 discerning between alternative scenarios is still precluded by uncertainties in dating these early
456 divergences. Echinoid diversification during the Triassic was relatively fast (figure 4B) and involved rapid
457 divergence among its major clades. Even many lineages presently classified at the ordinal level trace
458 their origins to this initial pulse of diversification following the P-T mass extinction.

459 The late Palaeozic and Triassic origins inferred for the crown group and many euechinoid orders
460 prompts a re-evaluation of fossils from this interval of time. Incompletely known fossil taxa such as the
461 Pennsylvanian *?Eotiaris meurevillensis*, with an overall morphology akin to that of crown group
462 echinoids, has a stratigraphic range consistent with our inferred date for the origin of the echinoid
463 crown group [20]. Additionally, the Triassic fossil record of echinoids has been considered to be
464 dominated by stem group cidaroids, with the first euechinoids not known until the Late Triassic [17, 94].
465 However, the Triassic origins of many euechinoid lineages supported by our analyses necessitates that
466 potential euechinoid affinities should be re-considered for this diversity of Triassic fossils. This is
467 especially the case for the serpianotiarids and triadocidarids, abundant Triassic families variously
468 interpreted as cidaroids, euechinoids, or even stem echinoids [3, 17, 94, 95]. A reinterpretation of any of
469 these as stem euechinoids would suggest that the long-implied gap in the euechinoid record [16, 18] is
470 caused by our inability to correctly place these key fossils, as opposed to an incompleteness of the fossil
471 record itself.

472 While our phylogenomic approach is the first to resolve the position of all major cassiduloid
473 lineages, the inferred ages for many nodes within Neognathostomata remain in strong disagreement
474 with the fossil record. No Mesozoic fossil can be unambiguously assigned to either sand dollars or sea
475 biscuits, a surprising situation given the good fossilization potential and highly distinctive morphology of
476 these clades [19, 21, 25, 96]. While molecular support for a sister group relationship between scutelloids
477 and echinolampadids already implied this clade (Echinolampadacea) must have split from clypeasteroids
478 by the Late Cretaceous [3, 4, 14, 19], this still left open the possibility that the crown groups of sand
479 dollars and sea biscuits radiated in the Cenozoic. Under this scenario, the Mesozoic history of these
480 groups could have been comprised of forms lacking their distinctive morphological features,
481 complicating their correct identification. This hypothesis is here rejected, with the data unambiguously
482 supporting the radiation of both crown groups preceding the K-Pg mass extinction (figure 4C). While it

483 remains possible that these results are incorrect even after such a thorough exploration of the time-
484 calibration toolkit (see for example [90, 97]), these findings also call for a critical reassessment of the
485 Cretaceous fossil record, and a better understanding of the timing and pattern of morphological
486 evolution among fossil and extant neognathostomates. For example, isolated teeth with an overall
487 resemblance to those of modern sand dollars and sea biscuits have been found in Lower Cretaceous
488 deposits [98], raising the possibility that other overlooked and disarticulated remains might close the
489 gap between rocks and clocks.

490

491 **Conclusions**

492 Although echinoid phylogenetics has long been studied using morphological data, the position
493 of several major lineages (e.g., aspidodiadematoids, micropygoids, salenioids, apatopygids) remained to
494 be confirmed with the use of phylogenomic approaches. Our work not only greatly expands the
495 available genomic resources for the clade, but finds novel resolutions for some of these lineages,
496 improving our understanding of their evolutionary history. The most salient aspect of our topology is the
497 splitting of the extant "cassiduloids" into two distantly related clades, one of which is composed
498 exclusively of apatopygids. This result is crucial to constrain the ancestral traits shared by the main
499 lineages of neognathostomates, helping unravel the evolutionary processes that gave rise to the unique
500 morphology of the sand dollars and sea biscuits [3, 24, 25].

501 Although divergence time estimation is known to be sensitive to many methodological
502 decisions, systematically quantifying the relative impact of these on inferred ages has rarely been done.
503 Here we propose an approach based on chronospaces (implemented in R and available at
504 https://github.com/mongiardino/chronospace) that can help visualize key effects and determine the

505 sensitivity of node dates to different assumptions. Our results shed new light on the early evolutionary
506 history of crown group echinoids and its relationship with the P-T mass extinction event, a point in time
507 where the fossil record provides ambiguous answers. They also establish with confidence a Cretaceous
508 origin for the sand dollars and sea biscuits, preceding their first appearance in the fossil record by at
509 least 40 to 50 Ma, respectively (and potentially up to 65 Ma). These clades, therefore, join several well-
510 established cases of discrepancies between the fossil record and molecular clocks, such as those
511 underlying the origins of placental mammals [99] and flowering plants [100].

512

513 **Author contributions**

514 NMK, RM, and GWR conceived and designed the study. NMK, ASH, MFM, AFA, SEC, FA, OB, and
515 AK performed extractions, prepared libraries, and sequenced transcriptomic/genomic datasets. NMK
516 and AK generated *de novo* assemblies. JRT and RM designed fossil calibrations. NMK run all phylogenetic
517 analyses. NMK and JRT wrote the manuscript with input from all other authors. All authors gave final
518 approval for publication and agree to be held accountable for the work performed therein.

519

520 **Data, code and materials**

521 Raw transcriptomic reads have been deposited in NCBI under Bioproject accession number XXX.
522 Assemblies are available from the authors upon request. R code to plot chronospaces and explore the
523 effects of methodological decisions on divergence time estimation can be found at
524 https://github.com/mongiardino/chronospace.

525

Competing interests

We declare no competing interests.

Acknowledgments

We are grateful to the crew of R/Vs Antéa, Atlantis, Falkor and Mirai, HOV Alvin, and the crew and PIs of BIOMAGLO deep-sea cruises (Laure Corbari, Karine Olu-Le Roy, Sarah Samadi) for assistance in specimen collection. We would also like to thank Libby Liggins, Wilma Blom, Owen Anderson, Francisco Solís-Marín, Carlos A. Conejeros-Vargas, and Jih-Pai Lin for the collection and shipment of specimens. We gratefully acknowledge assistance from Thomas Saucède (Univ. Burgundy), Marc Eléaume (MNHN), and Charlotte Seid (Scripps) in accessing, cataloging and processing museum specimens. Tim Ravasi provided resources and collection facilities to GWR via King Abdullah University of Science and Technology (KAUST). Institutional support was provided by the Central Research Laboratories and the Department of Geology and Palaeontology at the Natural History Museum Vienna, Austria. This work was supported by a Yale Institute for Biospheric Studies Doctoral Dissertation Improvement Grant and a Society of Systematic Biologists Graduate Student Research Award to NMK. AK received funding from an Austrian Science Fund project (FWF, project number P 29508-B25), FA from Agencia Nacional de Investigación (ANID/PAI Inserción en la Academia, project number PAI79170033). JRT was supported by a Royal Society Newton International Fellowship and a Leverhulme Trust Early Career Fellowship. NMK was supported by a Yale University fellowship.

References

[1] Donoghue PC, Benton MJ. 2007 Rocks and clocks: calibrating the Tree of Life using fossils and molecules. *Trends Ecol. Evol.* **22**, 424–431.

[2] Smith AB, Peterson KJ. 2002 Dating the time of origin of major clades: molecular clocks and the fossil record. *Annu. Rev. Earth Planet. Sci.* **30**, 65–88.

[3] Mongiardino Koch N, Thompson JR. 2021 A Total-Evidence Dated Phylogeny of Echinoidea Combining Phylogenomic and Paleontological Data. *Syst. Biol.* **70**, 421–439.

[4] Mongiardino Koch N, Coppard SE, Lessios HA, Briggs DEG, Mooi R, Rouse GW. 2018 A phylogenomic resolution of the sea urchin tree of life. *BMC Evol. Biol.* **18**, 189.

[5] Kroh A, Mooi R. 2020 World Echinoidea Database. Accessed at http://www.marinespecies.org/echinoidea on 2021-07-13. (doi:10.14284/355)

[6] Schultz HA. 2015 *Echinoidea: with pentameral symmetry*. Walter de Gruyter GmbH & Co KG.

[7] Edmunds PJ, Carpenter RC. 2001 Recovery of *Diadema antillarum* reduces macroalgal cover and increases abundance of juvenile corals on a Caribbean reef. *Proc. Natl. Acad. Sci.* **98**, 5067–5071.

[8] Harrold C, Pearse JS. 1987 The ecological role of echinoderms in kelp forests. *Echinoderm Studies* **2**, 137–233.

[9] Zonneveld JP, Furlong CM, Sanders SC. 2016 Triassic echinoids (Echinodermata) from the Aksala Formation, north Lake Laberge, Yukon Territory, Canada. *Pap. Palaeontol.* **2**, 87–100.

[10] Thompson JR, Petsios E, Bottjer DJ. 2017 A diverse assemblage of Permian echinoids (Echinodermata, Echinoidea) and implications for character evolution in early crown group echinoids. *J. Paleontol.* **91**, 767–780.

[11] Kier PM. 1977 Triassic echinoids. *Smithson. Contrib. Paleobiol.* **30**, 1–88.

[12] Twitchett RJ, Oji T. 2005 Early Triassic recovery of echinoderms. *C. R. Palevol* **4**, 531–542.

[13] Thompson JR, Erkenbrack EM, Hinman VF, McCauley BS, Petsios E, Bottjer DJ. 2017 Paleogenomics of echinoids reveals an ancient origin for the double-negative specification of micromeres in sea urchins. *Proc. Natl. Acad. Sci.* **114**, 5870–5877.

[14] Smith AB, Pisani D, Mackenzie-Dodds JA, Stockley B, Webster BL, Littlewood DTJ. 2006 Testing the molecular clock: molecular and paleontological estimates of divergence times in the Echinoidea (Echinodermata). *Mol. Biol. Evol.* **23**, 1832–1851.

[15] Smith AB, Hollingworth N. 1990 Tooth structure and phylogeny of the Upper Permian echinoid *Miocidaris keyserlingi*. *Proc. Yorkshire Geol. Soc.* **48**, 47–60.

[16] Thompson JR, Petsios E, Davidson EH, Erkenbrack EM, Gao F, Bottjer DJ. 2015 Reorganization of sea urchin gene regulatory networks at least 268 million years ago as revealed by oldest fossil cidaroid echinoid. *Sci. Rep.* **5**, 15541.

[17] Smith AB. 2007 Intrinsic versus extrinsic biases in the fossil record: contrasting the fossil record of echinoids in the Triassic and early Jurassic using sampling data, phylogenetic analysis, and molecular clocks. *Paleobiology* **33**, 310–323.

[18] Thompson JR, Hu S-X, Zhang Q-Y, Petsios E, Cotton LJ, Huang J-Y, Zhou C-Y, Wen W, Bottjer DJ. 2018 A new stem group echinoid from the Triassic of China leads to a revised macroevolutionary history of echinoids during the end-Permian mass extinction. *R. Soc. Open Sci.* **5**, 171548.

[19] Kroh A, Smith AB. 2010 The phylogeny and classification of post-Palaeozoic echinoids. *J. Syst. Palaeontol.* **8**, 147–212.

588 [20] Thompson JR, Mirantsev GV, Petsios E, Bottjer DJ. 2020 Phylogenetic analysis of the
589 Archaeocidaridae and Palaeozoic Miocidaridae (Echinodermata, Echinoidea) and the origin of crown
590 group echinoids. *Pap. Palaeontol.* **6**, 217–249.

591 [21] Kier PM. 1982 Rapid evolution in echinoids. *Palaeontology* **25**, 1–9.

592 [22] Barras CG. 2008 Morphological innovation associated with the expansion of atelostomate irregular
593 echinoids into fine-grained sediments during the Jurassic. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **263**,
594 44–57.

595 [23] Saucède T, Mooi R, David B. 2007 Phylogeny and origin of Jurassic irregular echinoids
596 (Echinodermata: Echinoidea). *Geol. Mag.* **144**, 333–359.

597 [24] Hopkins MJ, Smith AB. 2015 Dynamic evolutionary change in post-Paleozoic echinoids and the
598 importance of scale when interpreting changes in rates of evolution. *Proc. Natl. Acad. Sci.* **112**, 3758–
599 3763.

600 [25] Mooi R. 1990 Paedomorphosis, Aristotle's lantern, and the origin of the sand dollars
601 (Echinodermata: Clypeasteroida). *Paleobiology* **16**, 25–48.

602 [26] Nebelsick JH. 2020 Ecology of clypeasteroids. In: *Sea Urchins: Biology and Ecology, 4th edition* (ed.
603 JM Lawrence), pp. 315–331. Cambridge, Academic Press.

604 [27] Smith AB. 2001 Probing the cassiduloid origins of clypeasteroid echinoids using stratigraphically
605 restricted parsimony analysis. *Paleobiology* **27**, 392–404.

606 [28] Souto C, Mooi R, Martins L, Menegola C, Marshall CR. 2019 Homoplasy and extinction: the
607 phylogeny of cassidulid echinoids (Echinodermata). *Zool. J. Linn. Soc.* **187**, 622–660.

608 [29] Kier PM. 1962 Revision of the cassiduloid echinoids. *Smithson. Misc. Collect.* **144**, 1–262.

609 [30] Littlewood DTJ, Smith AB. 1995 A combined morphological and molecular phylogeny for sea urchins
610 (Echinoidea: Echinodermata). *Philos. Trans. R. Soc. B* **347**, 213–234.

611 [31] Mongiardino Koch N. 2021 Exploring adaptive landscapes across deep time: A case study using
612 echinoid body size. *Evolution* **75**, 1567–1581.

613 [32] Mooi R. 1990 Progenetic miniaturization in the sand dollar Sinaechinocyamus: implications for
614 clypeasteroid phylogeny. In: *Echinoderm Research* (eds. C De Ridder, P Dubois, M Lahaye, M Jangoux),
615 pp. 137–143. Rotterdam, AA Balkema.

616 [33] Dunn CW, Zapata F, Munro C, Siebert S, Hejnol A. 2018 Pairwise comparisons across species are
617 problematic when analyzing functional genomic data. *Proc. Natl. Acad. Sci.* **115**, 409–417.

618 [34] Smith SD, Pennell MW, Dunn CW, Edwards SV. 2020 Phylogenetics is the new genetics (for most of
619 biodiversity). *Trends Ecol. Evol.* **35**, 415–425.

620 [35] Kroh A. 2020 Phylogeny and classification of echinoids. In: *Sea Urchins: Biology and Ecology, 4th*
621 *edition* (ed. JM Lawrence), pp. 1–17. Cambridge, Academic Press.

622 [36] Bolger AM, Lohse M, Usadel B. 2014 Trimmomatic: a flexible trimmer for Illumina sequence data.
623 *Bioinformatics* **30**, 2114–2120.

624 [37] Dunn CW, Howison M, Zapata F. 2013 Agalma: an automated phylogenomics workflow. *BMC*
625 *Bioinformatics* **14**, 330.

626 [38] Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury
627 R, Zeng Q, et al. 2011 Full-length transcriptome assembly from RNA-Seq data without a reference
628 genome. *Nature Biotechnology* **29**, 644–652.

629 [39] Modolo L, Lerat E. 2015 UrQt: an efficient software for the Unsupervised Quality trimming of NGS
630 data. *BMC Bioinformatics* **16**, 137.

631 [40] Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. 2015 MEGAHIT: an ultra-fast single-node solution for
632 large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676.

633 [41] Smit A, Hubley R, Green P. 2015 RepeatMasker Open-4.0. 2013–2015.

634 [42] Hoff KJ, Stanke M. 2019 Predicting genes in single genomes with AUGUSTUS. *Curr. Protoc.*
635 *Bioinformatics* **65**, e57.

636 [43] Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. 2006 AUGUSTUS: ab initio
637 prediction of alternative transcripts. *Nucleic Acids Res.* **34**, 435–439.

638 [44] Simion P, Belkhir K, François C, Veyssier J, Rink JC, Manuel M, Philippe H, Telford MJ. 2018 A
639 software tool 'CroCo' detects pervasive cross-species contamination in next generation sequencing data.
640 *BMC Biol.* **16**, 28.

641 [45] Ryan JF. 2014 Alien Index: identify potential non-animal transcripts or horizontally transferred
642 genes in animal transcriptomes. (doi: 10.5281/zenodo.21029)

643 [46] Guang A, Howison M, Zapata F, Lawrence C, Dunn CW. 2021 Revising transcriptome assemblies with
644 phylogenetic information. *PLoS One* **16**, e0244202.

645 [47] Ranwez V, Harispe S, Delsuc F, Douzery EJ. 2011 MACSE: Multiple Alignment of Coding SEquences
646 accounting for frameshifts and stop codons. *PLoS One* **6**, e22594.

647 [48] Talavera G, Castresana J. 2007 Improvement of phylogenies after removing divergent and
648 ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**, 564–577.

649 [49] Morel B, Kozlov AM, Stamatakis A. 2018 ParGenes: a tool for massively parallel model selection and
650 phylogenetic tree inference on thousands of genes. *Bioinformatics* **35**, 1771–1773.

651 [50] Mai U, Mirarab S. 2018 TreeShrink: fast and accurate detection of outlier long branches in
652 collections of phylogenetic trees. *BMC Genomics* **19**, 272.

653 [51] Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018 ASTRAL-III: polynomial time species tree
654 reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**, 153.

655 [52] Sayyari E, Mirarab S. 2016 Fast coalescent-based computation of local branch support from quartet
656 frequencies. *Mol. Biol. Evol.* **33**, 1654–1668.

657 [53] Aberer AJ, Kobert K, Stamatakis A. 2014 ExaBayes: massively parallel Bayesian tree inference for the
658 whole-genome era. *Mol Biol Evol* **31**, 2553–2556.

659 [54] Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2014 IQ-TREE: a fast and effective stochastic
660 algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274.

661 [55] Kalyaanamoorthy S, Minh BQ, Wong TK, von Haeseler A, Jermiin LS. 2017 ModelFinder: fast model
662 selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589.

663 [56] Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. 2017 UFBoot2: improving the ultrafast
664 bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522.

665 [57] Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019 RAxML-NG: a fast, scalable and user-
666 friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**, 4453-4455.

667 [58] Wang H-C, Minh BQ, Susko E, Roger AJ. 2018 Modeling site heterogeneity with posterior mean site
668 frequency profiles accelerates accurate phylogenomic estimation. *Syst. Biol.* **67**, 216–235.

669 [59] Mongiardino Koch N. 2021 Phylogenomic subsampling and the search for phylogenetically reliable
670 loci. *Mol. Biol. Evol.* **msab151**. (doi:10.1093/molbev/msab151).

671 [60] Lartillot N, Rodrigue N, Stubbs D, Richer J. 2013 PhyloBayes MPI. Phylogenetic reconstruction with
672 infinite mixtures of profiles in a parallel environment. *Syst. Biol.* **62**, 611–615.

673 [61] Strimmer K, Von Haeseler A. 1997 Likelihood-mapping: a simple method to visualize phylogenetic
674 content of a sequence alignment. *Proc. Natl. Acad. Sci.* **94**, 6815–6819.

675 [62] Shen X-X, Hittinger CT, Rokas A. 2017 Contentious relationships in phylogenomic studies can be
676 driven by a handful of genes. *Nat. Ecol. Evol.* **1**, 0126.

677 [63] Yang Z, Rannala B. 2006 Bayesian estimation of species divergence times under a molecular clock
678 using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* **23**, 212–226.

679 [64] Telford MJ, Lowe CJ, Cameron CB, Ortega-Martinez O, Aronowicz J, Oliveri P, Copley RR. 2014
680 Phylogenomic analysis of echinoderm class relationships supports Asterozoa. *Proc. R. Soc. B* **281**,
681 20140479.

682 [65] Schlager S. 2017 Morpho and Rvcg–Shape Analysis in R. In: *Statistical shape and deformation*
683 *analysis* (eds. G Zheng, S Li, GJ Székely), pp. 217–256. Academic, London.

684 [66] R Core Team. 2019 *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R
685 Foundation for Statistical Computing. See https://www.R-project.org.

686 [67] Paradis E, Schliep K. 2018 ape 5.0: an environment for modern phylogenetics and evolutionary
687 analyses in R. *Bioinformatics* **35**, 526–528.

688 [68] Lin J-P, Tsai M-H, Kroh A, Trautman A, Machado DJ, Chang L-Y, Reid R, Lin K-T, Bronstein O, Lee S-J,
689 et al. 2020 The first complete mitochondrial genome of the sand dollar *Sinaechinocyamus mai*
690 (Echinoidea: Clypeasteroida). *Genomics* **112**, 1686–1693.

691 [69] Dell'Ampio E, Meusemann K, Szucsich NU, Peters RS, Meyer B, Borner J, Petersen M, Aberer AJ,
692 Stamatakis A, Walzl MG, et al. 2014 Decisive data sets in phylogenomics: lessons from studies on the
693 phylogenetic relationships of primarily wingless insects. *Mol. Biol. Evol.* **31**, 239–249.

694 [70] Zamora S, Lefebvre B, Álvaro JJ, Clausen S, Elicki O, Fatka O, Jell P, Kouchinsky A, Lin J-P, Nardin E.
695 2013 Cambrian echinoderm diversity and palaeobiogeography. *Geol. Soc. Lond. Mem.* **38**, 157–171.

696 [71] Bottjer DJ, Davidson EH, Peterson KJ, Cameron RA. 2006 Paleogenomics of echinoderms. *Science*
697 **314**, 956–960.

698 [72] Pu JP, Bowring SA, Ramezani J, Myrow P, Raub TD, Landing E, Mills A, Hodgin E, Macdonald FA. 2016
699 Dodging snowballs: Geochronology of the Gaskiers glaciation and the first appearance of the Ediacaran
700 biota. *Geology* **44**, 955–958.

701 [73] Reich M. 2015 Different pathways in early evolution of the holothurian calcareous ring. In: *Progress*
702 *in Echinoderm Palaeobiology* (eds. S Zamora, I Rábano), pp. 137–145. Madrid: Instituto Geológico y
703 Minero de España.

704 [74] Miller AK, Kerr AM, Paulay G, Reich M, Wilson NG, Carvajal JI, Rouse GW. 2017 Molecular phylogeny
705 of extant Holothuroidea (Echinodermata). *Mol. Phylogenet. Evol.* **111**, 110–131.

706 [75] Rahman IA, Thompson JR, Briggs DEG, Siveter DJ, Siveter DJ, Sutton MD. 2019 A new ophiocistioid
707 with soft-tissue preservation from the Silurian Herefordshire Lagerstätte, and the evolution of the
708 holothurian body plan. *Proc. R. Soc. B* **286**, 20182792.

[76] Rouse GW, Jermiin LS, Wilson NG, Eeckhaut I, Lanterbecq D, Oji T, Young CM, Browning T, Cisternas P, Helgen LE, et al. 2013 Fixed, free, and fixed: the fickle phylogeny of extant Crinoidea (Echinodermata) and their Permian–Triassic origin. *Mol. Phylogenet. Evol.* **66**, 161–181.

[77] Brosseau O, Murienne J, Pichon D, Vidal N, Eléaume M, Ameziane N. 2012 Phylogeny of Cidaroida (Echinodermata: Echinoidea) based on mitochondrial and nuclear markers. *Org. Divers. Evol.* **12**, 155–165.

[78] Durham JW, Melville R. 1957 A classification of echinoids. *J. Paleontol.* **31**, 242–272.

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

[80] Kier PM. 1966 Cassiduloids. In: *Treatise on Invertebrate Paleontology, Part U* (ed. R Moore), pp. 492–523. Lawrence, University of Kansas Press.

[81] Mortensen T. 1948 *A Monograph of the Echinoidea. IV, 1 Holectypoida, Cassiduloida*. Copenhagen, CA Reitzel.

[82] King B. 2020 Bayesian Tip-dated Phylogenetics in Paleontology: Topological Effects and Stratigraphic Fit. *Syst. Biol.* **70**, 283–294.

[83] Turner AH, Pritchard AC, Matzke NJ. 2017 Empirical and Bayesian approaches to fossil-only divergence times: a study across three reptile clades. *PLoS One* **12**, e0169885.

[84] Mongiardino Koch N, Garwood RJ, Parry LA. 2021 Fossils improve phylogenetic analyses of morphological characters. *Proc. R. Soc. B* **288**, 2021004.

[85] Dos Reis M, Gunnell GF, Barba-Montoya J, Wilkins A, Yang Z, Yoder AD. 2018 Using phylogenomic data to explore the effects of relaxed clocks and calibration strategies on divergence time estimation: primates as a test case. *Syst. Biol.* **67**, 594–615.

[86] Warnock RC, Yang Z, Donoghue PC. 2011 Exploring uncertainty in the calibration of the molecular clock. *Biol. Lett.* **8**, 156–159.

[87] Dos Reis M, Thawornwattana Y, Angelis K, Telford MJ, Donoghue PC, Yang Z. 2015 Uncertainty in the timing of origin of animals and the limits of precision in molecular timescales. *Curr. Biol.* **25**, 2939–2950.

[88] Sauquet H. 2013 A practical guide to molecular dating. *C. R. Palevol* **12**, 355–367.

[89] Carruthers T, Scotland RW. 2021 Uncertainty in divergence time estimation. *Syst. Biol.* **70**, 855–861.

[90] Carruthers T, Sanderson MJ, Scotland RW. 2020 The implications of lineage-specific rates for divergence time estimation. *Syst. Biol.* **69**, 660–670.

[91] Tao Q, Barba-Montoya J, Huuki LA, Durnan MK, Kumar S. 2020 Relative efficiencies of simple and complex substitution models in estimating divergence times in phylogenomics. *Mol. Biol. Evol.* **37**, 1819–1831.

[92] Tao Q, Tamura K, Battistuzzi FU, Kumar S. 2019 A machine learning method for detecting autocorrelation of evolutionary rates in large phylogenies. *Mol. Biol. Evol.* **36**, 811–824.

[93] Raup DM, Sepkoski JJ. 1982 Mass extinctions in the marine fossil record. *Science* **215**, 1501–1503.

[94] Kier PM. 1984 Echinoids from the Triassic (St. Cassian) of Italy, their lantern supports, and a revised phylogeny of Triassic echinoids. *Smithson. Contrib. Paleobiol*. **56**, 1–41.

[95] Smith AB. 1994 Triassic echinoids from Peru. *Palaeontogr. Abt. A* **233**, 177–202.

[96] Kier PM. 1977 The poor fossil record of the regular echinoid. *Paleobiology* **3**, 168–174.

[97] Field D, Berv J, Hsiang A, Lanfear R, Landis M, Dornburg A. 2020 Timing the Extant Avian Radiation: The Rise of Modern Birds, and the Importance of Modeling Molecular Rate Variation. *Bull. Am. Mus. Nat. Hist.* **440**, 150–181.

[98] Caramés A, Adamonis S, Concheyro A, Remírez, M. 2019 New Finding of Regular Echinoid Elements and Microfossils from the Pilmatué Member, Agrio Formation (Early Cretaceous), Neuquén Basin, Argentina. In: *Advances in South American Micropaleontology* (eds. G Cusminsky, E Bernasconi, G Concheyro), pp. 1–20. Springer Earth System Sciences. Springer, Cham.

[99] Ronquist F, Lartillot N, Phillips MJ. 2016 Closing the gap between rocks and clocks using total-evidence dating. *Phil. Trans. R. Soc. B* **371**, 20150136.

[100] Coiro M, Doyle JA, Hilton J. 2019 How deep is the conflict between molecular and fossil evidence on the age of angiosperms? *New Phytol.* **223**, 83–99.

761 # Supplementary Information for:

762 N Mongiardino Koch, JR Thompson, AS Hatch, MF McCowin, AF Armstrong, SE Coppard, F Aguilera, O
763 Bronstein, A Kroh, R Mooi, GW Rouse – Phylogenomic analyses of echinoid diversification prompt a re-
764 evaluation of their fossil record

765

766 **Details on tissue sampling, extraction, library preparation, sequencing and bioinformatic**
767 **approaches**

768 *Apatopygus recens*, *Aspidodiadema hawaiiense*, *Fellaster zelandiae*, *Histocidaris variabilis*,
769 *Rhyncholampas pacificus*, *Sinaechinocyamus mai*, *Stereocidaris neumayeri*, *Tromikosoma hispidum*.
770 Tissue subsamples were finely chopped with a scalpel and preserved in RNA*later* (Ambion) buffer
771 solution for 1 day at 4°C to allow the RNAl*ater* to effectively penetrate the tissues, followed by long-
772 term storage at -80°C until RNA extraction. Total RNA was extracted from 1.5 mm Triple-Pure High
773 Impact Zirconium beads (Benchmark Scientific) in Trizol (Ambion), using Direct-zol RNA Miniprep Kit
774 (Zymo Research) with in-column DNase I incubation to remove genomic DNA. Prior to mRNA capture,
775 total RNA concentration was estimated using Qubit RNA HS Assay Kit (Invitrogen; range = 8.33-96
776 ng/µL), and quality was assessed using either High Sensitivity RNA ScreenTape or RNA ScreenTape with
777 an Agilent 4200 TapeStation. Mature mRNA was isolated from total RNA and libraries were prepared
778 using KAPA mRNA HyperPrep Kit (KAPA Biosystems) following the manufacturer's instructions (including
779 sample customization based on total RNA quantity and quality values), targeting an insert size circa 500
780 base pairs (bp), and using custom 10-nucleotide Illumina TruSeq style adapters [101]. Post-amplification,
781 DNA concentration was estimated using Qubit dsDNA BR Assay Kit (Invitrogen; range = 6.06-13.4 ng/µL).
782 Concentration, quality, and molecular weight distribution of libraries (range = 547-978 bp) was also

783 assessed using Genomic DNA ScreenTape with an Agilent 4200 TapeStation. Ten libraries (including two
784 annelids not employed here) were sequenced on one lane of a multiplexed run using NovaSeq S4
785 platform with 100 bp paired-end reads at the IGM Genomics Center (University of California San Diego).

786 The assembled transcriptomes were sanitized from cross-contaminant reads product of
787 multiplexed sequencing using CroCo v. 1.1 [44]. These eight datasets, as well as two annelid
788 transcriptomes not employed in this study but sequenced in the same lane, were employed. Transcripts
789 considered over or under-expressed across samples (as defined by default parameters), were kept.
790 This resulted in an average removal of 1.26% of assembled transcripts (range: 0.4% for
791 *Aspidodiadema* to 3.06% for *Histocidaris*).

792

793 *Clypeaster japonicus*, *Encope emarginata*, *Leodia sexiesperforata*, *Peronella japonica*. Eggs were
794 collected from a single female specimen of each species and immediately placed into RNA later for
795 preservation. The samples were left in RNAlater at 4°C for at least 1 day and then transferred to a -80°C
796 freezer until RNA extraction. RNA extraction was performed using Trizol and was then treated with
797 Ambion's Turbo DNA-free kit. RNA was quantified using both Nanodrop and an Agilent 2100
798 Bioanalyzer. Only RNA samples with an RNA integrity number (RIN) of > 7 were used. RNA-Seq libraries
799 were prepared using the NEB Ultra Directional kit using the standard protocol and multiplexed with 6bp
800 NEB multiplex primers. Paired-end 100 libraries were generated from the resultant libraries at the UC
801 Berkeley Genome Center using an Illumina HighSeq 2500.

802

803 *Bathysalenia phoinissa*, *Micropyga tuberculata*: The material was collected by the deep-sea
804 cruise BIOMAGLO [102] conducted in 2017 jointly by the French National Museum of Natural History

805 (MNHN) as part of the Tropical Deep-Sea Benthos program, the French Research Institute for
806 Exploitation of the Sea (IFREMER), the "Terres Australes et Antarctiques Françaises" (TAAF), the
807 Departmental Council of Mayotte and the French Development Agency (AFD), with the financial support
808 of the European Union (Xe FED). Specimens were sorted into taxonomic bins at class level and
809 collectively fixed in 70% ethanol at room temperature. Following taxonomic identification, tissue
810 samples were obtained using sterilized forceps and disposable scalpel blades. Tissue subsamples were
811 collected in high-purity 96% ethanol and stored at -40°C until DNA extraction. Total genomic DNA was
812 extracted using the DNeasy Blood and Tissue Kit (QIAGEN) following the manufacturer's instructions,
813 complemented by the addition of 4 µl RNAse A (QIAGEN, 100 mg/ml) for RNA digestion. Elution buffer
814 volume was reduced to 60 µl and the elution step repeated twice per sample in order to maximize DNA
815 concentration and yield. After collection of the first elution, spin-column membranes were rinsed again
816 twice with fresh elution buffer (40 µl) to further increase yield. Elutions were collected separately, with
817 the first one used for sequencing and the second for quality control. For the latter, a portion of the
818 mitochondrial cytochrome c oxidase subunit I (COI) was amplified using PCR. Amplifications were
819 conducted with TopTaq DNA Polymerase (QIAGEN) using 1 µl of extracted genomic DNA (approx. 10–15
820 ng). Details on primers and protocols can be found in [103]. PCR products were visualized on a 1%
821 agarose gel, purified using ExoSAP-IT (Affymetrix), and sequenced at Microsynth GmbH (Vienna, Austria)
822 with the same primers. Sequences are deposited in GenBank under the accession numbers XXX and XXX.

823 Prior to library preparation, total genomic DNA concentration was estimated using Qubit DNA
824 BR Assay Kit (Invitrogen) in order to determine library preparation strategy. This step was carried out by
825 Macrogen (Korea), using an Illumina TruSeq DNA PCR-free kit (for *Micropyga*) and an Illumina TruSeq
826 Nano DNA kit (for *Bathysalenia*). Library preparation followed the manufacturer's instructions, targeting
827 an insert size of 350 base pairs (bp). The libraries were sequenced by Macrogen (Korea) on a shared lane
828 of a multiplexed run using an Illumina HiSeq X instrument with 150 bp paired-end reads. Sequencing

829 depth was 134.1 and 161.3 million reads for *Bathysalenia* and *Micropyga*, respectively. Pre-processing of
830 Illumina shotgun data was carried out by removal of adapters using BBDuk
831 (https://sourceforge.net/projects/bbmap/) with settings: minlen=50 ktrim=r k=23 mink=11 tpe tbo);
832 followed by quality trimming of reads using UrQt v. 1.0.18 [39], discarding regions with phred quality
833 score < 28 and sequences of size < 50. Assembly of the data was carried out with MEGAHIT v. 1.1.2 [40],
834 in an iterative approach followed by assessment of assembly quality using BUSCO v. 3.0.2. scores [104]
835 using the metazoan dataset ODB v. 9 [105]. Final BUSCO scores of the draft assemblies were 28.4%
836 (S:27.8%, D:0.6%, F:48.2%, M:23.4%, n:978) for *Bathysalenia* and 17.2% (S:16.8%, D:0.4%, F:59.0%,
837 M:23.8%, n:978) for *Micropyga*. As recommended by Hoff and Stanke [42], draft genomes were masked
838 with RepeatMasker v. 4.1.0 [41] prior to gene prediction, using settings: -norna -xsmall; resulting in 3.82
839 % and 2.9% masked bases for *Bathysalenia* and *Micropyga*, respectively. Gene prediction was carried
840 out using AUGUSTUS v. 3.2.3 [43] using settings: --strand=both --singlestrand=false --genemodel=partial
841 --codingseq=on --sample=0 --alternatives-from-sampling=false --exonnames=on --softmasking=on. This
842 step relied on a training dataset composed of universal single-copy orthologs (USCOs) derived from the
843 most recent *Strongylocentrotus purpuratus* genome (Spur v.5).

844

845 *Echinothrix calamaris*, *Tripneustes gratilla*. Tissue samples were taken from live sea urchins
846 housed in laboratory aquaria and total RNA was immediately extracted from 150 mg of tissue using a
847 PureLink™ RNA extraction kit (Invitrogen™). The quality of total RNA was assessed on a BioAnalyzer
848 2100 (Agilent Technologies, Santa Clara, CA, USA) to ensure a RIN > 9 for all samples. Mature mRNA was
849 extracted from 1ug of total RNA and cDNA libraries were constructed using a TruSeq kit (Illumina).
850 Quality control of libraries was assessed on a BioAnalyzer and quantification measured using qPCR.
851 *NEBNext* Multiplex adaptors were added via ligation, and the cDNA libraries were sequenced at Genome
852 Quebec, McGill University. Three libraries were multiplex sequenced on one lane of Illumina at a

853 concentration of 8 pM per cDNA library using HiSeq2500 transcriptome sequencing to generate 125bp
854 paired-end reads. This resulted in approximately 80 million reads for each transcriptome.

855

856 *Tetrapygus niger*. Adult specimens were collected and transported alive to the University of
857 Concepcion. Tissue samples (female gonad and tube feet) from one individual were finely chopped with
858 a scalpel, and total RNA was extracted immediately using TriReagent (Sigma), following the
859 manufacturer's instructions. Total RNA concentration was estimated using QuantiFluor® RNA System
860 (111.84-339.04 ng/µL) and quality was assessed using Agilent RNA 6000 Nano kit with an Agilent 2100
861 TapeStation (RIN: 5.4-8.4). cDNA libraries were prepared at Genoma Mayor SpA (Chile) using TruSeq
862 Stranded mRNA Kit (Illumina) and sequenced on a multiplexed run using Illumina HiSeq 4000 platform
863 with 150 paired-end reads, resulting in 54.1 M reads for gonad female and 52.7 M reads for tube feet,
864 respectively. These two transcriptomic datasets were combined before performing all steps of
865 bioinformatic processing.

866

867 *Eucidaris metularia*. A single specimen was preserved in RNA*later* after being starved overnight
868 and kept for 1 day at 4°C before long-term storage at -80°C. Total RNA was extracted using Direct-zol
869 RNA Miniprep Kit (with in-column DNase treatment; Zymo Research) from Trizol. mRNA was isolated
870 with Dynabeads mRNA Direct Micro Kit (Invitrogen). RNA concentration was estimated using Qubit RNA
871 broad range assay kit, and quality was assessed using RNA ScreenTape with an Agilent 4200 TapeStation.
872 Values were used to customize downstream protocols following manufacturers' instructions. Library
873 preparation was performed with KAPA-Stranded RNA-Seq kits, targeting an insert size in the range of
874 200–300 base pairs. The quality and concentration of libraries were assessed using DNA ScreenTape.
875 The library was sequenced in a multiplexed pair-end runs using Illumina HiSeq 4000 with 7 other

876 libraries in the same lane (all previously published in [4]). In order to minimize read crossover, 10 bp
877 sequence tags designed to be robust to indel and substitution errors were employed [101].

878 The assembled transcriptome was filtered from cross-contaminant reads product of multiplexed
879 sequencing using CroCo v. 1.1 [44]. Seven other transcriptomes sequenced together were also
880 employed, and results of this step are reported in [3] (figure S1).

881

882 **Fossil constraints on node ages**

883 All clade names used below refer to the corresponding crown groups.

884 Ambulacraria (Echinodermata-Hemichordata divergence) – **Younger bound**: 518 Ma, Middle Atdabanian
885 (age 3, Series 3 of Cambrian). **Older bound**: 636.1 Ma, Lantian Biota, Ediacaran. **Reference:** [106]. **Notes:**
886 The root of our tree is constrained with a younger bound concurrent with the earliest occurrence of
887 stereom in the fossil record (see below). The older bound is based on the maximum age of the Lantian
888 Biota, a Lagerstätte lacking any trace of eumetazoan fossils.

889 Echinodermata (Pelmatozoan-Eleutherozoan divergence) – **Younger bound**: Unconstrained. **Older**
890 **bound:** 515.5 Ma, Middle Atdabanian (age 3, Series 3 of Cambrian). **Reference**: [106]. **Notes:** This
891 divergence (which represents the divergence of crown group echinoderms), has an older bound based
892 upon the earliest occurrence of stereom in the fossil record. Stereom constitutes an echinoderm
893 synapomorphy readily recognizable in the fossil record due to its unique mesh-like structure [71]. The
894 first records of stereom point to a sudden and global appearance within the middle Atdabanian [70].
895 Note that this is also used as younger bound for the divergence between echinoderms and
896 hemichordates.

Echinozoa (Echinoidea-Holothuroidea divergence) – **Younger bound:** 469.96 Ma, Top of *Pseudoclamacograptus decorates* graptolite zone. **Older bound:** 461.95 Ma, Top Floian. **Reference:** [107, 108]. **Notes:** The oldest crown eleutherozoan fossils are disarticulated holothurian calcareous ring elements which are known from the Red *Orthoceras* limestone of Sweden. This falls within the *P. decorates* graptolite zone.

Asterozoa (Ophiuroidea-Asteroidea divergence) – **Younger bound**: 521 Ma, Top of *Psigraptus jacksoni* Zone. **Older bound**: 480.55 Ma, Base of Series 2 Cambrian Period. **Reference**: [107, 108]. **Taxon**: *Maydenia roadsidensis.* **Notes**: The divergence of asterozoan classes is calibrated based upon the stratigraphic occurrence of *Maydenia roadsidensis*, the oldest asterozoan, which is a stem group ophiuroid [109] and occurs in the *Psigraptus jacksoni* Zone of the Floian.

Asteroidea – **Younger bound:** 239.10 Ma, Top Fassanian, Ladinian, Mo3, *nodosus* biozone, Triassic. **Older bound**: 252.16 Ma, P-T Boundary. **Reference**: [110]. **Taxon**: *Trichasteropsis weissmanni*. **Notes**: Fossil evidence suggests the divergence of the asterozoan crown group took place sometime in the Early or Middle Triassic [110]. Based upon its phylogenetic position and stratigraphic occurrence as the oldest crown group asteroid, we use the forcipulatacean *T. weissmanni* from the Middle Triassic Muschelkalk as the soft bound on this divergence.

Crinoidea – **Younger bound**: 247.06 Ma, Top Spathian (Paris Biota, Lower Shale unit, Thaynes Group, Spathian, Triassic). **Older bound**: Unconstrained. **Reference**: [111]. **Taxon**: *Holocrinus*. **Notes**: Crown group Crinoidea is difficult to define, though a rapid post-Palaeozoic morphological diversification is supported by fossil evidence. We use *Holocrinus* from the Early Triassic of the Thaynes group to calibrate the younger bound on the divergence of the crown group (i.e., the split between Isocrinida and all other extant crinoids).

Ophiuroidea (Euryophiurida-Ophintegrida divergence) – **Younger bound**: 247.06 Ma, Top Spathian (Paris Biota, Lower Shale unit, Thaynes Group, Spathian, Triassic). **Older bound**: Unconstrained. **Reference**: [112]. **Taxon**: *Shoshonura brayardi*. **Notes**: *S. brayardi* from the Spathian Thaynes Group of the Western USA is a member of the crown group ophiuroid clade Ophiodermatina. It is thus the currently oldest described member of the ophiuroid crown group, setting its minimum age of divergence.

Pneumonophora (Holothuriida-Neoholothuriida divergence) – **Younger bound**: 259.8 Ma, S*pinosus* zone of Early Ladinian (used base Ladinian). **Older bound**: 241.5 Ma, Base Wuchiapingian. **Reference**: [74, 107]. **Notes**: The oldest (undescribed) holothuriid calcareous ring ossicles are from the *Spinosus* zone of the Ladinian of Germany and calibrate its divergence from neoholothuriids.

Echinoidea (Cidaroidea-Euechinoidea divergence) – **Younger bound**: 298.9 Ma, Base Carnian. **Older bound**: 237 Ma, Base Permian. **Reference**: [113]. **Taxon**: *Triassicidaris? ampezzana*. **Notes**: The recent phylogenetic analysis of Mongiardino Koch and Thompson [3] found many traditional early crown group echinoids as members of the stem group. Given this result, *Triassicidaris? ampezzana*, a cidaroid which is known from the St. Cassian beds of Italy, becomes the current oldest crown group echinoid.

Pedinoida-Aspidodiadematoida – **Younger bound**: 209.46 Ma, Top of Norian. **Older bound**: 252.16 Ma, P-T boundary. **Reference**: See discussion in supplement of Thompson et al. [13]. **Taxon**: *Diademopsis* ex. gr. *heberti*. **Notes**: The oldest euechinoid, *Diademopsis* ex. gr. *heberti*, which is a pedinoid, is known from the Norian of Peru. This fossil calibrates the younger bound on the pedinoid-aspidodiadematoid divergence.

Carinacea (Echinacea-Irregularia divergence) – **Younger bound**: 234.5 Ma, Top Sinemurian. **Older bound**: 190.8 Ma, Top Julian 1 ammonoid zone of Norian. **Reference**: Li et al. [114]. **Taxon**: *Jesionekechinus*. **Notes**: The divergence of Carinacea is calibrated based upon the oldest irregular

echinoid, *Jesionekechinus*, which occurs above the Sinemurian of New York Canyon, Nevada. For a more detailed discussion of this divergence see the supplement of Thompson et al. [13].

Salenioida-(Camarodonta+Stomopneustoida) divergence– **Younger bound**: 228.35 Ma, Base Hettangian. **Older bound**: 201.3 Ma, Top Carnian. **Reference**: [14]. **Taxon**: *Acrosalenia chartroni*. **Notes**: The stem group salenioid *Acrosalenia chartroni* from the Hettangian of France is the earliest representative of Echinacea. Given the topology recovered by our analyses, this fossil was used to calibrate the divergence between salenioids and the clade composed of stomopneustoids and camarodonts.

Stomopneustoida-Camarodonta divergence – **Younger bound**: 201.3 Ma, Top Pliensbachian. **Older bound**: 182.7 Ma, Base Jurassic. **Reference**: See discussion in supplement of Thompson et al. [13]. **Taxon**: Stomechinids from Morocco like *Diplechinus hebbriensis* [115]. **Notes**: The oldest stomechinids, which are stem group stomopneustoids, are from the Early Jurassic of Morocco.

Neognathostomata-Atelostomata divergence – **Younger bound**: 234.5 Ma, Base of Toarcian. **Older bound**: 182.7 Ma, St. Cassian Beds, bottom of Julian 2 ammonoid Zone. **Reference**: [14, 114]. **Taxon**: Younger bound is set by *Galeropygus sublaevis* (older bound is relatively uninformative). **Notes**: The base of the Toarcian (which contains the oldest neognathostomate *G. sublaevis*) is used as the lower bound on the divergence between the neognathostomata and all other crown irregular echinoids. Note that depending on the position of the unsampled echinoneoids, this split might also correspond to the origin of crown group Irregularia (see recent topologies recovered by Lin et al. [68] and Mongiardino Koch & Thompson [3]).

Atelostomata (Holasteroida-Spatangoida divergence) – **Younger bound**: 137.68 Ma, Bottom of *Campylotoxus* zone in Valanginian. **Older bound**: 201.3 Ma, Base of Jurassic. **Reference**: [116]. **Taxon**: Younger bound is based on *Toxaster* (older bound is loosely uninformative). **Notes**: The toxasterids are stem group spatangoids that calibrate the divergence between spatangoids and holasteroids. The oldest

964 *Toxaster* are from the *Campylotoxus* zone in the Valanginian, and set the younger bound on the
965 divergence.

966 Echinolampadacea (Cassiduloida-Scutelloida divergence) – **Younger bound**: 113 Ma, Base Albian. **Older**
967 **bound**: 145 Ma, Base Cretaceous. **Reference**: [117]. **Taxon**: Younger bound is based on *Eurypetalum*
968 *rancheriana* (older bound is loosely uninformative). **Notes**: The oldest member of Echinolampadacea is
969 the faujasiid *Eurypetalum rancheriana* (originally placed in the genus *Faujasia* and later transferred, see
970 [28, 29]) from the Late Albian of Colombia.

971 Scutelloida (Scutelliformes-Laganiformes divergence) – **Younger bound**: 56.0 Ma, Bottom of Eocene.
972 **Older bound**: Unconstrained. **Reference**: [14, 118]. **Taxon**: Younger bound is based on *Eoscutum*
973 *doncieuxi*. **Notes**: We use the base of the Eocene as the younger bound on the divergence between
974 scutelliforms and laganiforms. This is based upon the occurrence of *E. doncieuxi*, the oldest scutelliform,
975 which is known from the Lower Eocene. The older bound is left unconstrained in order to account for
976 the current uncertainty in the origin of the clade (i.e., allow for Mesozoic ages).

977 *Leodia*-(*Encope*+*Mellita*) divergence – **Younger bound**: 23.03 Ma, Base Miocene. **Older bound**: 33.9 Ma,
978 Base of Oligocene. **Reference**: [119, 120]. **Taxon**: Younger bound is based on *Encope michoacanensis*
979 (older bound is loosely uninformative). **Notes**: This node is constrained using the oldest representative
980 of these three genera of mellitids.

981 Clypeasteroida – **Younger bound**: 47.8 Ma, Base Lutetian. **Older bound**: Unconstrained. **Reference**:
982 [121, 122]. **Taxon**: Younger bound is based on *Clypeaster calzadai* and *Clypeaster moianensis*. **Notes**:
983 The oldest known fossil species of Clypeasteroida (the clypeasterids *C. calzadai* and *C. moianensis*) are
984 from the middle Eocene, upper Lutetian (Biarritzian) of Cataluña, Spain. The older bound is left
985 unconstrained in order to account for the current uncertainty in the origin of the clade (i.e., allow for
986 Mesozoic ages).

Strongylocentrotidae-Toxopneustidae divergence – **Younger bound**: 44.4 Ma, Eocene Planktonic Zones E9-E12. **Older bound**: 56.0 Ma, Base Eocene. **Reference**: [123, 124]. **Taxon**: Younger bound is based on *Lytechinus axiologus* (older bound is loosely uninformative). **Notes**: The oldest member of the clade comprising toxopneustids and strongylocentrotids is the toxopneustid *Lytechinus axiologus.* This taxon is known from the Eocene Yellow Limestone Group of Jamaica. The Yellow Limestone group represents the Eocene Planktonic Zones E9-E12.

Supplementary Figures and Tables

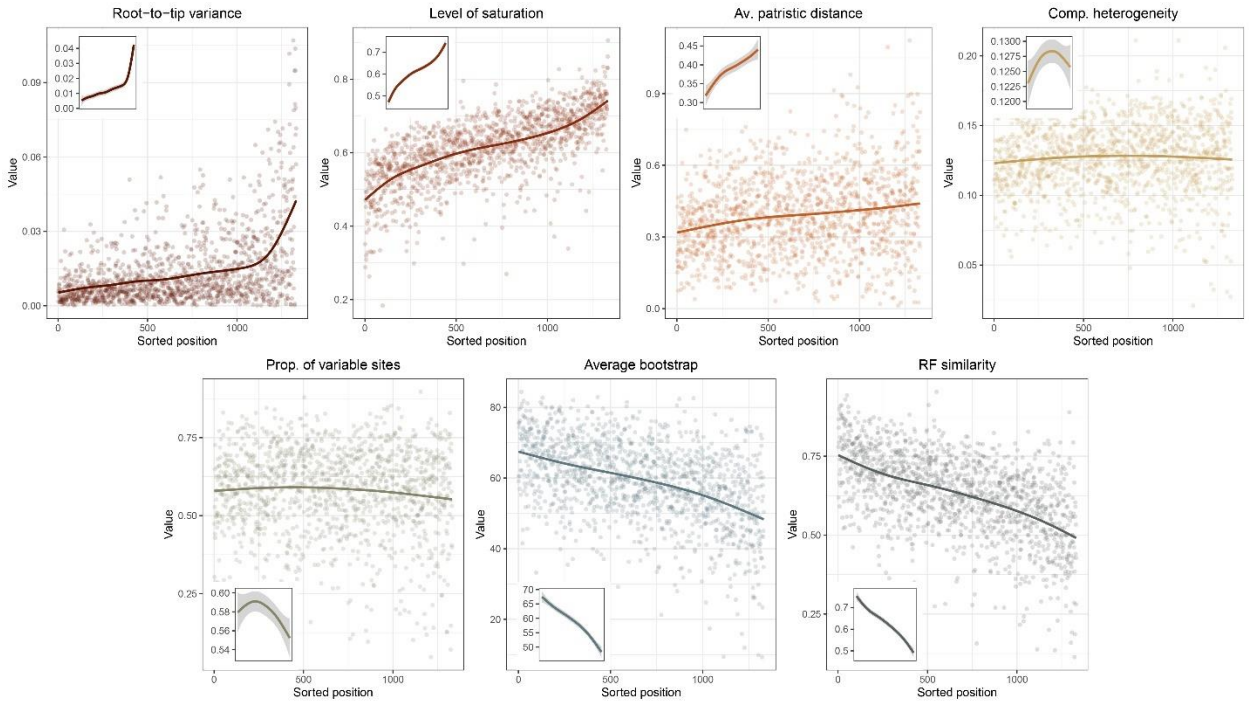
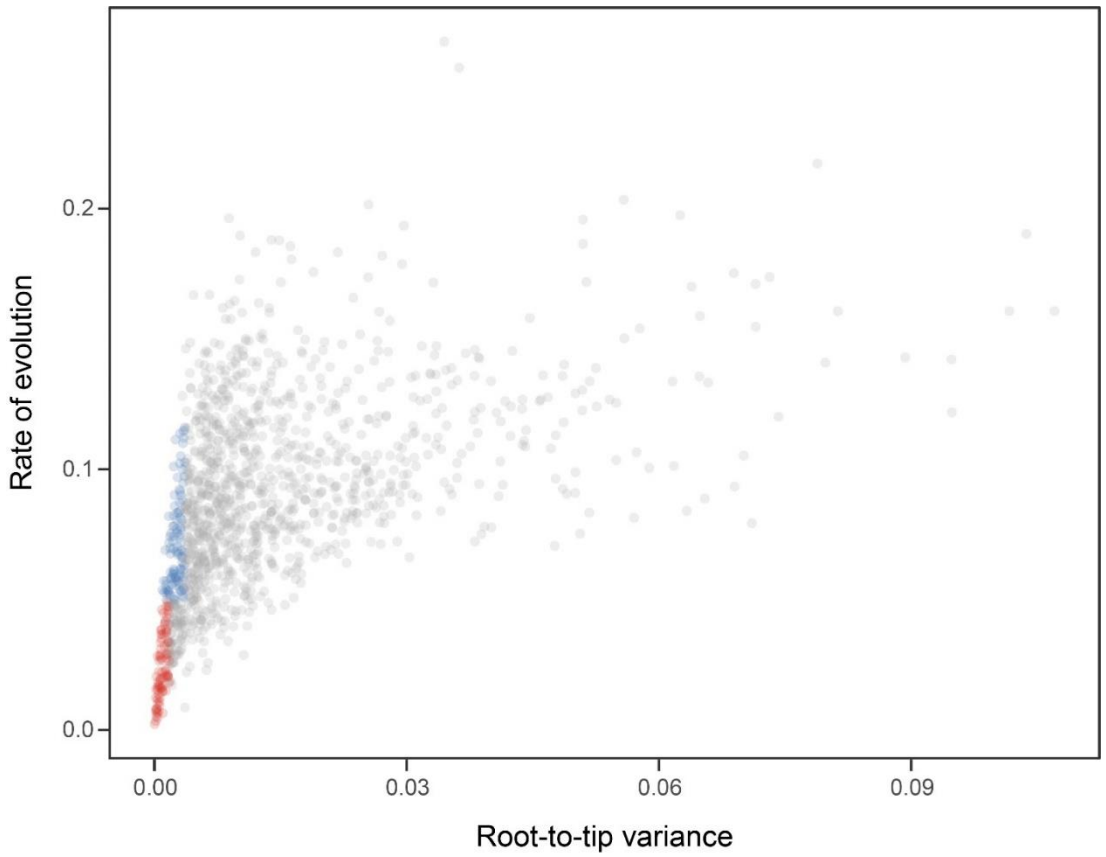


Figure S1: Ordering of loci enforced using *genesortR* [59] and its relationship to the seven gene properties employed. High ranking loci (i.e., the most phylogenetically useful) show low root-to-tip variances (or high clock-likeness), low saturation and compositional heterogeneity, as well as high average bootstrap and Robinson-Foulds similarity to a target topology (in this case, with the contentious relationship among major lineages of Echinacea collapsed).



1000

1001 **Figure S2:** Relationship between the root-to-tip variance (a proxy for the clock-likeness of loci) and the
1002 rate of evolution (estimated as the total tree length divided by the number of terminals). The most
1003 clock-like loci (shown in red), which are often favored for the inference of divergence times (e.g., [90,
1004 125]), are among the most highly conserved and can provide little information for constraining node
1005 ages [59]. Highly clock-like genes with a higher information content were used instead by choosing the
1006 loci with the lowest root-to-tip variance from among those that were within 1 standard deviation from
1007 the mean evolutionary rate (shown in blue).

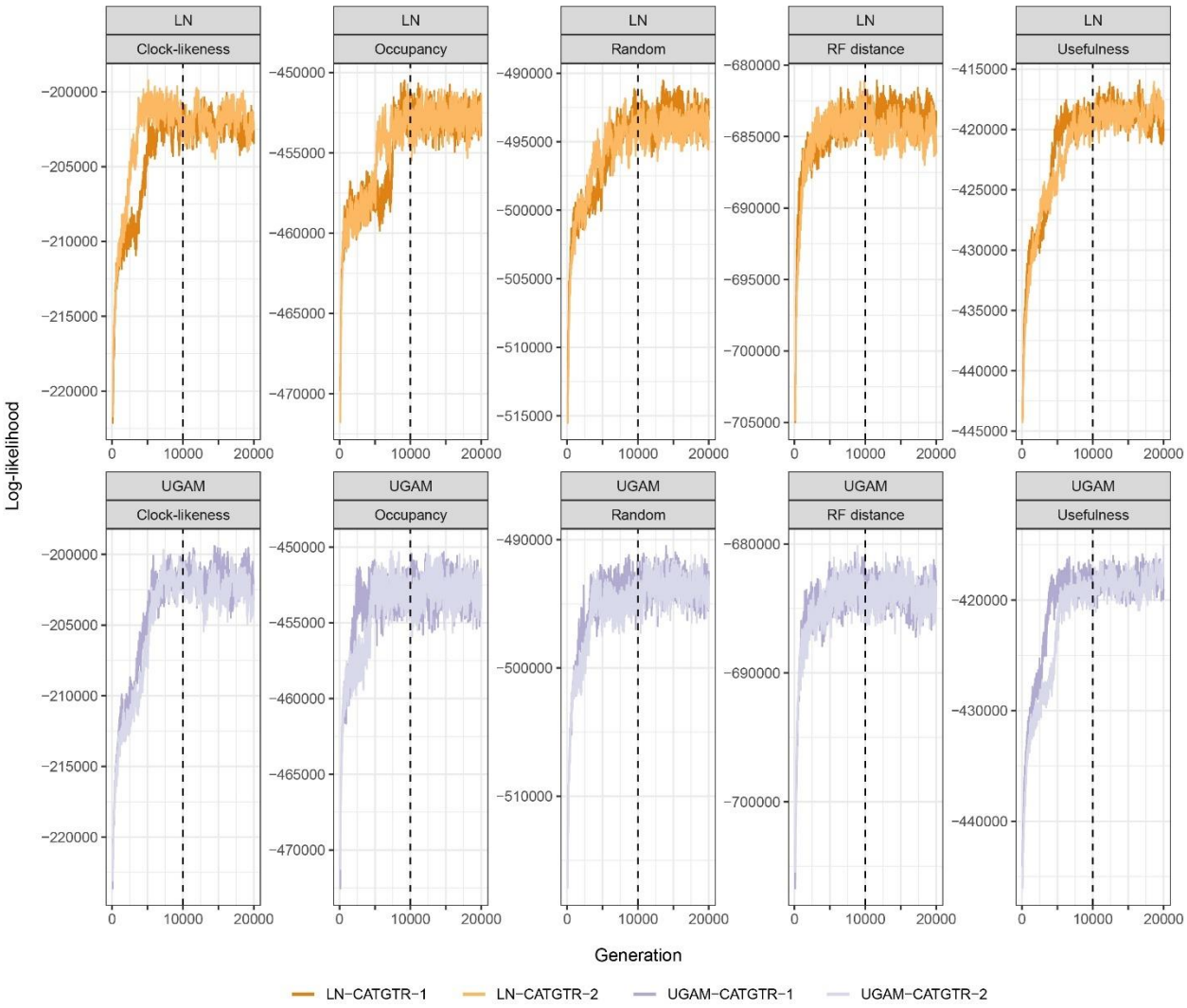


Figure S3: Trace plots of the log-likelihood of different time-calibration runs. All runs show evidence of reaching convergence and stationarity before our imposed burn-in fraction of 10,000 generations. For simplicity, only runs under the CAT+GTR+G model are plotted. Those run under GTR+G converged much faster.

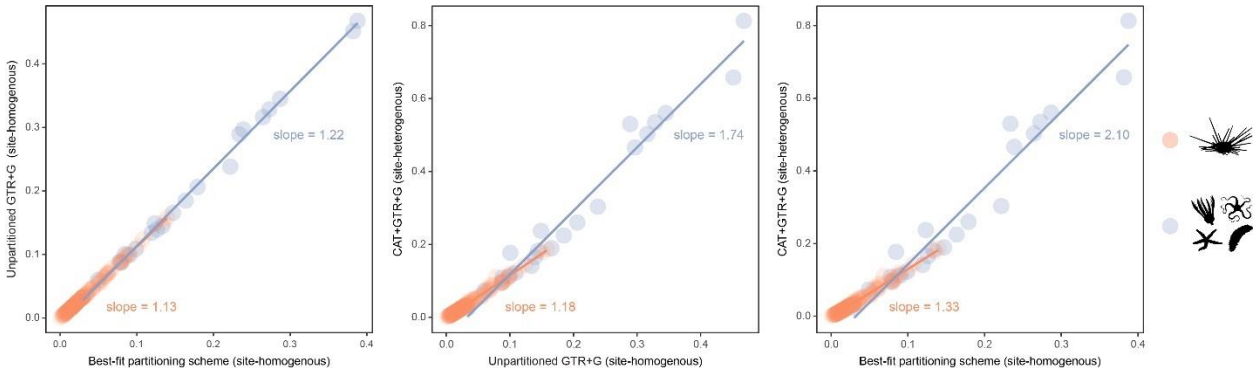


Figure S4: Estimated branch lengths across different models of molecular evolution. Different site-homogenous models (left) infer similar levels of divergence, and the choice between them induces little distortion in the general tree structure. Site-heterogenous models on the other hand, not only infer a larger degree of divergence between terminals relative to site-homogenous ones (center and left), but they also distort the tree (i.e., impose a non-isometric stretching), with branch lengths connecting outgroup taxa expanding much more than those within the ingroup clade.

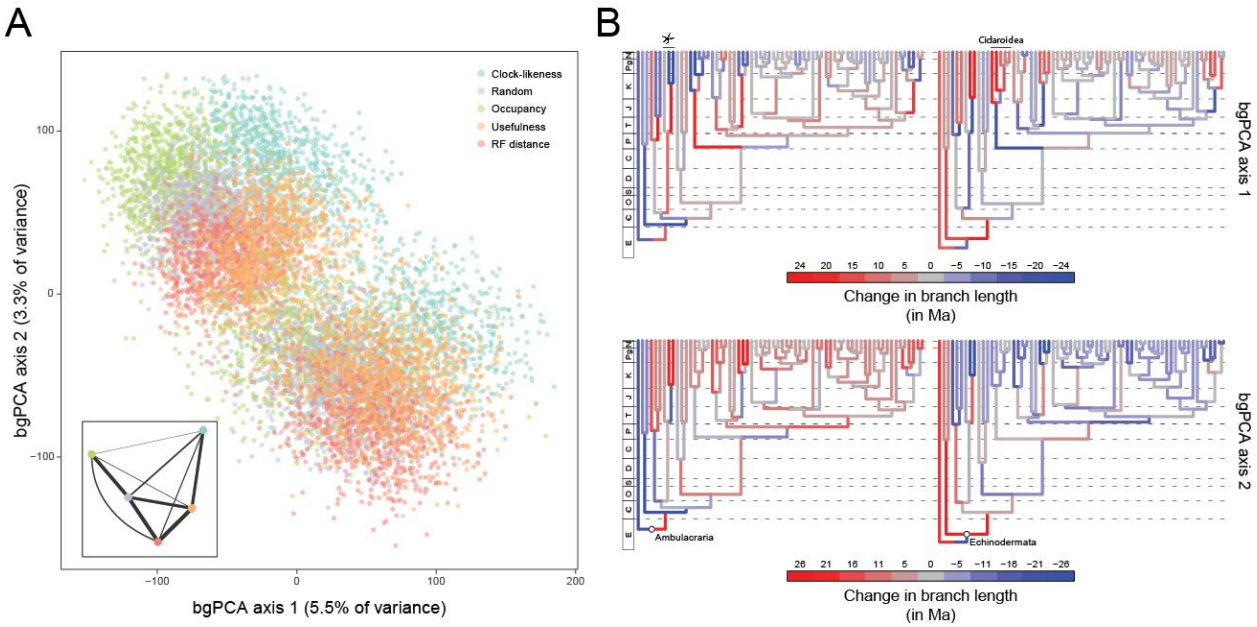


Figure S5: Effect of gene sampling strategy on inferred times of divergence. **A.** A chronospace, depicting the differences in inferred ages between chronograms obtained using different loci. Axes constitute the first two dimensions of the bgPCA (capturing 8.8% out of a total of 10.5% of variances summarized by all four dimensions). The inset shows the centroid for each dataset, and the width of the lines connecting them are scaled to the inverse of the Euclidean distances that separates them (as a visual summary of overall similarity). The ages most different to those obtained under random loci sampling were found when selecting clock-like genes. **B.** Change in branch lengths along the first two bgPC axes. Branches are colored according to their contraction/expansion given a score of – 1 (top) or + 1 (bottom) standard deviations, relative to the mean. Clades/nodes showing age differences larger than 20 Ma are labelled.

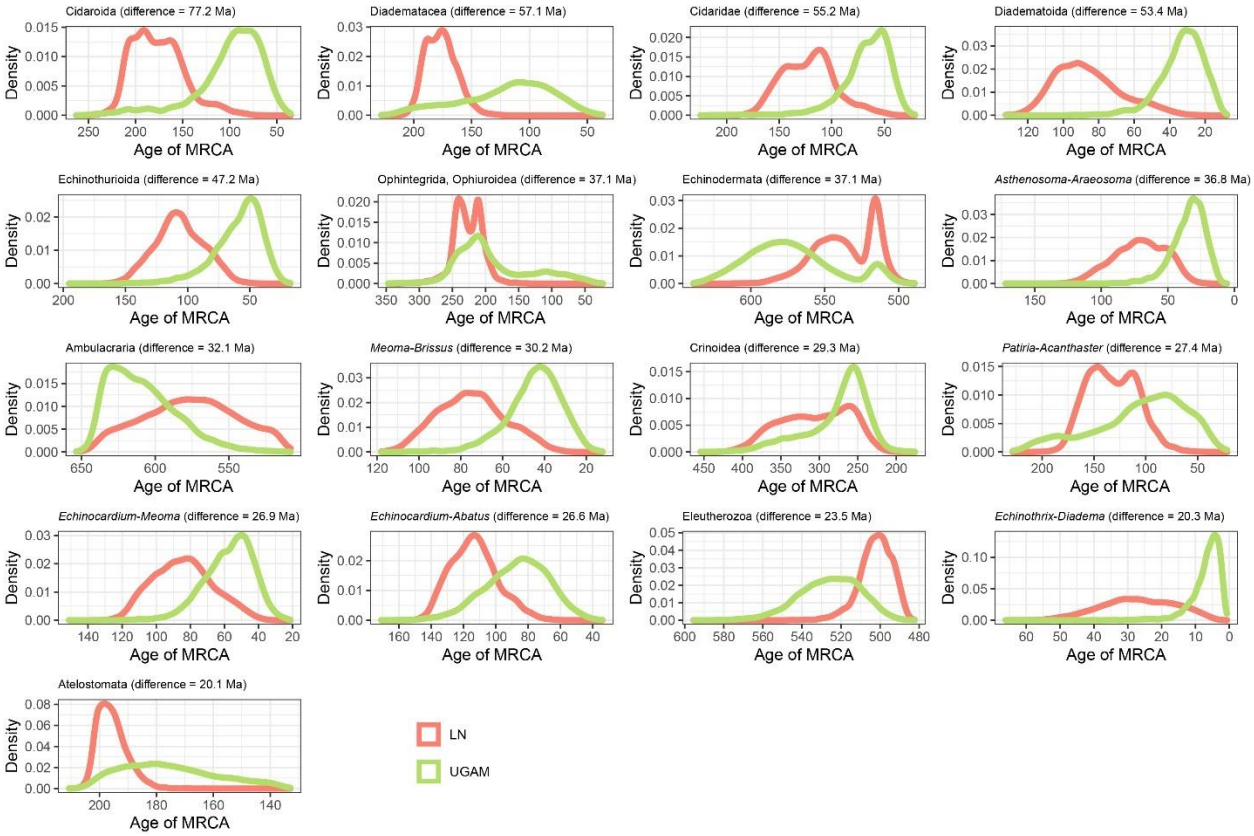


Figure S6: Distribution of posterior probabilities for node ages that show an average difference larger than 20 Ma depending on the choice of clock prior.

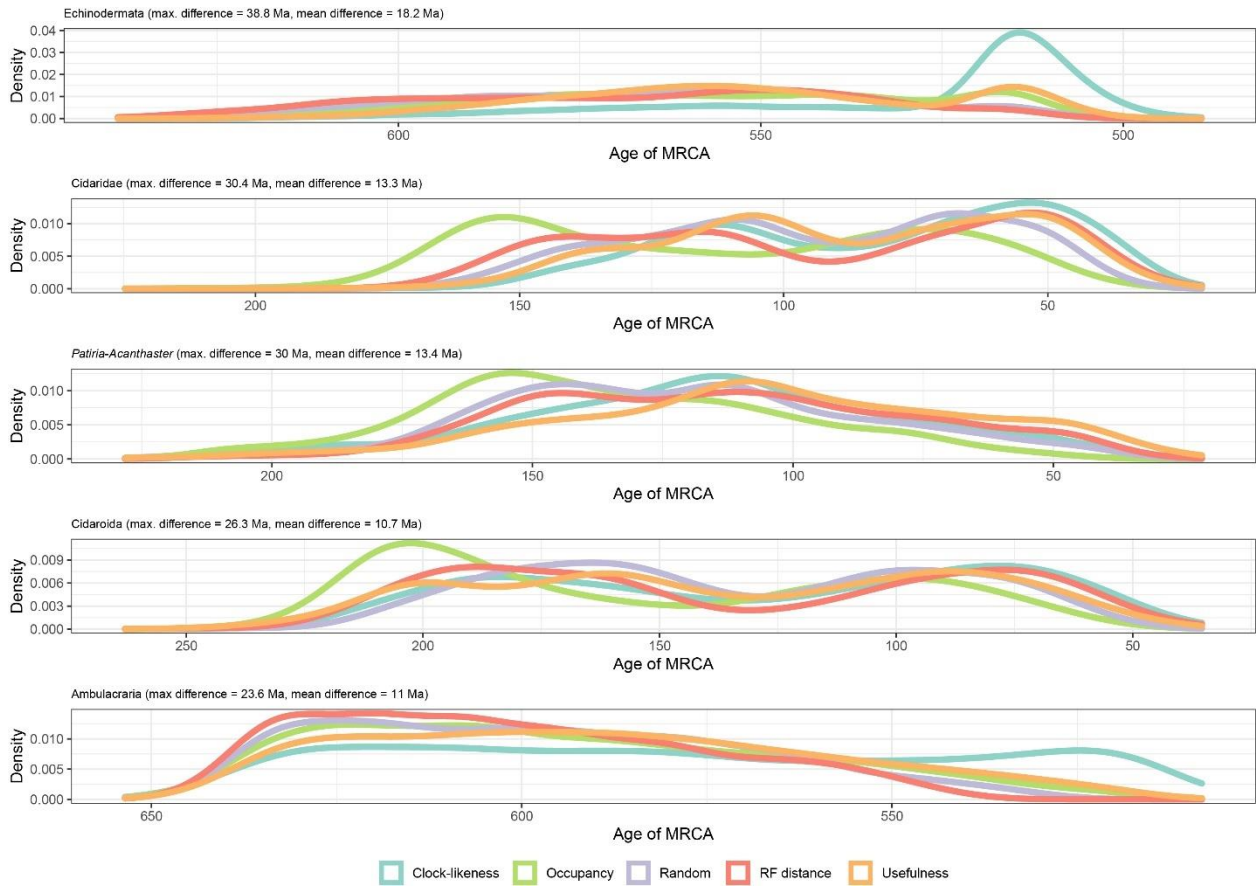


Figure S7: Distribution of posterior probabilities for node ages that show a maximum difference larger than 20 Ma depending on the choice of clock prior (no node showed average differences larger than 20 Ma). The largest differences can be seen in the relatively younger ages of Ambulacraria and Echinodermata when using clock-like genes, and in the relatively older ages for some nodes within cidaroids and starfish when using loci with high occupancy. Other sampling criteria largely agree on inferred node ages, as can also be seen in figure S5 as short distances between their centroids in chronospace.

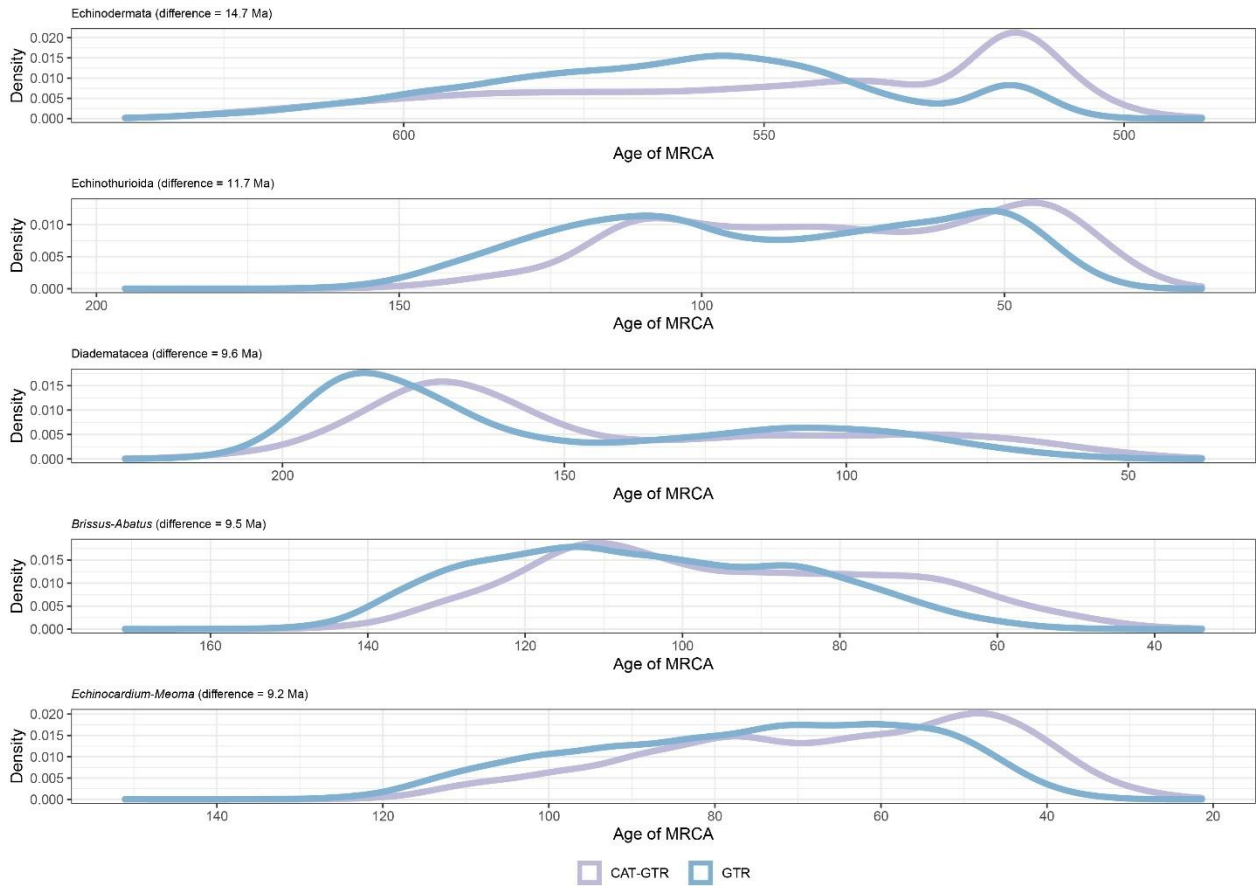
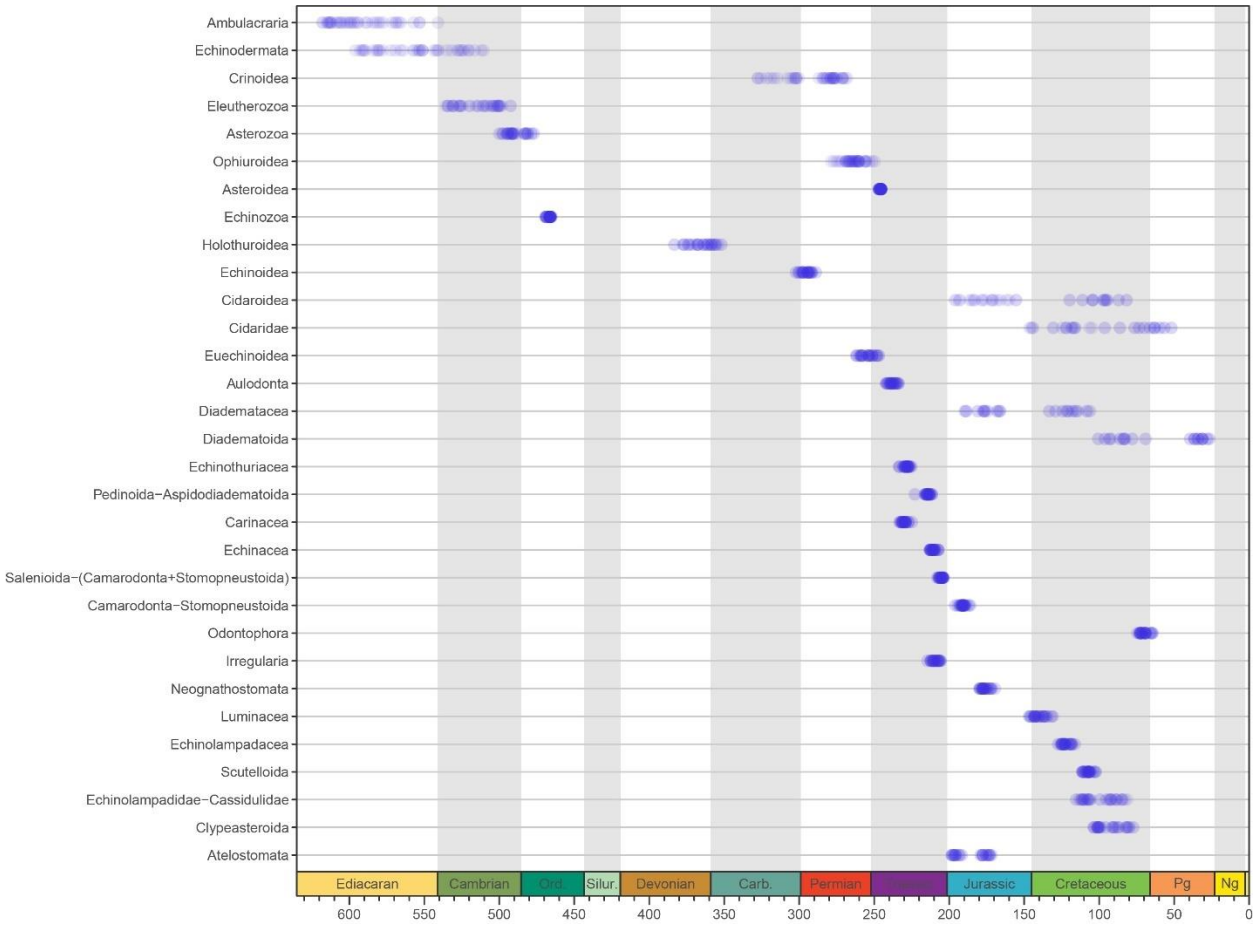


Figure S8: Distribution of posterior probabilities for node ages that are the most affected by the choice of model of molecular evolution. No node showed average differences larger than 20 Ma, so those with the largest effect are shown.



1045

1046 **Figure S9:** Median ages for selected clades across the consensus trees of the 40 time-calibrated runs

1047 performed.

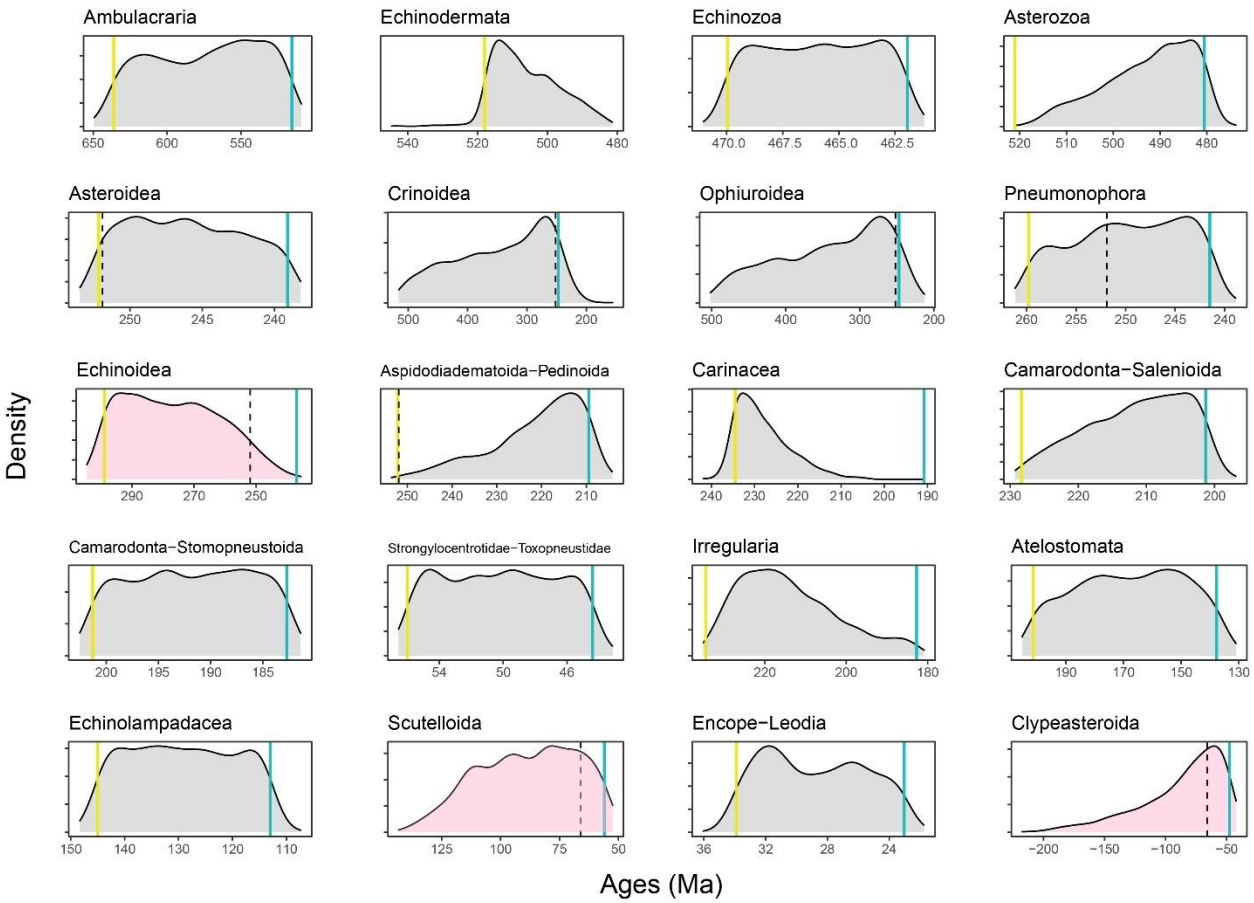


Figure S10: Prior distributions of all constrained nodes. Five hundred replicates were sampled from the joint prior, showing appropriately broad distributions of node ages. Blue lines show minima and yellow ones maxima (when enforced); dotted lines show the age of the Permian-Triassic (251.9 Ma) and Cretaceous-Paleogene (66 Ma) mass extinction events. Nodes whose ages are of special interest (Echinoidea, Scutelloida, and Clypeasteroida) are shown in pink, revealing large prior probabilities of the divergence occurring at either side of mass extinction events.

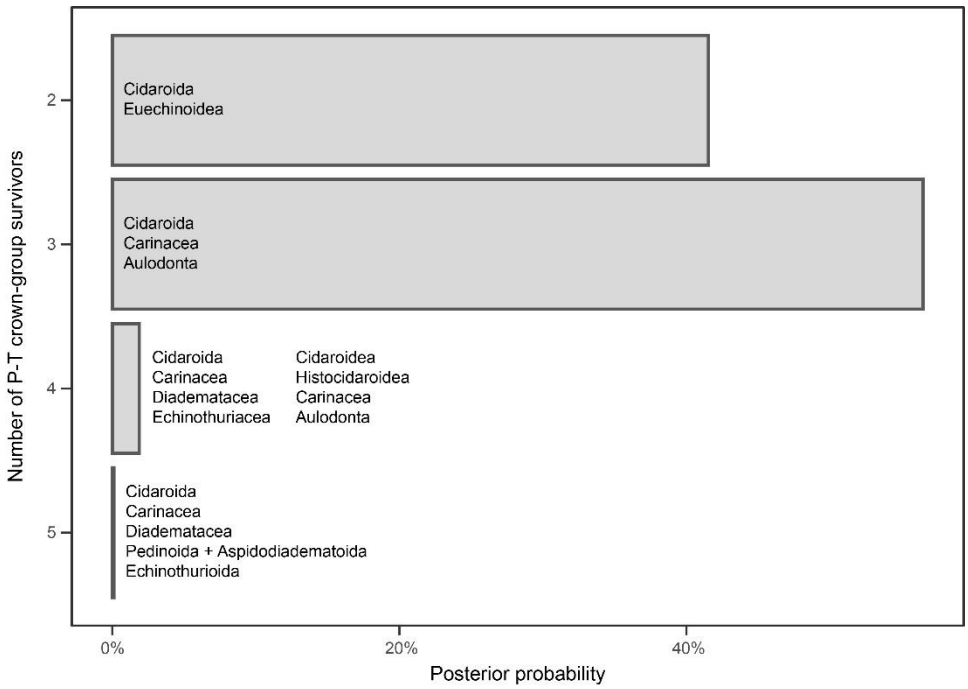


Figure S11: Number of lineages inferred to have crossed the Permian-Triassic (P-T) boundary. The probabilities of each scenario are estimated from the inferred divergence times of the 10,000 chronograms sampled across all of the analyses performed (250 for the two runs of each combination of sampled loci, model of molecular evolution, and clock prior). The probability of three or more crown group lineages surviving the Permian-Triassic extinction is 58.47%.

1061 **Table S1:** Transcriptomic/genomic datasets added in this study relative to the taxon sampling of
1062 Mongiardino Koch et al. [4] and Mongiardino Koch & Thompson [3]. Taxa with citations were taken from
1063 the literature, and details can be found in the corresponding studies and associated NCBI BioProjects.
1064 Taxa are shown in alphabetical order; those identified with "OG" are outgroup taxa (i.e., non-echinoids).
1065 Voucher specimens are deposited at the Benthic Invertebrate Collection, Scripps Institution of
1066 Oceanography (SIO-BIC), and the Echinoderm Collection, Muséum National d'Histoire Naturelle (MNHN-
1067 IE). If multiple accession numbers are shown for a given taxon, these datasets were merged before
1068 assembly. Similar details for all other specimens can be found in [3, 4].

Taxon	**Citation**	**Tissue type**	**Collector**	**Locality (depth)**	**Voucher**	**GenBank accession number**
Amphiura filiformis (OG)	[126]	-	-	-	-	SRX2255774
Apatopygus recens	This study	Mixed	Owen Anderson	Foveaux Strait, South Island, New Zealand	SIO-BIC E7142	
Aspidodiadema hawaiiense	This study	Mixed	Greg Rouse, Avery Hatch	Seamount 4, Costa Rica, Pacific Ocean (1,003 m)	SIO-BIC E7363	
Asterias rubens (OG)	[127]	-	-	-	-	SRX445860
Astrophyton muricatum (OG)	[128, 129]	-	-	-	-	SRX1391908
Bathysalenia phoinissa	This study	Tube feet, spine muscles	BIOMAGLO Cruise Team	Mozambique Channel, Mayotte (295-336 m)	MNHN-IE-2016-23	

Clypeaster japonicus	This study	Eggs	Frances Armstrong	Misaki Marine Biological Station, Kanagawa, Japan	-	
Echinothrix calamaris	This study	Male gonad	-	Phillipines	-	
Encope emarginata	This study	Eggs	Gulf Specimen Marine Lab	Apalachee Bay, Florida, USA	-	
Eucidaris metularia	This study	Mixed	Greg Rouse	Al-Fahal Reef, Red Sea, Makkah, Saudi Arabia	SIO-BIC 2017-008	
Fellaster zelandiae	This study	Spines, tube feet, gut, joining tissue of Aristotle's lantern	Wilma Blom	Western end of Cornwallis Beach, Auckland, New Zealand	SIO-BIC E7920	
Histocidaris variabilis	This study	Gonad	Greg Rouse, Avery Hatch	Las Gemelas Seamount, near Isla del Coco, Costa Rica, Pacific Ocean (571 m)	SIO-BIC E7350	
Lamberticrinus messingi (OG)	[129, 130]	-	-	-	-	SRX1397823
Leodia sexiesperforata	This study	Eggs	Frances Armstrong	Bocas del Toro, Panama	-	
Metacrinus rotundus (OG)	[131]	-	-	-	-	DRX021520
Micropyga tuberculata	This study	Tube feet, spine muscles	BIOMAGLO Cruise Team	Mozambique Channel, Îles Glorieuses (385-410 m)	MNHN-IE-2016-39	

Ophiocoma echinata (OG)	[127]	-	-	-	-	SRX445856
Peronella japonica	This study	Eggs	Frances Armstrong	Kanazawa, Japan	-	
Rhyncholampas pacificus	This study	Mixed	Carlos A. Conejeros-Vargas	Panteón Beach, Puerto Ángel Bay, Oaxaca, Mexico	SIO-BIC E7918	
Scaphechinus mirabilis	[132]	-	-	-	-	DRX187887 DRX187888
Sinaechinocyamus mai	This study	Gregory's diverticulum	Jih-Pai Lin	Miaoli County, Taiwan	SIO-BIC E7919	
Sterechinus neumayeri	[133]	-	-	-	-	ERX3498697 ERX3498698 ERX3498699 ERX3498700 ERX3498701
Stereocidaris nascaensis	This study	Muscle surrounding Aristotle's lantern	Charlotte Seid	Off San Ambrosio, Desventuradas Islands, Chile (215 m)	SIO-BIC E7154	
Tetrapygus niger	This study	Female gonad, tube feet	Felipe Aguilera	Dichato Bay, Chile	-	
Tripneustes gratilla	This study	Pedicellaria	-	Phillipines	-	

Tromikosoma hispidum	This study	Tube feet	Lisa Levin, Todd Litke	Quepos Plateau, Costa Rica, Pacific Ocean (2067 m)	SIO-BIC E7251

1069

1070 **Table S2:** Details of molecular datasets and supermatrix. Statistics for raw reads and assemblies are
1071 shown for datasets incorporated in this study relative to the sampling of [3, 4] (where similar statistics
1072 can be found for the other datasets). Taxa are shown in alphabetical order; those identified with "OG"
1073 are outgroup taxa (i.e., non-echinoids). Novel datasets correspond to Illumina pair-end transcriptomes,
1074 except for two draft genomes (*Bathysalenia phoinissa* and *Micropyga tuberculata*). Mean insert size is
1075 expressed in number of base pairs. For transcriptomes, read pairs (initial) shows numbers input into
1076 Agalma [37], (i.e., after processing with Trimmomatic [36] or UrQt [39]), while read pairs (retained)
1077 shows those that passed the Agalma curation checks (including ribosomal, adapter, quality, and
1078 compositional filters), and represent the final number of read pairs used for assembly. For genomes, see
1079 information in the bioinformatic details above. Assemblies were further sanitized with either
1080 alien_index [45] alone or in combination with CroCo [44] (see bioinformatic details), and the number of
1081 assembled transcripts/contigs shows the size of datasets after these curation steps. The number of loci
1082 shows the occupancy of terminals in the supermatrix output by Agalma (1,346 loci at 70% occupancy),
1083 after which loci were further removed by TreeShrink [50], resulting in the final occupancy scores.

Species	Mean insert size	Read pairs (initial)	Read pairs (retained)	Assembled transcripts/contigs	Number of loci	Removed by TreeShrink	Final occupancy
Abatus agassizii	-	-	-	-	522	5	38.4
Abatus cordatus	-	-	-	-	497	2	36.8
Acanthaster planci (OG)	-	-	-	-	864	8	63.6
Amphiura filiformis (OG)	180.1	61,558,173	52,206,727	416,946	761	8	55.9
Apatopygus recens	415.3	74,315,687	56,898,850	274,380	1152	5	85.2
Apostichopus japonicus (OG)	-	-	-	-	849	9	62.4
Araeosoma leptaleum	-	-	-	-	1184	2	87.8
Arbacia lixula	-	-	-	-	1234	1	91.6
Aspidodiadema hawaiiense	228.7	109,716,219	104,518,714	311,032	1060	11	77.9
Asterias rubens (OG)	230.4	31,890,613	25,495,009	103,090	805	9	59.1
Asthenosoma varium	-	-	-	-	1254	7	92.6

Astrophyton muricatum (OG)	195.9	25,191,954	22,281,829	149,146	478	5	35.1
Bathysalenia phoinissa	-	-	-	154,120	605	6	44.5
Brissus obesus	-	-	-	-	773	0	57.4
Caenopedina hawaiiensis	-	-	-	-	1094	2	81.1
Clypeaster japonicus	198.6	10,505,520	9,298,316	74,743	829	1	61.5
Clypeaster rosaceus	-	-	-	-	1242	2	92.1
Clypeaster subdepressus	-	-	-	-	1233	3	91.4
Colobocentrotus atratus	-	-	-	-	1158	0	86.0
Conolampas sigsbei	-	-	-	-	1001	5	74.0
Cystechinus c.f. *giganteus*	-	-	-	-	661	0	49.1
Dendraster excentricus	-	-	-	-	337	1	25.0
Diadema setosum	-	-	-	-	305	1	22.6
Echinarachnius parma	-	-	-	-	1187	3	88.0
Echinocardium cordatum	-	-	-	-	1012	2	75.0
Echinocardium mediterraneum	-	-	-	-	1185	1	88.0
Echinocyamus crispus	-	-	-	-	709	6	52.2
Echinometra mathaei	-	-	-	-	1055	0	78.4
Echinothrix calamaris	203.9	68,871,087	60,443,904	251,788	1252	4	92.7
Encope emarginata	206.2	12,015,888	10,461,870	66,241	1076	1	79.9
Eucidaris metularia	321.6	38,802,159	29,047,401	83,318	412	2	30.5
Eucidaris tribuloides	-	-	-	-	414	0	30.8
Evechinus chloroticus	-	-	-	-	1196	1	88.8
Fellaster zelandiae	313.9	62,619,791	51,742,748	168,598	1269	1	94.2
Heliocidaris erythrogramma	-	-	-	-	931	0	69.2
Histocidaris variabilis	329.2	84,103,672	73,884,180	144,044	1189	5	88.0
Holothuria forskali (OG)	-	-	-	-	743	8	54.6
Lamberticrinus messingi (OG)	195.0	22,989,820	20,234,597	59,049	351	3	25.9
Leodia sexiesperforata	203.8	16,211,182	14,174,475	68,964	1064	1	79.0
Leptosynapta clarki (OG)	-	-	-	-	403	4	29.6
Lissodiadema lorioli	-	-	-	-	770	0	57.2
Loxechinus albus	-	-	-	-	1265	1	93.9

Lytechinus variegatus	-	-	-	-	1257	2	93.2
Mellita tenuis	-	-	-	-	1002	1	74.4
Meoma ventricosa	-	-	-	-	1053	0	78.2
Mesocentrotus nudus	-	-	-	-	1247	0	92.6
Metacrinus rotundus (OG)	198.5	23,791,832	20,900,345	83,718	642	7	47.2
Micropyga tuberculata	-	-	-	170,514	415	4	30.5
Ophiocoma echinata (OG)	278.5	28,427,026	24,836,025	130,153	712	7	52.4
Paracentrotus lividus	-	-	-	-	1266	0	94.1
Patiria miniata (OG)	-	-	-	-	740	8	54.4
Peronella japonica	208.3	17,696,287	15,707,931	106,110	1043	6	77.0
Prionocidaris baculosa	-	-	-	-	794	3	58.8
Psammechinus miliaris	-	-	-	-	1173	1	87.1
Rhyncholampas pacificus	346.8	63,116,413	52,160,741	234,910	1070	2	79.3
Saccoglossus kowalevskii (OG)	-	-	-	-	668	6	49.2
Scaphechinus mirabilis	401.7	10,169,195	8,796,067	127,664	974	4	72.1
Sinaechinocyamus mai	297.2	60,208,172	52,265,201	164,646	1233	1	91.5
Sphaerechinus granularis	-	-	-	-	1214	1	90.1
Sterechinus neumayeri	207.3	26,376,279	21,308,840	122,001	1097	1	81.4
Stereocidaris nascaensis	327.0	69,962,495	60,316,050	121,508	1200	6	88.7
Stomopneustes variolaris	-	-	-	-	908	0	67.5
Strongylocentrotus purpuratus	-	-	-	-	1266	5	93.7
Tetrapygus niger	202.5	63,040,541	52,761,671	163,084	1287	2	95.5
Tripneustes gratilla	154.6	62,962,239	57,015,692	130,376	1309	1	97.2
Tromikosoma hispidum	300.4	57,208,357	47,909,038	234,502	1238	7	91.5

1084

1085 **Supplementary References (references 1-100 in main text)**

1086 [101] Faircloth BC, Glenn TC. 2012 Not all sequence tags are created equal: designing and validating
1087 sequence identification tags robust to indels. *PLoS One* **7**, e42543.

1088 [102] Corbari L, Samadi S, Olu K. 2017 BIOMAGLO cruise, RV Antea. (doi: doi.org/10.17600/17004000)

1089 [103] Bronstein O, Kroh A, Miskelly AD, Smith SD, Dworjanyn SA, Mos B, Byrne M. 2019 Implications of
1090 range overlap in the commercially important pan-tropical sea urchin genus *Tripneustes* (Echinoidea:
1091 Toxopneustidae). *Mar. Biol.* **166**, 34.

1092 [104] Seppey M, Manni M, Zdobnov EM. 2019 BUSCO: assessing genome assembly and annotation
1093 completeness. Methods Mol. Biol. **1962**, 227–245.

1094 [105] Kriventseva EV, Tegenfeldt F, Petty TJ, Waterhouse RM, Simao FA, Pozdnyakov IA, Ioannidis P,
1095 Zdobnov EM. 2015 OrthoDB v8: update of the hierarchical catalog of orthologs and the underlying free
1096 software. *Nucleic Acids Res.* **43**, 250–256.

1097 [106] Benton MJ, Donoghue PC, Asher RJ, Friedman M, Near TJ, Vinther J. 2015 Constraints on the
1098 timescale of animal evolutionary history. *Palaeontol. Electron.* **18**, 1–106.

1099 [107] Erkenbrack EM, Thompson JR. 2019 Cell type phylogenetics informs the evolutionary origin of
1100 echinoderm larval skeletogenic cell identity. *Comm. Biol.* **2**, 160.

1101 [108] Cooper R, Sadler P, Hammer O, Gradstein F. 2012 The Ordovician Period. In: *The geologic time*
1102 *scale* (eds. FM Gradstein, JG Ogg, M. Schmitz, G. Ogg), pp. 489–523. Elsevier.

1103 [109] Hunter AW, Ortega-Hernández J. 2021 A new somasteroid from the Fezouata Lagerstätte in
1104 Morocco and the Early Ordovician origin of Asterozoa. *Biol. Lett.* **17**, 20200809.

[110] Villier L, Brayard A, Bylund KG, Jenks JF, Escarguel G, Olivier N, Stephen DA, Vennin E, Fara E. 2018 *Superstesaster promissor* gen. et sp. nov., a new starfish (Echinodermata, Asteroidea) from the Early Triassic of Utah, USA, filling a major gap in the phylogeny of asteroids. *J. Syst. Palaeontol.* **16**, 395–415.

[111] Saucède T, Vennin E, Fara E, Olivier N, Team TPB. 2019 A new holocrinid (Articulata) from the Paris Biota (Bear Lake County, Idaho, USA) highlights the high diversity of Early Triassic crinoids. *Geobios* **54**, 45–53.

[112] Thuy B, Escarguel G. 2019 A new brittle star (Ophiuroidea: Ophiodermatina) from the Early Triassic Paris Biota (Bear Lake County, Idaho, USA). *Geobios* **54**, 55–61.

[113] Kroh A. 2011 Echinoids from the Triassic of St. Cassian-A review. *Geo. Alp.* **8**, 136–140.

[114] Li Z, Chen Z-Q, Zhang F, Ogg JG, Zhao L. 2020 Global carbon cycle perturbations triggered by volatile volcanism and ecosystem responses during the Carnian Pluvial Episode (late Triassic). *Earth-Sci. Rev.* **211**, 103404.

[115] Smith AB. 2011 *Gymnodiadema* and the Jurassic roots of the Arbacioida (stirodont echinoids). *Swiss J. Palaeontol.* **130**, 155–171.

[116] François É, Marchand D, David B. 2003 Fluctuations morphologiques et hétérochronies chez *Toxaster* (échinides, Crétacé inférieur). *C. R. Palevol* **2**, 597–605.

[117] Cooke CW. 1955 Some Cretaceous echinoids from the Americas. *US Geol. Surv. Prof. Pap.* **264**, 87–112.

[118] Roman J. 1989 L'ancêtre éocène des Scutellidae (Echinoidea; Clypeasteroida). In: *Echinoderm Research. Proceedings of the Second European Echinoderm Conference, Brussels* (ed. C De Ridder), pp. 18–21. Rotterdam, AA Balkema.

[119] Durham JW. 1994 Fossil *Encope* (Echinoidea) from the Pacific coast of southern Mexico. *Rev. Mex. Cienc. Geol.* **11**, 113–116.

[120] Coppard SE, Lessios HA. 2017 Phylogeography of the sand dollar genus *Encope*: implications regarding the Central American Isthmus and rates of molecular evolution. *Sci. Rep.* **7**, 11520.

[121] Ali MSM. 1983 The paleogeographic distribution of *Clypeaster* (Echinoidea) during the Cenozoic Era. *Neues Jahrb. Geol. Paläontol.* **8**, 449–464.

[122] Via L, Padreny J. 1970 Dos nuevas especies de *Clypeaster* del Eoceno de Cataluña. *Publ. Inst. Invest. Geol. Diput. Prov.* **24**, 89–118.

[123] Wade BS, Pearson PN, Berggren WA, Pälike H. 2011 Review and revision of Cenozoic tropical planktonic foraminiferal biostratigraphy and calibration to the geomagnetic polarity and astronomical time scale. *Earth-Sci. Rev.* **104**, 111–142.

[124] Gold DP, Fenton JP, Casas-Gallego M, Novak V, Pérez-Rodríguez I, Cetean C, Price R, Nembhard N, Thompson H. 2018 The biostratigraphic record of Cretaceous to Paleogene tectono-eustatic relative sea-level change in Jamaica. *J. S. Am. Earth Sci.* **86**, 140–161.

[125] Smith SA, Brown JW, Walker JF. 2018 So many genes, so little time: A practical approach to divergence-time estimation in the genomic era. *PLoS One* **13**, e0197433.

[126] Dylus DV, Czarkwiani A, Stångberg J, Ortega-Martinez O, Dupont S, Oliveri P. 2016 Large-scale gene expression study in the ophiuroid Amphiura filiformis provides insights into evolution of gene regulatory networks. *Evodevo* **7**, 1–17.

[127] Reich A, Dunn CW, Akasaka K, Wessel, G. 2015 Phylogenomic analyses of Echinodermata support the sister groups of Asterozoa and Echinozoa. *PLoS One* **10**, e0119627.

1147 [128] Linchangco Jr GV, Foltz DW, Reid R, Williams J, Nodzak C, Kerr AM, Miller AK, Hunter R, Wilson NG,
1148 Nielsen WJ, et al. 2017 The phylogeny of extant starfish (Asteroidea: Echinodermata) including *Xyloplax*,
1149 based on comparative transcriptomics. *Mol. Phylogenet. Evol* **115**, 161–170.

1150 [129] Janies DA, Witter Z, Linchangco GV, Foltz DW, Miller AK, Kerr AM, Jay J, Reid RW, Wray GA. 2016
1151 EchinoDB, an application for comparative transcriptomics of deeply-sampled clades of echinoderms.
1152 *BMC Bioinformatics* **17**, 17–48.

1153 [130] Clouse RM, Linchangco Jr GV, Kerr AM, Reid RW, Janies DA. 2015 Phylotranscriptomic analysis
1154 uncovers a wealth of tissue inhibitor of metalloproteinases variants in echinoderms. *R. Soc. Open Sci.* **2**,
1155 150377.

1156 [131] Koga H, Fujitani H, Morino Y, Miyamoto N, Tsuchimoto J, Shibata TF, Nozawa M, Shigenobu S,
1157 Ogura A, Tachibana K, et al. 2016 Experimental approach reveals the role of *alx1* in the evolution of the
1158 echinoderm larval skeleton. *PLoS One* **11**, e0149067.

1159 [132] Simakov O, Kawashima T, Marlétaz F, Jenkins J, Koyanagi R, Mitros T, Hisata K, Bredeson J,
1160 Shoguchi E, Gyoja F, et al. 2015 Hemichordate genomes and deuterostome origins. *Nature* **527**, 459–
1161 465.

1162 [133] Collins M, Peck LS, Clark MS. 2021 Large within, and between, species differences in marine
1163 cellular responses: Unpredictability in a changing environment. *Sci. Total Environ.* **794**, 148594.